Assessing inter-organ crosstalk in the development of obesity-associated non-alcoholic fatty liver disease

Irene Vañó Segarra

Bacherlor's Degree Final Project Biochemistry and Molecular Biology

Directed by Dr. Helena Torrell, Prof. Jorge Joven, Helena Castañé

UNIVERSITAT ROVIRA I VIRGILI

Tarragona, June 2022

Treball realitzat a partir dels resultats obtinguts en les Pràctiques Externes realitzades al Grup de Recerca Biomèdica, sota la tutorització del Dr Jorge Joven i Helena Castañé.

CONTENT

ABREVIATIONS:

AB	antibody
ACE-ARBS	angiotensin convertin enzyme and angiotensin receptor blockers
ALT	alanine aminotransferase
ANOVA	analysis of variance
AST	aspartate aminotransferase
AT	adipose tissue
AUC	area under the curve
BA	bile acids
BAT	brown adipose tissue
BMI	body max index
ChREBP	carbohydrate response element binding protein
CPAP	continuous positive airway pressure
CVD	cardiovascular disease
DAP	diastolic arterial pressure
DLP	dyslipdemia
ELISA	enzyme-linked immunosorbent assay
FFA	free fatty acids
FGF	fibroblast growth factor
GAL 3	galectina 3
GGT	gammaglutamyl transferase
HCC	hepatocellular carcinoma
HDL	high density lipoprotein
HT	hypertension
IL-6	interleukin-6
IQR	inter quartil range
IR	insuline resistence
KNN	k-nearest neighbors
LDL	low density lipoprotein
MS	metabolic syndrome
NAFLD	non-alcholic fatty liver disease
NAS	NAFLD activity score
NASH	non-alcoholic steatohepatitits
NF-κB	nuclear factor kappa-light-chain-enhancer of activated B cells
OSA	obstructive sleep apnea
PLS-DA	partil least square discriminant analysis
ROC	receiver operatin characteristics
SAT	subcutaneous adipose tissue
SREBP	sterol regulatory element-binding protein
SREBP1c	sterol regulatory element binding protein 1c
T2DM	diabetes mellitus
TG	triglycerides
TNF-α	tumor necrosis factor-α
VAT	visceral adipose tissue
VIP	variable importance in projection
WHO	world health organitzation

ABSTRACT:

Backgroud: Non-alcoholic fatty liver disease (NAFLD) is closely related to obesity, and both diseases have become so prevalent over time that they are considered a global pandemic. NAFLD begins with simple steatosis and, if not treated, it turns more severe reaching non-alcoholic steatohepatitis (NASH) in which inflammation and ballooning are also present. Nowadays, the gold standard technique for the diagnosis of NAFLD/NASH progression, is liver biopsy. During the development and progression of this disease, the organism tries to modulate the metabolic and physiologic status of the organs. These organs exchange information using different compounds such as cytokines, commonly named as organokines. Nevertheless, there is also a concern about the complexity of inter-tissue communication through these molecules.

Objective: To determine the concentration of six organokines in plasma and assess their role in the obesity-associated NAFLD progression.

Methods: Plasma samples from n=1068 patients with obesity and different stages of NAFLD and n=414 healthy volunteers were used to perform sandwich Enzyme-Linked ImmunoSorbent Assay (ELISA) to determine the concentration of Adiponectin, Leptin, Irisin, Galectin3, Fibroblast growth factor 19 (FGF19) and Fibroblast growth factor 21 (FGF21).

Results: We found significative differences in all analysed organokines profile between healthy volunteers and obesity-associated NAFL patients. However, when considering the different profile of organokines along NAFLD severity, we only observed significative differences in Adiponectin and Leptin. When studying the ability of the organokines to differ between the NASH, uncertain NASH and non-NASH groups, only Adiponectin and FGF21 had this capability. Adiponectin, Leptin and FGF21 showed significative differences when classifying the patients according to liver histological characteristics, however they did not have enough power to serve as biomarkers of fibrosis.

Conclusions: Significant differences where found between healthy volunteers and NAFL patients regarding analyzed organokines' profile. However, only Adiponectin and FGF21 were found to be associated to NAFLD progression. Organokines' concentration were not significantly related to hepatic histological features.

Keywords: Non-alcoholic fatty liver disease, organokines, obesity, crosstalk

INTRODUCTION:

The liver is involved in essential biological functions that are important to maintain metabolic homeostasis. Alterations in the normal liver functioning can promote a high number of pathologies.

With the increasing prevalence of obesity and weight-related metabolic comorbidities, non-alcoholic fatty liver disease (NAFLD) has become the most usual liver disease worldwide. In fact, it has been currently recognized as the most prevalent chronic liver disease, and globally affects around 25% of the adult population. It is one the most prominent cause of liver related morbidity and mortality (1).

This disease can progress to non-alcoholic steatohepatitis (NASH). This progression, occurs to one third of the NAFLD population so, owing to its high prevalence and potential risk of progression to cirrhosis and hepatocellular carcinoma (HCC), has become a major health concern worldwide (2). Moreover, both diseases NAFLD and NASH are closely related to obesity, 51% of NAFLD and 82% of NASH patients suffer this disease. This fact is important to be considered due to the increasing proportion of overweight population.

1. NAFLD

NAFLD is the liver component of a cluster of conditions that are associated with metabolic dysfunction. It remains asymptomatic and it is defined by the presence of steatosis in more than 5% of hepatocytes (Figure 1) in association with metabolic risk factors such as obesity and type 2 diabetes, and it appears in the absence of excessive alcohol consumption (≥30 g per day for men and ≥20 g per day for women) or other chronic liver diseases.

This disease was first described in 1980 by Ludwing et al. after finding a group of patients that, without having a significant alcohol consumption, showed the same hepatic histopathological changes as those who had a liver disease associated with alcoholism. Although the leading causes of death in people with NAFLD are cardiovascular disease and extrahepatic malignancy, advanced liver fibrosis is a key prognostic marker for liver-related outcomes and overall mortality. Patients with cirrhosis should be screened for hepatocellular carcinoma and esophageal varices (3).

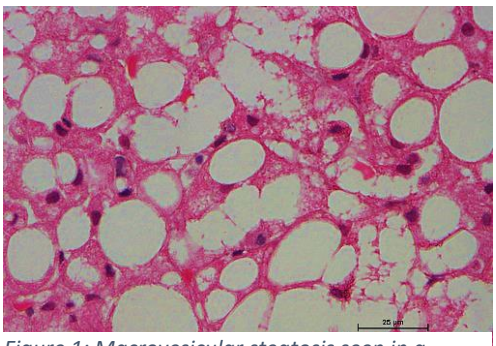

Figure 1: Macrovesicular steatosis seen in a sample stained with Hematoxilin-Eosin

The underlying mechanism for the development and progression of NAFLD is complex and multifactorial. One of the theories, discusses the role of communication between organs through cytokines specifically synthesized by them, namely organokines.

1.1. Prevalence

NAFLD has a global prevalence of 25%, which varies depending on ethnic, with the highest prevalence among Hispanics and lowest among Africans and Americans, whereas Caucasian and Asian ethnicities having an intermediate prevalence. In fact, in recent meta-analyisis, the prevalence of NAFLD was reported as follows: Asia, 27.4%; the Middle East, 31.8%; North America, 24.1%; South America, 30.5%; Europe, 23.7% and Africa, 13.5 (4). Because of its close association with the metabolic syndrome (MS), 63.7% of people with type 2 diabetes (T2DM) and up to 80% of people with obesity also suffer from NAFLD. However, some people with a normal body-mass index (eg, <21-25 kg/m^2) can still develop NAFLD, often described as non-obese or lean NAFLD. These patients usually have central obesity or other metabolic risk factors (3).

Besides, one third of the NAFLD population have developed NASH, which is also associated with different comorbidities. Table 1 shows how these diseases affect the prevalence of NAFLD or NASH in the population.

Table 1: Obesity, T2DM, Metabolic Syndrome and dyslipidemia (DLP) are risk factors to develop NAFLD and NASH. Specially, obesity is the one which has a more negative impact over both diseases.

Comorbidities	NAFLD	NASH
Obesity	51%	82%
T2DM	23%	47%
Metabolic syndrome and DLP	41%	71%

** NAFLD: nonalcoholic fatty liver disease; NASH: non-alcoholic steatohepatitits; T2DM: diabetes mielitus; DLP: dyslipidemia*

NAFLD affects differently depending on age and sex of the patients: it has been found that the prevalence of NAFLD increases with advancing age and it is more frequent in men (31% of prevalence) than in women (16%) (5).

1.2 NAFLD and obesity

Obesity is defined by the World Health Organization (WHO) as abnormal and/or excessive fat accumulation that may damage the health. Obesity is diagnosed by the body mass index (BMI), which is the result of dividing a person's weight in kilograms by the square of their height in meters. Specifically, an adult has overweight if their BMI is 25 kg/m^2 or over, and obesity if their BMI is 30 kg/m^2 or over. A BMI greater than 40 kg/m^2 is considered a sign of morbid obesity (6).

Although there are multiple factors associated with the cause of obesity, the most frequent cause is the overnutrition. This positive energy balance is accumulated in patient's bodies and transformed into deposits of triglycerides (TG) in their adipose tissue. This situation leads to adipocyte's growth in size (hypertrophy). Moreover, there is a differentiation of preadipocytes

into new adipocytes, a process named as adipogenesis. Subsequently, it leads to hyperplasia, which is an increase of the number of adipocytes, resulting in the dysfunction and eventual adipocyte death (7).

Obesity is associated with different comorbidities. As mentioned above, the adipose tissue hyperplasia and hypertrophy cause a dysfunctional tissue. Due to this, circulating free fatty acids (FFA) and adipokines increase. These FFAs are stored in the liver ectopically (and other tissues), giving rise to NAFLD. Not only does the amount of adipose tissue influence the patient, but its location, which can be classified into subcutaneous adipose tissue (SAT) or visceral adipose tissue (VAT). SAT is located under the skin and is the first place where fat accumulates, while VAT is surrounding the internal organs and only appears when SAT is completely full (8). In addition, patients with obesity may lead to insulin resistance (IR), T2DM and dyslipidemia (DLP). These comorbidities are closely linked to the NAFLD/NASH progression.

1.3 Spectrum of NAFLD/NASH

The progression of NAFLD is explained by the multiple hit hypothesis. As mentioned above, the NAFLD disease begins with fat accumulation in the liver, a condition often related to MS, DLP, hypertension and T2DM which leads to a steatosis accumulation of > 5% in hepatocytes. Evolution from simple steatosis to NASH occurs to 12-40% of population, and is the result from a complex interplay that involves either liver cell population (both parenchymal and nonparenchymal) and pathological signals coming from other organs, such as adipose tissue and the gut. NASH is characterized by inflammation and ballooning of hepatocytes, in addition to a more severe steatosis. Then, 15-25% of population reaches the worst stage of the spectrum, which are cirrhosis and, in the worst cases, (7% of population) develops hepatocellular carcinoma .The progression to cirrhosis is caused because of the replacement of the hepatocytes by scar tissue made of type I collagen in order to regenerate new cells (9). Patients suffering from these lates stages of NAFLD, have been shown to represent the most growing indicator for liver transplantation (Figure 2).

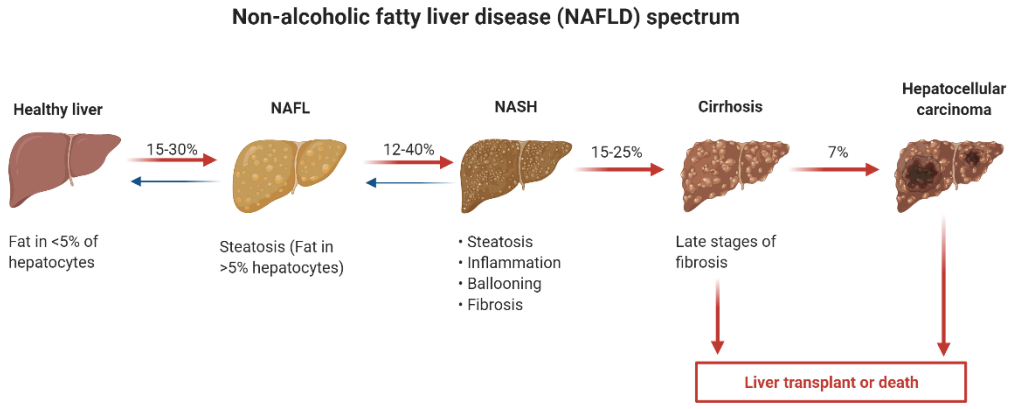

Figure 2: NAFLD progression can lead patients to death. Until NASH stage, the damage is reversible

1.4 Physiopathology

Different theories have emerged in order to understand the underlying mechanism for the development and progression of NAFLD. According to the traditional "two-hit" hypothesis, hepatic accumulation of lipids acts as the first hit, sensitizing the liver to a second hit that activates inflammatory cascades and fibrogenesis. However, a more accurate explanation of NAFLD pathogenesis contemplates that several molecular and metabolic changes take place synergistically in its development and progression. For this reason, it is more correct to explain NAFLD disease as a "multiple hit" hypothesis that considers multiple factors acting together to induce it. Such hits include IR and lipotoxicity which are one of the comorbidities related to obesity caused by overnutrition. This explains the great importance of the obesity impact in the development of this disease. But also, are important other factors such as the genetic background or the gut microbiota (10).

As mentioned above, the primary driver of NAFLD is the excess of lipid accumulation into the liver. This accumulation is closely related to an increased visceral adipose tissue lipolysis (59% of the cases), hepatic de novo lipogenesis activation (26% of cases) and high content of calories and/or fat in the diet (15% of cases) (11). The combination of these processes causes an expansion of adipose depots as well as accumulation of ectopic fat in the liver.

The fat accumulation leads to immune system infiltration into the visceral adipose tissue, creating a proinflammatory state that, in turn, promotes its insulin resistance (3).

The inability to respond to insulin by the target cells leads to hyperinsulinemia in the blood, which alters glucose metabolism and, consequently, the lipid synthesis and storage. Due to IR, adipose tissue becomes resistant to the antilipolytic effect of insulin, leading to TG breakdown and final formation of FFAs and glycerol (increased lipolysis) (Figure 3).

These FFAs released are taken by the liver where they will be accumulated as TG. This is known as *de novo* lipogenesis, and it is a crucial pathway that contributes to lipid storage. This process is under the control of 2 transcriptional factors, the sterol regulatory element binding protein 1c (SREBP1c) and the carbohydrate response element binding protein (ChREBP) (11). Under IR, this hormone activates lipogenesis by induction of sterol regulatory element-binding protein (SREBP) and fatty acid synthase, which all along cause an increase of FFAs (Figure 3).

Finally, the release of fatty acids, can also come from the dietary FFAs, absorbed in the gut and incorporated as TGs into chylomicrons. Chylomicrons can reach the liver to be accumulated (Figure 3). That is not the only role of gut related with NAFLD: lipid metabolism is also modulated by intestinal products that derive from bacteria fermentation of polysaccharides in the course of dysbiosis, and this disturbance in gut also plays an important role in progression of NAFLD/NASH by activating proinflammatory and profibrogenic intracellular pathways (11).

Such amount of FFAs released into to the liver and other peripheral tissues overwhelms its metabolic capacity and leads to a situation of lipotoxicity. This imbalance in lipid metabolism, leads to the formation of lipotoxic lipids which contributes to cellular stress in the hepatocyte.

As a result, inflammation, tissue regeneration, apoptosis and fibrogenesis are stimulated, not only in hepatocytes, but also in other tissues (3).

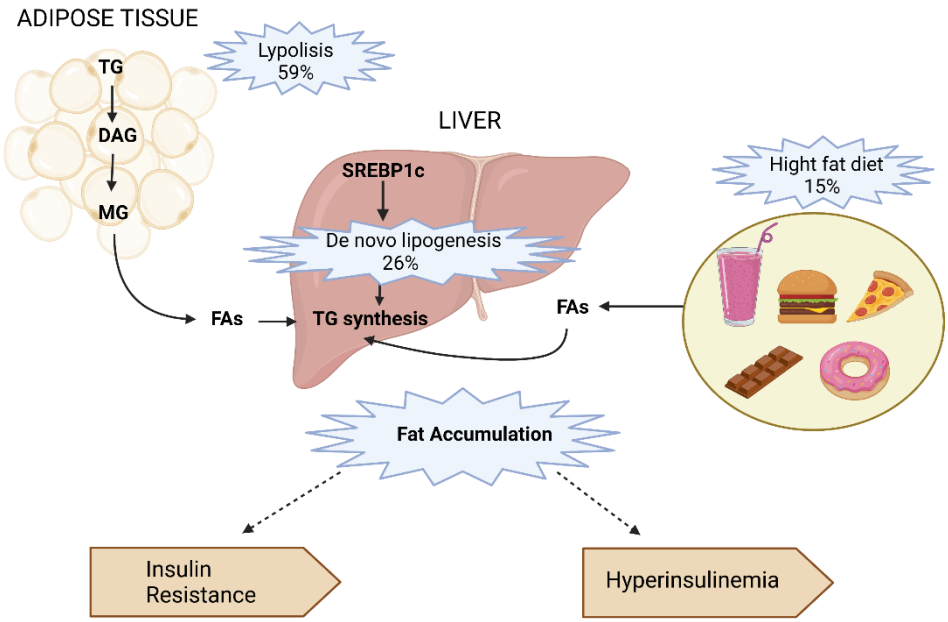

Figure 3 Three principal causes of fat accumulation (Lipolysis, de novo lipogenesis and high fat diet) and the consequences of these accumulation (insulin resistance and hyperinsulinemia). Modified from: (9)
**TG: triglycerides, DAG: diacylglycerides, MG: Monoglycerides, FAs: fatty acids*

1.5 Diagnostic

To avoid too much liver damage, it is crucial to know how NAFLD manifests and to establish methods to predict liver severity. The main difficulty resides in that most NAFLD patients remain asymptomatic. Nowadays, ultrasound offers 60-94% sensitivity and 66-97% specificity determining the presence of NAFLD. What this technique allows us to do is to establish if patients have or not steatosis.

However, ultrasound cannot be used in morbidly obese patients due to the visceral fat, and in lean individuals, it does not inform about the inflammation state of the liver. This is the reason why, to date, the gold standard technique for the diagnosis of NAFLD/NASH progression is liver biopsy. This is an invasive procedure that poses potential sampling errors and inconsistencies in histopathological interpretation (12). When the hepatic biopsy is done, each formalin-fixed paraffin-embedded sample is stained with Hematoxylin and Eosin and Masson's Trichrome. Then, through the study of the microscopic slices, it is observed if the sample presents certain histological characteristics that are part of a punctuation system. This punctuation system was described by Kleiner et al., in 2005 and it is known as NAS (NAFLD Activity Score) (4). It consists on evaluating the steatosis degree, necroinflammatory lesions, and ballooning from the samples and gives a score from 0 to 8 (Figures 4-6). This score allows to determine in which state of the disease the patient is. Samples rated 5 or over are diagnosed as NASH, while the ones rated less than 3 are labeled as non-NASH. Samples that score 3 or 4 are in an intermediate state, considered uncertain or borderline (Table 2).

These characteristics are: steatosis (0-3), ballooning (0-2) and lobular inflammation (0-3). The sum of all scores will result in the NAS score, i.e. from 0 to 8 (0-2 non-NASH, 3-4 Uncertain, 5-8 NASH) (Table 2).

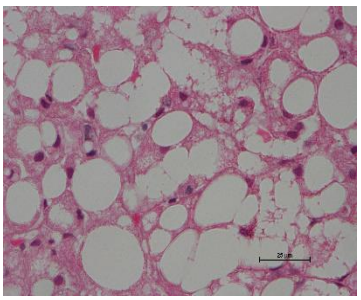

Figure 4 sample stained with Hematoxilin-Eosin where it is shown the steatosis

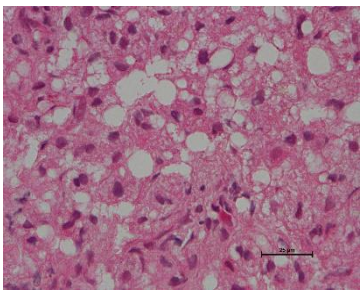

Figure 5 sample stained with Hematoxilin-Eosin where it is shown the balloning

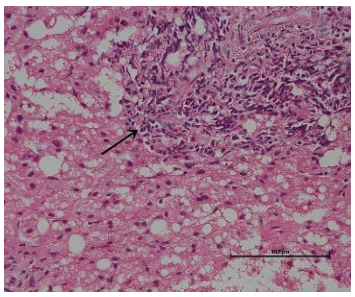

Figure 6 sample stained with Hematoxilin-Eosin where it is shown the lobular inflammation

Table 2: Kleiner et al. System of punctuation to determine if patients have NASH.

Steatosis	Score S	Lobula Inflammation	Score L	Hepatocytic Ballooning	Score B
<5%	0	None	0	None	0
5-33%	1	<2	1	A few balloon cells	1
34-66%	2	2-4	2	Many abalone cells	2
>66%	3	>4	3		

Σ

NASH
0-2 non-NASH
3,4 Uncertain NASH
<5 NASH

Although fibrosis is not included in the NAS score, its assessment also contributes to the understanding of the liver severity state from the patient (Table 3).

Table 3: System of punctuation to determine the fibrosis stage in NASH using adipose tissue samples.

Fibrosis stage in NASH	Stage
No fibrosis	0
Mild fibrosis, perisinusoidal zone 3	1
Moderate fibrosis, perisinusoidal zone 3	1
Portal/periportal fibrosis only	1
Perisinusoidal fibrosis zone 3 and portal/peri portal	2
Bidging fibrosis	3
Cirrhosis	4

2. Organokines

To modulate the metabolic and physiologic status of the organs, they exchange information using different compounds such as cytokines.

In the 1990s after the discovery of leptin as an adipocyte-secreted factor that regulates body fat stores, the concept of “metabolic cross talk” arose. It was seen that, these molecules were used as connectors, and they received the name of organokines (14).

Nevertheless, there is also a concern about the complexity of inter-tissue communication though these organokines. This is because there is a wide variability with regards to the types of molecules secreted, the mechanisms regulating their production and secretion, and the vesicles in which these molecules are transported (13). However, what seems to be clear is that the combined action of organokines is related to health or to the genesis of several diseases.

2.1 Organokines and NAFLD

It has been seen that adipose and hepatic tissues have an important endocrine function which seems to be influencing NAFLD (and also NASH) development. These tissues are able to produce different organokines (adipokines and hepatokines, respectively) which through autocrine, endocrine and paracrine pathways connect different organs.

As mentioned above, what generally happens during the progression of NAFLD is an expansion of adipose tissue, which generates lipotoxicity. This results on the release of FFAs, and due to this, the variety of immune cells there initiate a vicious cycle with the production and secretion of pro-inflammatory organokines (especially hepatokines) (13). These have crucial roles in the pathogenesis and progression of NAFLD, causing systemic inflammatory reactions and contributing to metabolic disorders such as IR, T2DM and MS (14).

The called “gut-liver axis” interaction seems to play a critical role in NAFLD onset and progression too. These refers to the link between the gut microbiome, the intestinal barrier integrity, the bile acid and the NAFLD (2). It is well known that gut plays a critical role in human physiology in terms of digestion, absorption of nutrients and the excretion of waste (8). Hence, the gut microbiota role in the pathophysiology of NASH is important (14).

On the one hand, some dietary compounds are metabolized by the gut microbiota and result in metabolites. Those microbiome-derived metabolites reach the liver though the portal system, being this organ considered the first line of defense of intestinal-derived toxins. These derived metabolites can affect the liver leading to interconnected effects such as liver inflammation and fibrosis, causing an altered secretion of organokines (especially hepatokines), and a disrupted croostalk between other organs.

On the other hand, diet can modify the microbiota composition and consequently exert beneficial or harmful effects on the host (14). On the basis, comparing individuals with NAFLD and non-NAFLD controls, both microbiome and bacterial density are different (8). Excessive ingestion of foods high in fat is associated with a significant depletion of bacterial species. Furthermore, diet regulates some organokines, such as FGF21, amplifying the process of a

vicious cycle that is also influencing the establishment of NASH and the secretion of hepatokines. As hepatokines participate in crosstalk, then indirectly, the metabolism of other organs/tissues such as the pancreas or adipose tissue (14).

However, not only those tissues named before are related to the NAFLD, but also skeletal muscle and bone tissue seems to be part of this cross talk too. These two tissues have a proximate relationship which goes beyond anatomy as they are also physiologically connected by the endocrine system through organokines (myokines and osteokines, respectively). The crosstalk between both tissues is shown during a resistance exercise. In this situation, the load applied to skeletal muscle is transferred to bone, which in addition to initiating muscle protein production, also signals a high energy requirement to facilitate bone formation. In that moment, the bone and muscle are a bidirectional pathway for biochemical signals (14).

An improved knowledge of the pathogenic cross-talk between the liver and extra-hepatic organs will not only help to modulate known risk factors associated with the onset of NAFLD and/or its progression to end-stage liver disease, but may also provide insight for the development of new pharmacological treatments for NAFLD (Figure 7) (15).

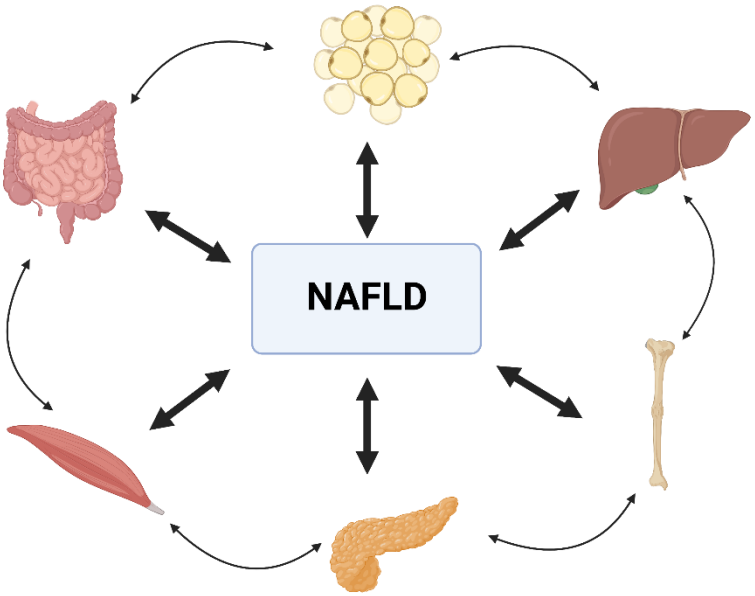

Figure 7: Cross talk between the adipose tissue, the liver, the bond, the pancreas, the muscle and the intestine in relation with the NAFLD.

2.1.1 Adiponectin

Adiponectin is the most abundant adipokine, produced and secreted mainly by white adipose tissue. This organokine is more expressed when the smallest is the adipose tissue mass and when it circulates at very high concentrations (2-30 µg/mL). It has different functions such as anti-inflammatory action due to the secretion of anti-inflammatory cytokines (IL-10), the ability to block the activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) and the inhibition of the release of tumor necrosis factor-α (TNF-α) and interleukin-6 (IL-6). It also

increases fatty acid oxidation and glucose uptake in skeletal muscle and inhibits hepatic gluconeogenesis.

This adipokine is influenced by obesity, IR, NAFLD and other conditions of MS. In fact, is has been seen that its concentration in plasma is lower in NAFLD patients than in control individuals (15).

The administration of adiponectin inhibits hepatic glucose production and favors the reduction of plasma levels of fatty acids and their beta-oxidation in muscle (15).

2.1.2 Leptin

Leptin is an adipokine, which circulating concentrations are associated with whole-body fat mass, as they reflect the amount of body energy storage as well as sense changes in energy intake. For this reason, this organokine is important in the regulation of appetite, body fat mass and controls energy balance in the hypo- and normoleptinemic state, specially by suppressing appetite (16).

Increased levels of leptin are associated with obesity, cardiovascular health and NAFLD. In fact, high levels of this organokine contribute to different comorbidities strongly associated with the progression of NAFLD. For example, it contributes to the action of pro-inflammatory cytokines, stimulates the division of stellate liver cells, the production of pro-fibrinogenic factors, etc. Moreover, it has been seen that patients with fibrosis present high levels of leptin. It is also related to the IR and the failure in the antisteatotic effect, caused on NAFLD (14).

2.1.3 Irisin

Irisin is a myokine produced by muscles and usually secreted after physical exercise. But this organokine is also recognized as adipokine identified in adipose tissue near human skeletal muscle and it is synthetized in many other tissues too.

Its principal aim is to induce “browning” of white adipose tissue, leading to an increase in the thermogenesis and the energy expenditure. Moreover, Irisin is related to the regulation of energy homeostasis and metabolism, and in the interactions between skeletal muscle and other tissues (17).

Its secretion is stimulated by oxygen consumption, and its concentration is related to loss of body mass, improvement of glucose tolerance, and decrease of insulin secretion.

Irisin has been linked to favorable effects on metabolic diseases, including T2DM, cardiovascular disease (CVD) and NAFLD. Although it has been detected at higher levels in lean patients than in obese, NASH or NAFLD patients, there are contradictory results in the published studies and more data is needed. For example, it has been identified that, depending on the evolution of portal inflammation, an increase in Irisin may occur, probably as a compensatory mechanism to limit inflammation (14).

2.1.6 Galectin3

Galectin3 is a beta-galactoside binding protein and it participates in diverse biological processes, including cell adhesion and growth. This organokine is considered a drug target for many diseases such as CVD, fibrotic or neurodegeneration diseases (18).

The expression of Galectin3, protein considered essential to the development of hepatic fibrosis, was increased in NASH with the highest expression in macrophages surrounding lipid-laded hepatocytes (18). Moreover, the results of several studies (17) suggest that the association of this organokine with NAFLD, may reflect reverse causality. This means that, disease processes increase Galectin3 levels, whereas higher levels of this organokine do not increase disease risk.

Some authors have tried to use Galectin3 as a therapeutic target to treat NAFLD and other diseases, however to date, it seems to be more useful as a biomarker of disease.

2.1.5 Fibroblast Growth Factor 19 (FGF19)

This is an intestine-derived hormone, which principal activities are related to metabolic regulation. In fact, this orgakine is considered a key factor that modulates bile acids (BA)/cholesterol synthesis. It is known that FGF19 is elevated in human plasma postprandially via activation of bile acids-farnesoid X receptor axis to repress the expression of the rate-limiting enzyme CYPA1 in the liver that controls BA synthesis (20). Moreover, FGF19 is a potential molecular target for T2DM and NAFLD.

In NAFLD patients, circulating FGF19 is found at low concentration. This shows altered enterohepatic BA homeostasis in NAFLD and also, an association with an abnormal hepatic lipid metabolism. In particular, human data support a co-relation between low serum FGF19 levels and hepatic steatosis. The most biologically plausible explanation of this relationship is that FGF19 deficiency precedes the development of steatosis because this deficiency decreases hepatic triglyceride oxidation while simultaneously increasing *de novo* lipogenesis. However, this hypothesis contradicts those studies which showed elevated FGF19 levels in patients with alternative etiologies of lives disease, such as alcoholic hepatitis and cholestasis (18).

2.1.4 Fibroblast Growth Factor 21 (FGF21)

FGF21 is an hepatokine/adipokine which is released by the TA and the liver. Its main activity is to regulate glucose and lipid metabolism (20). Due to this, it can adjust insulin sensitivity, body weight, lipoprotein profile and may have a role in inflammatory modulation, which is related to its potential to reduce IR (14).

This organokine is significantly elevated upon prolonged fasting in humans and its circulating levels are elevated in NAFLD patients. It is considered to play a protective role against NAFLD because it can increase the oxidation of FFAs in the liver, reducing glucose production and consequently, delaying steatohepatitis development (14).

Moreover, its high levels in NAFLD correlate with the amount of hepatic triglycerides. Therefore, it is considered as an important emerging biomarker of NAFLD (14).

HYPOTHESIS AND OBJECTIVES

Obesity is characterized by an excessive accumulation of fat, which ends up producing ectopic storage of lipids in the liver, with the consequent appearance of non-alcoholic fatty liver disease (NAFLD), and the further progression to non-alcoholic steatohepatitis (NASH). These diseases are very prevalent in the glboal population and they have many associated comorbidities in common.

The most problematic side of these diseases resides in their long-term consequences if not treated, as patients may have a very poor quality of life, and even increased risk of early death. In fact, the development of NASH is characterized for being silent until it turns irreversible, and the liver transplantation becomes the only option left for the patients. Moreover, currently a liver biopsy, besides being a very invasive procedure, is the only method to diagnose the severity of NAFLD.

But also, another problematic point is the unknown physiopathologic progression of NAFLD. We know that the liver is one of the most damaged organs because of this disease but we hypothesize if other organs could be also affected. We believe that there is a cross talk between organs such as intestines, pancreas, skeletal muscle, adipose tissue…which might be sharing information of damage through organokines.

The aim of this study was to assess concentration levels of 6 organokines (Adiponectin, FGF21, FGF19, Irisin and Galectin3) in plasma from obesity-associated NAFLD patients and normoweight healthy individuals in order to acquire more information to better characterize the physiopathology and progression of the disease. We believe that this information could be useful in medical decision making and, in some cases, it would avoid the invasive procedure of biopsy.

Specific objectives:

- To determine the concentration of six different organokines (Galectin3, Adiponectin, Irisin, FGF21, FGF19, Leptin) in plasma samples of patients with severe obesity who had undergone bariatric surgery, and suffer different stages of NAFLD.
- To compare different variables from these patients to find:
    - If organokines are affected by obesity and different obesity-associated comorbidities.
    - If organokines serve as biomarkers of NAFLD.
- To compare the concentration of organokines when patients are classified according to NASH diagnostic to establish a disease-specific organokine profile.
- To assess whether the organokine profile is related to the eight different scores of NAFLD and also if they are associated with its hepatic histological features.

MATERIALS AND METHODS

1. Study design

There were a total of 1,482 participants from whom BMI and NAFLD Activity Score (NAS) were employed to make groups. Among those patients, n= 1,068 had morbid obesity, and they were classified in three groups, depending on their disease's grade: 235 were diagnosed as NASH (NAS ≥ 5), 416 were classified as non-NASH (NAS ≤ 2), and 417 formed an intermediate group, named uncertain NASH (3 < NAS 4) (Figure 8).

Those patients with morbid obesity (BMI>30kg/m^2) met the same inclusion criteria to go under bariatric surgery in Sant Joan Hospital (Reus, Spain). All these patients were over 18 years old and were included both, men and women. The main exclusion criteria were having clinical or analytical evidence of severe illness, chronic or acute inflammation, infectious diseases, or terminal illness.

For the control group, we used age- and sex-matched blood samples from healthy normoweight volunteers previously recruited in our area from the Bank of biological samples from Institut d'Investigació Sanitària Pere Virgili (Biobanc-IISPV) (n= 434) (Figure 8).

The Ethics Committee of the Hospital Universitari Sant Joan (Reus, Spain) approved the study (OBESPAD/14.07-31proj3) and all the participants signed an informed consent (OM-NAFLD, ESO3/18012013 project).

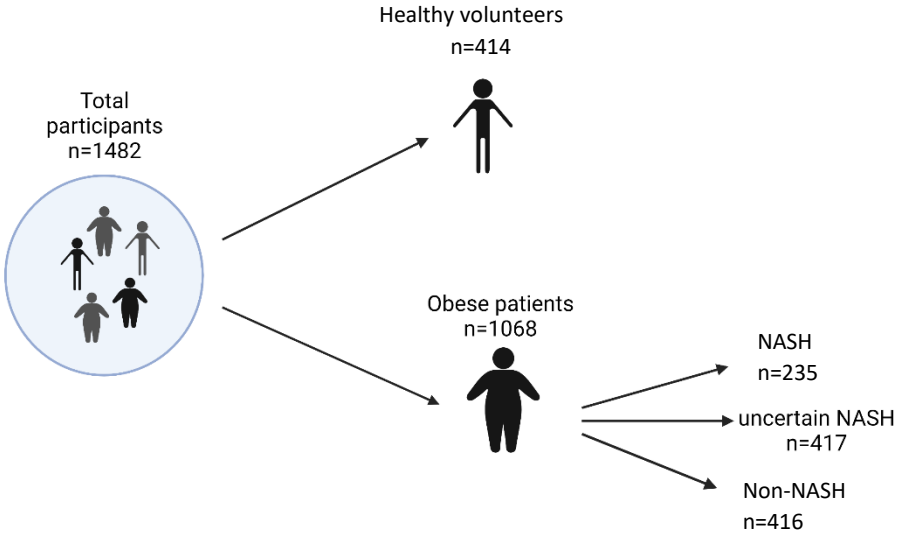

Figure 8: Schematic representation of the study design and distribution of patients

2. Sampling

We obtained blood samples from patients with obesity just before the bariatric surgical intervention. These samples were collected in tubes and centrifuged at 2500rpm for 15 minutes at 4ºC to obtain plasma and serum. Finally, we stored the samples at -80ºC for further analyses.

During bariatric surgery, a liver and adipose tissue biopsies were performed. These samples were fixed for 24 hours in formaldehyde. Then, we paraffin-embedded the samples to perform histological analysis.

3. Biochemical characteristics

We subjected the serum samples to biochemical analyses. An automatic analyzer (COBAS 8000, Roche Farma) was used for this purpose, to determine biochemical parameters:
glucose, insulin, total cholesterol, HDL-cholesterol, LDL-cholesterol, triglycerides and transaminases. Samples were subjected to chemiluminescence, immunoassays or enzymatic assays following the standard procedures from Laboratori de Referència del Camp de Tarragona i Terres de l'Ebre.

Homeostatic model assessment of insulin resistance (HOMA-IR) was used to estimate the insulin resistance of each patient (21).

4. Histological analyses

The paraffin-embedded samples were cut in 2 µm slices using the microtome.
From the sample of each patient, we obtained three different slices. We dewaxed them and performed three different stains: Masson's Trichrome, Hematoxylin and Eosin stainings for liver samples and Sirius Red staining for adipose tissue samples (Figures 9-11).

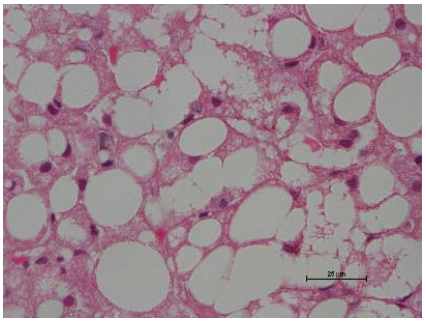

Figure 99: liver sample stained with Hematoxilin-Eosin where it can be seen the steatosis

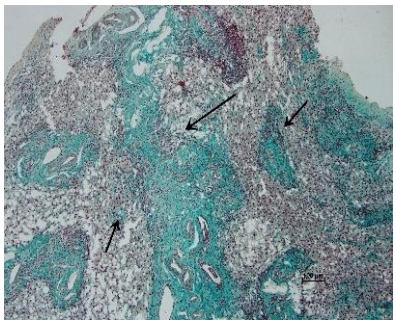

Figure 10: liver simple stained with Masson's Trichrome where it can be seen the perisinusoidal fibrosis bridges

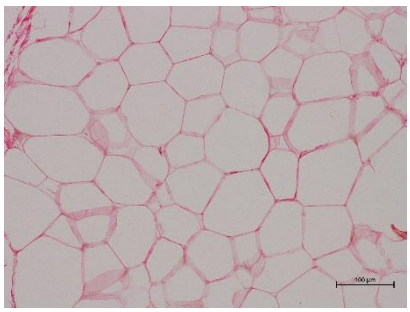

Figure 10: adipose tissue sample stained with Sirius Red where it can be seen the fibrosis

4.1 NAFLD Activity Score (NAS score)

For the determination of the NAS score, we observed in an optical microscope (Nikon, Eclipse E600, Madrid, Spain) the slices and followed the basis previous established by Kleiner (4). From the ones stained with Hematoxylin, we could evaluate the degree of steatosis, ballooning and lobular inflammation. From the ones stained with Masson's Trichrome, we could observe the

degree of fibrosis in each sample. All evaluations were carried out by an expert pathologist blinded to clinical data.

4.2 Adipocyte size

For the determination of the adipocyte area, we observed in an optical microscope (Nikon, Eclipse E600, Madrid, Spain) the slices stained with Hematoxylin and Eosin. We took from 7 to 9 photos of each sample and used the Image J software to determine its average area size with the specific tool "MRI_Adipocyte_Tool"(23).

5. Cytokine's determinations

To determine cytokines' concentration present in plasma samples, we performed an indirect sandwich Enzyme-linked immunosorbent assay (ELISA) (Figure 12).
This method allows quantification of a marker of interest, which in our case were cytokines, from a biological sample.

For our purpose, we used the following kits:

- Human Irisin/FNDC5 Duo Set ELISA Development Kit (R&D Systems, Catalog (#DY9420-05)
- Human Galectin3 ELISA Development Kit (R&D Systems, Catalog (#DY1154)
- Human Leptin/OB ELISA Development Kit (R&D Systems, Catalog (#DY398)
- Human Adiponectin/Acrp30 Duo Set ELISA Development Kit (R&D Systems, Catalog (#DY1065)
- Human FGF21 ELISA Development Kit (R&D Systems, Catalog #DY2539)
- Human FGF19 ELISA Development Kit (R&D Systems, Catalog #DY969)

The plasma samples had never been thawed before. Before being used for ELISA assay, they were centrifuged for 30 seconds at 800g. Then, we followed the manufacturer protocol. Moreover, if necessary, a dilution was performed as follows:

- Irisin: no dilution
- Leptin: dilution 1/120
- Adiponectin: dilution 1/6500
- Galectin3: dilution 1/7
- FGF21: no dilution
- FGF19: no dilution

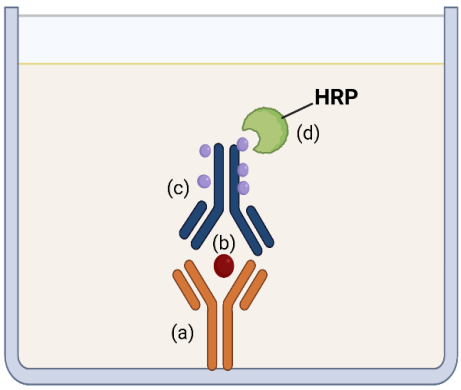

Figure 12: Representative sandwitch ELISA scheme. (a)Capture Ab. (b) Sample (c) Biotin-conjugated capture antibody (d) Avidin conjugated with HRP enzym

6. Data analysis and statistics

Statistical analyses were performed on MetaboAnalyst 5.0 (www.metaboanalyst.ca), GraphPad Prism 6 (GraphPad Software, San Diego, CA, USA) and RStudio (R version 4.0.2).

Outlier values were detected using the interquartile range method and were replaced by the median of the group.
The missing values were replaced using the k-nearest neighbors (KNN) method. Moreover, two-group comparisons were performed using t-test and more-than-two group comparisons were performed using analysis of variance (ANOVA). Differences between groups were considered significant if p-value <0.05.

For data management in excel format and dataframes within the R software, we used the package "Readxl" (24) and "dplyr" (25).

Population study:
To study the normality of the variables, we performed the Saphiro test. It is based on the correlation between the data and the corresponding normal scores.

Once studied the normality of our variables, we used the *"Tableone"* (26) package to study the significance of each variable. *"Tableone"* is an R package that eases the construction of tables of patients baseline characteristics commonly found in biomedical research papers. The package can summarize both continuous and categorical variables mixed within one table. Categorical variables were summarized as percentages. Continuous variables were summarized as mean and inter quartile range (IQR). Package "knitr" (28) was used to generate a report of obtained data.

Correlation analysis:
Different correlations analyses were performed. On the one hand, univariant correlations were studied using "ggplot2" and "ggpur" packages (30) that helped us mapping variables to aesthetics; "corrplot" package (29), provide us a visual exploratory tool on correlation matrix that supports automatic variable reordering to help detect hidden patterns among variables.

On the other hand, multivariant correlations were displayed to study the relation between all variables at the same time. For this purpose, we used the "Qgrapgh" package (31) as the creator specifies.

Logistic regression
We performed a logistic regression to obtain the odds ratio, that helped us knowing how strong were our variables in group differentiation. For this purpose, we used "mice" package (32) for the imputation of data, "questionr" package (33) which is a set of functions to make the processing and analysis of surveys easier and the "sjPlot" (34) package for plotting our results.

Box and radar plots

GraphPad software was used to plot the results in Bar Plots. Moreover, we also used it to obtain the significance through t-test and ANOVA.

We also used "fmsb" package (35) from R to represent the data in a Radar Chart. For this purpose, we calculated the mean of each organokin in each group of NASH, uncertain and non-NASH and established a minimum and maximum value for each organokine. This way, the RadarChart lines represents a different ratio for each organokine in each group of NASH. Comparing these ratios, we can find an organokines' profile in each group.

Metaboanalyst

This software was used to perform Partial Least Square Discriminant Analysis (PLS-DA) from which we showed the three-dimensional scores plots and the Variable Importance in Projection (VIP) scores; the heat maps provided direct visualization of all data points in the form of colors squares and also dendograms. To evaluate the biomarker ability of the organokines we performed a Monte Carlo cross validation model that randomly and repeatedly combined organokines and calculated Receiver Operating Characteristics (ROC) curves.

RESULTS

1. Comparison between healthy individuals and NAFL patients

1.1 Clinical and biochemical characteristics of the population

The main clinical and biochemical characteristics of the study participants are described in table 4, showed below.

In the following table 4, the population was stratified according to being whether normoweight or obese. It was shown that age was similar in both groups whereas sex was significantly different ($p<0.001$), being the women proportion greater in the obese group. Regarding the rest of the clinical characteristics, including concomitant diseases, biochemical characteristics, and transaminases, the results were all statistically significantly different between both groups.

Regarding to the percentage of treatments usually prescribed to patients with hypertension and dyslipidemia, it was seen stadistically significant different between obese patients and healthy controls. Treatment prescribed to patients with diabetes, showed no stadistically significant differences when both groups were compared.

Table 4: Main clinical and biochemical characteristics of the control population vs obese patients

	Healthy volunteers	Patients with obesity	p-value
Clinical characteristics			
n	414	1,068	
Sex (Woman %)	231 (56.9)	776 (72.7)	<0.001
Age (years)	46.00 (35.00, 59.50)	49.00 (41.00, 56.00)	0.047
BMI (kg/m2)	26.78 (23.34, 30.12)	44.14 (40.33, 48.52)	<0.001
Heart rate (bpm)	76.00 (68.00, 80.00)	75.00 (67.00, 84.00)	0.091
SAP(mmHg)	126.00 (116.00, 140.00)	129.00 (118.00, 141.00)	0.014
DAP (mmHg)	80.00 (70.00, 90.00)	77.00 (69.00, 86.00)	0.019
hipP (cm)	101.00 (95.00, 108.00)	136.00 (127.00, 145.00)	<0.001
waistP (cm)	90.00 (79.00, 99.00)	130.00 (120.00, 140.00)	<0.001
T2DM (%)	26 (6.3)	274 (25.7)	<0.001
Hypertension (%)	62 (15.0)	445 (41.7)	<0.001
Dyslipidaemia (%)	36 (8.7)	248 (23.2)	<0.001
Biochemical varaibles			
Glucose (mmol/L)	4.70 (4.30, 5.20)	6.72 (5.50, 8.50)	<0.001
Insulin (pmol/L)	49.42 (31.93, 70.05)	70.84 (39.59, 121.89)	<0.001
HOMAIR	1.48 (0.95, 2.31)	3.32 (1.72, 6.03)	<0.001
Triglycerides (mmol/L)	1.10 (0.70, 1.50)	1.49 (1.15, 2.05)	<0.001
Total cholesterol (mmol/L)	5.20 (4.60, 5.90)	4.09 (3.52, 4.76)	<0.001
LDL-cholesterol (mmol/L)	3.13 (2.59, 3.77)	2.48 (1.94, 3.05)	<0.001
HDL-cholesterol (mmol/L)	1.44 (1.21, 1.73)	0.98 (0.83, 1.25)	<0.001
Transaminases			
ALT (µKat/L)	0.32 (0.23, 0.44)	0.56 (0.39, 0.90)	<0.001
AST (µKat/L)	0.35 (0.30, 0.41)	0.56 (0.39, 0.83)	<0.001
GGT (µKat/L)	0.24 (0.16, 0.39)	0.35 (0.23, 0.58)	<0.001
Medication			
Biguanides (%)	6 (1.4)	196 (18.4)	<0.001
Sulfonylureas (%)	10 (2.4)	21 (2.0)	0.734
Statins (%)	8 (1.9)	181 (16.9)	<0.001
Diuretics (%)	20 (4.8)	111 (10.4)	0.001
ACE_ARBS (%)	15 (3.6)	249 (23.3)	<0.001

NASH: Non-alcoholic steatohepatitis; BMI: Body mass index; hipP: hip perimeter; waistP: waist perimeter; SAP: systolic arterial pressure; DAP: diastolic arterial pressure; T2DM: type 2 diabetes mellitus; HOMA-IR: homeostatic model assessment for insulin resistance; LDL: low-density lipoproteins; HDL: high-density lipoproteins; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT: gammaglutamyl transferase; ACE-ARBS: angiotensin converting enzyme and angiotensin receptor blockers. Continuous variables are shown as mean and IQR (25-75%). Categoric variables are expressed as percentage (%) of participants. Groups were compared using the "Tableone" package which uses the ANOVA test to analyse the continuous normal variables, the Kruskal.test to analyse the continuous nonnormal variables and the chisq.test to analyse the continuous once. P values <0.05: significant differences between healthy patients and obese patients.

1.2 BMI and Leptin were significantly correlated

In order to check if there was any confounding variable, por example, any disease in particular that may be affecting the organokines, we performed a covariate plot (Figure 13). In it, we saw that leptin concentration may be affected by a disease in a significative way.

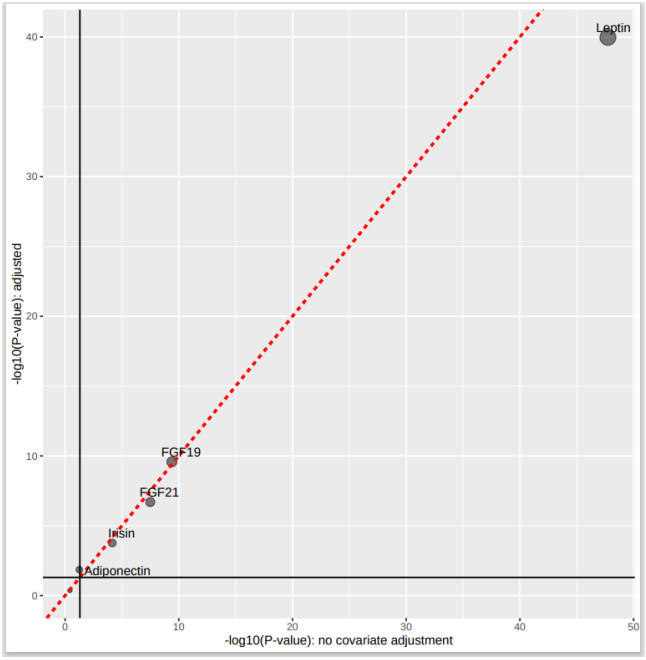

Figure 13: Covariate Plot were are represented the P-values for the organokines with and without covariate adjustment.

To get a better understanding of which variable might be and how significant the correlation was, we plotted a correlation matrix using Pearson's correlation method (Figure 14). At the individual correlation analysis (Figure 15) we obtained a significant statistic correlation (r=0.22) between BMI and Leptin.

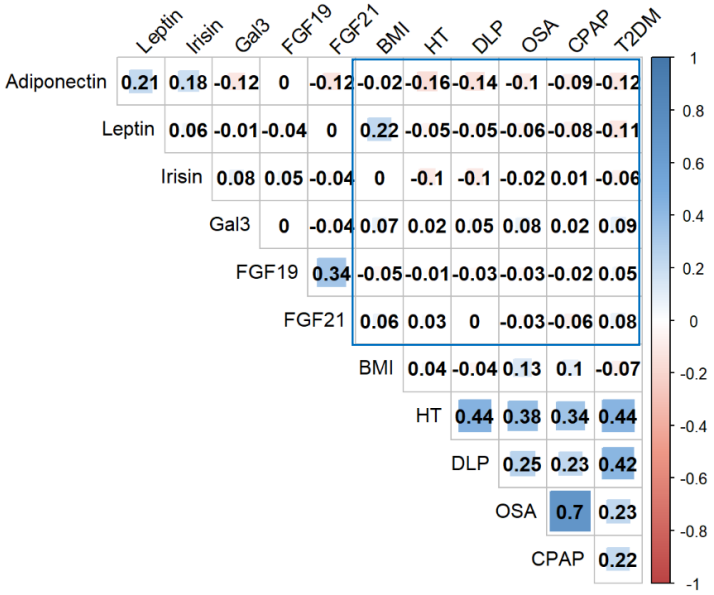

*Figure 14: Correlation plot, using Pearson's correlation of the organokines and the most frequent diseasas: HT, DLP, OSA, CPAP and BMI. * HT: hypertension, DLP: dyslipidemia; CPAP: continuous positive alrway pressure, BMI:body mass index*

Moreover, when plotting the correlation, we saw that it was significative ($p=2.2x10^{-9}$).

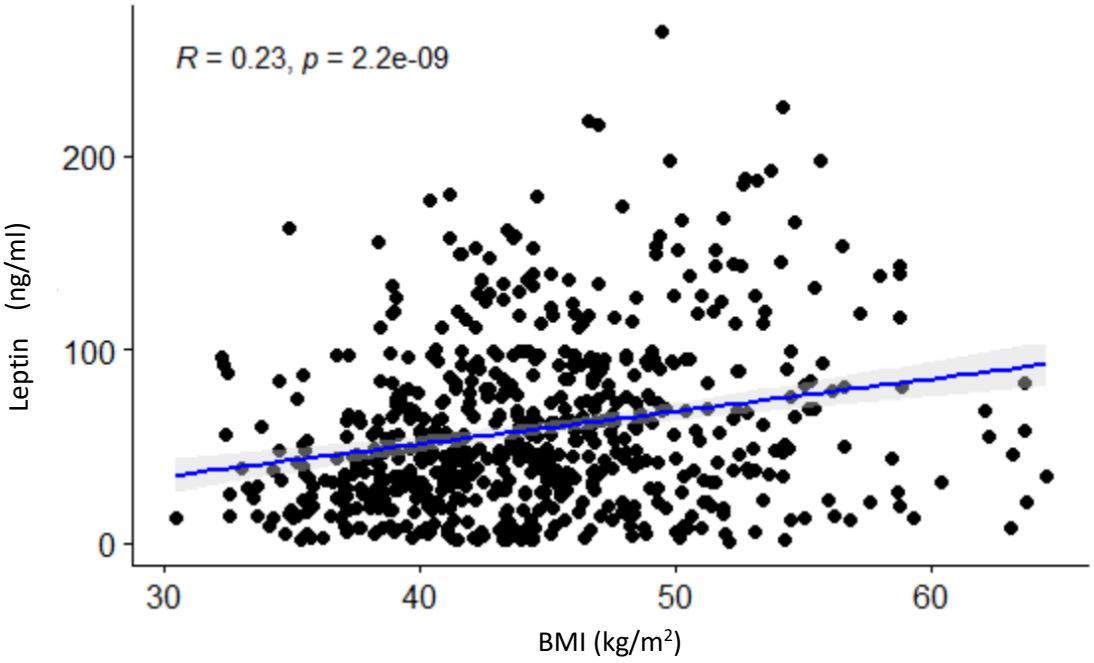

Figure 15: Correlation between Leptin and BMI. Filtering BMI<30 and Leptin<0. R indicates de correlation value and p de significance of the correlation.

In addition, to further explore the effect of BMI on Leptin, we plotted a heatmap where participants where organized according their BMI. (Figure 16) This showed how the concentration of only Leptin varied closely with BMI. Moreover, at figure 16 we could also appreciate the similitude between the variables. One example is that Irisin and Adiponectin were clustered close by, as well as it happened between Leptin and FGF19.

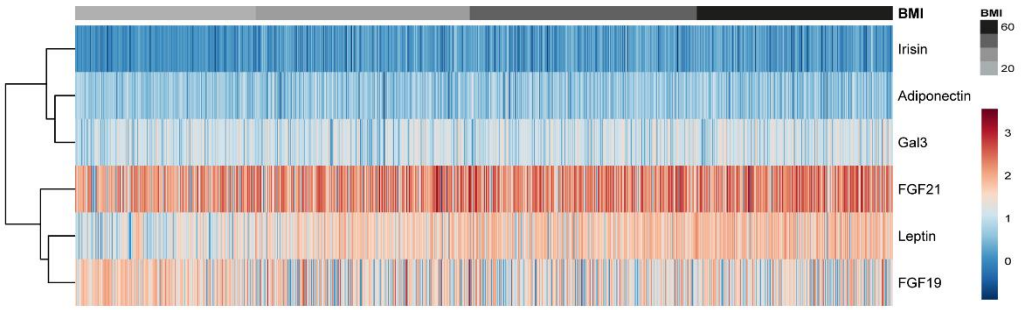

Figure 16: Heat map displaying how variates the Organokines concentration depending on the BMI variable. Moreover, the dendrogram is a tree diagram showing the groups formed by creating clusters of observations at each step and their levels of similarity.

1.3 Organokines were able to differentiate between NAFL patients and healthy individuals.

We wondered if organokines had the ability to distinguish between patients without any hepatic lesion (neither steatosis, lobular inflammation, ballooning nor fibrosis) and patients with a NAS score ranging from the least to the most severe (1 to 8). To do so, we first performed a logistic regression where the predictors were the concentration of each organokines. They were all significative despite Galectin3 and FGF19 concentrations. In figure 17 are represented the odds ratio and the strength of each organokine to differ between the control and the NAFL groups.

Figure 17: Odds ratio graphic representing the strength of each organokine.

Moreover, we performed a multivariate supervised analysis (PLS-DA) were it could be seen that, despite some overlapping, groups were separated only by using the organokines profile (Figure 18A).

We also performed a Monte Carlo cross validation, which tested the separation ability each organokines' concentration. We obtained a ROC curve (Figure 18B) and concluded that organokines were acceptable biomarkers for obesity-associated NAFL as the AUC (area under the curve) was 0.869.

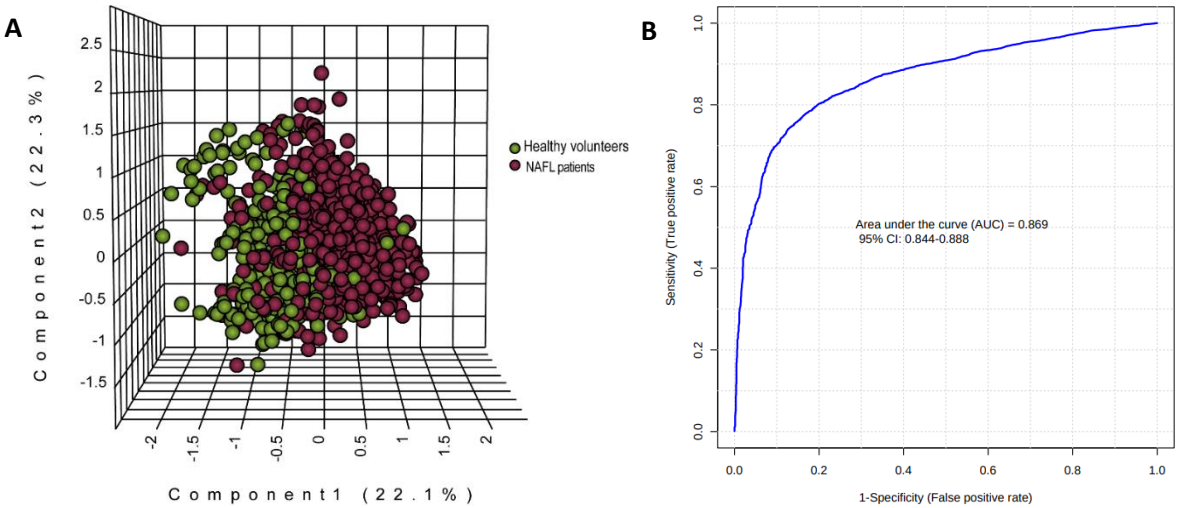

Figure 18: Discriminant Analysis (PLS-DA) of NAFL patients and Controls(A) ROC curve of the montacarlo cross validation representing the power of organokines as biomarkers (B)

2. Role of organokines in NASH: three NAFL groups comparative

2.1 Correlation network showed different associations between descriptive variables and organokines on the different stages of NAFLD

On table 5, the population was stratified according to NASH diagnosis in which we had three groups: NASH, non-NASH and uncertain NASH.

Among these patients, sex and BMI were similar in all groups whereas age was significantly different, being the NASH group the oldest. Following with the clinical characteristics we found that variables such as heart rate hip and waist perimeters, T2DM, hypertension, DLP, and obstructive sleep apnea (OSA) were significantly different among the three groups, all increased in NASH patients, whereas none of the variables related to the tiroid gland were significative.

Regarding to the biochemical variables glucose, insulin, HOMAIR and triglycerides showed significance too. The same happened with the histologic characteristics.

Finally, regarding to the percentage of treatments usually prescribed to each group of patients who suffer from DLP, hypertension and T2DM, all of them were significant different between these groups. The exception was the treatment with diuretics treatment for the hypertension and with fibrates for the DLP, in which case did not vary significantly.

Table 5: Main clinical and biochemical characteristics of the NASH patients, classified in three different groups: NASH, uncertain NASH and non-NASH

	Non-NASH (n=416)	Uncertain (n=417)	NASH (n=235)	p-value
Clinical characteristics				
Sex (Woman %)	319 (76.7)	293 (70.4)	164 (69.8)	0.067
Age (years)	47.00 (40.00, 56.00)	49.00 (42.00, 56.00)	51.00 (43.00, 57.00)	**0.02**
BMI (kg/m2)	43.56 (39.76, 47.62)	44.33 (40.78, 49.26)	44.79 (40.78, 48.95)	0.075
Heart rate (bpm)	74.00 (64.75, 81.00)	75.50 (69.00, 85.00)	78.00 (69.00, 92.50)	**0.016**
SAP (mmHg)	130.00 (120.00, 142.75)	128.00 (116.00, 140.50)	130.00 (119.00, 140.00)	0.309
DAP (mmHg)	78.00 (69.00, 87.00)	76.00 (68.00, 85.00)	79.50 (70.00, 86.00)	0.199
hipP (cm)	135.00 (125.00, 144.00)	135.00 (127.00, 145.00)	139.00 (131.00, 146.50)	**0.026**
waistP (cm)	129.00 (118.75, 139.00)	130.00 (120.50, 140.00)	134.00 (122.50, 141.00)	**0.017**
T2DM (%)	77 (18.6)	112 (26.9)	85 (36.2)	**<0.001**
HT (%)	142 (34.2)	177 (42.4)	126 (53.6)	**<0.001**
DLP (%)	73 (17.6)	99 (23.7)	76 (32.3)	**<0.001**
OSA (%)	72 (17.3)	113 (27.1)	75 (31.9)	**<0.001**
Hipotiroidism (%)	31 (7.5)	39 (9.4)	20 (8.5)	0.62
Biochemical varaibles				
Glucose (mmol/L)	6.47 (5.32, 8.01)	7.00 (5.61, 8.77)	7.00 (5.66, 9.09)	**0.002**
Insulin (pmol/L)	67.20 (35.42, 113.20)	70.84 (39.59, 127.09)	76.80 (47.23, 140.98)	**0.008**
HOMAIR	2.96 (1.50, 4.92)	3.32 (1.76, 6.50)	3.86 (2.18, 7.19)	**<0.001**
Triglycerides (mmol/L)	1.43 (1.11, 1.89)	1.55 (1.19, 2.05)	1.62 (1.26, 2.27)	**0.001**
Total cholesterol (mmol/L)	3.98 (3.47, 4.68)	4.16 (3.58, 4.80)	4.11 (3.57, 4.86)	0.294
LDL-cholesterol (mmol/L)	2.48 (1.97, 2.97)	2.47 (1.91, 3.13)	2.48 (1.93, 3.13)	0.904
HDL-cholesterol (mmol/L)	1.02 (0.85, 1.32)	0.98 (0.83, 1.22)	0.98 (0.80, 1.16)	**0.012**
Transaminases				
ALT (µKat/L)	0.50 (0.36, 0.70)	0.56 (0.39, 0.88)	0.83 (0.53, 1.31)	**<0.001**
AST (µKat/L)	0.49 (0.37, 0.70)	0.55 (0.39, 0.83)	0.76 (0.49, 1.14)	**<0.001**
GGT (µKat/L)	0.30 (0.21, 0.50)	0.35 (0.23, 0.52)	0.48 (0.34, 0.75)	**<0.001**
Treatments				
ACE_ARBS (%)	76 (18.3)	94 (22.5)	79 (33.6)	**<0.001**
Biguanides (%)	44 (10.6)	85 (20.4)	67 (28.5)	**<0.001**
Sulfonylureas (%)	5 (1.2)	9 (2.2)	7 (3.0)	0.274
Statins (%)	60 (14.4)	71 (17.0)	50 (21.3)	0.082
Diuretics (%)	40 (9.6)	41 (9.8)	30 (12.8)	0.4
Fibrates (%)	5 (1.2)	10 (2.4)	7 (3.0)	0.254
Tiroidal treatment (%)	30 (7.2)	34 (8.2)	18 (7.7)	0.878
CPAP (%)	48 (11.6)	70 (16.8)	52 (22.1)	**0.002**
Histologic characteristics				
Steatosis_score (%)				**<0.001**
0	300 (72.1)	102 (24.5)	-	
1	111 (26.7)	224 (53.7)	37 (15.7)	
2	5 (1.2)	85 (20.4)	119 (50.6)	
3	-	6 (1.4)	79 (33.6)	
Inflammation (%)				**<0.001**
0	125 (30.1)	15 (3.6)	-	
1	251 (60.5)	234 (56.4)	68 (28.9)	
2	39 (9.4)	138 (33.3)	130 (55.3)	
3	-	28 (6.7)	37 (15.7)	
Ballooning (%)				**<0.001**
0	295 (71.3)	107 (25.7)	6 (2.6)	
1	108 (26.1)	185 (44.5)	98 (41.7)	
2	11 (2.7)	124 (29.8)	131 (55.7)	
Fibrosis (%)				**<0.001**
0	76 (18.5)	20 (4.9)	14 (6.1)	
1	166 (40.5)	146 (35.9)	54 (23.7)	
2	142 (34.6)	181 (44.5)	93 (40.8)	
3	26 (6.3)	59 (14.5)	63 (27.6)	
4	-	1 (0.2)	4 (1.8)	
NAS (%)				**<0.001**
0	54 (13.0)	-	-	
1	144 (34.6)	-	-	
2	218 (52.4)	-	-	
3	-	229 (54.9)	-	
4	-	188 (45.1)	-	
5	-	-	131 (55.7)	
6	-	-	76 (32.3)	
7	-	-	24 (10.2)	
8	-	-	4 (1.7)	

NASH: Non-alcoholic steatohepatitis; BMI: Body mass index; hipP: hip perimeter; waistP: waist perimeter; SAP: systolic arterial pressure; DAP: diastolic arterial pressure; T2DM: type 2 diabetes mellitus; HOMA-IR: homeostatic

model assessment for insulin resistance; OSA: obstructive sleep apnea; LDL: low-density lipoproteins; HDL: high-density lipoproteins; ALT: alanine aminotransferase; AST: aspartate aminotransferase; GGT: gammaglutamyl transferase; ACE-ARBS: angiotensin converting enzyme and angiotensin receptor blockers; CPAP: continuous positive airway pressure; NAS: NAFLD Activity Score. Continuous variables are shown as mean and IQR (25, 75%). Categoric variables are expressed as number (percentage) (n (%)) of participants. Groups where compared using the "Tableone" package which uses the ANOVA test to analyse the continuous normal variables, the Kruskal.test to analyse the continuous nonnormal variables and the chisq.test to analyse the continuous once. P values <0.05: significant differences.

We also performed a correlation network analysis to better understand the relationship between descriptive continuous variables among the different stages of the disease and the control group (Figure 19).

Regarding to organokines composition, we observed how adiponectin was correlated with lipid and lipoprotein blood parameters in all groups and how FGF21 was correlated with steatosis in all groups too, being this correlation stronger in Uncertain (Figure 19B) and NASH (Figure 19C) groups. Also, when NASH was present, the waist and hip circumference and the BMI value were more correlated with FGF21. Moreover alkaline phosphatase (AP) and FGF21 were located close in the correlation zones.

Irisin was strongly correlated with VAT and SAT adipocyte areas in non and uncertain NASH groups, but not in NASH patients. On the other hand, Gal-3 behaves in an opposite manner: the correlation between this organokine and the area of adipocytes in adipose tissue, both SAT and VAT, is higher when NASH is developed. It is important to note that in the uncertain group, the correlation was positive, while in the NASH group, it was negative (Figure 19A-B).

Finally, we only found correlation between the organokines FGF19 and FGF21 in the non-NASH group (Figure 19A).

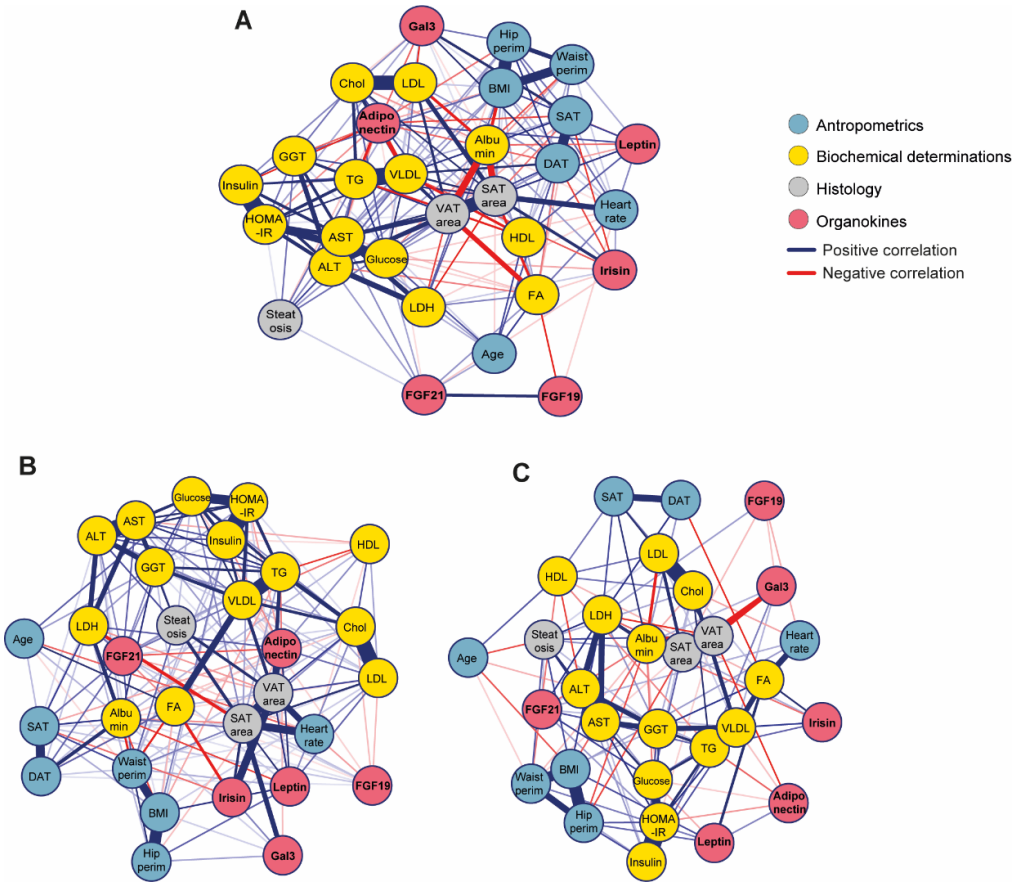

Figure 19: Correlation network analyses showed different associations between Organokines and clinical and biochemical characteristics depending on the studied population. Correlation networks from (A) patients with obesity and without NASH, (B) patients with obesity and with Uncertain diagnosis for NASH, and (C) patients with obesity and NASH. Thicker lines represent stronger correlations, closer localizations represent stronger correlations.

2.2 Only FGF-21 and Adiponectin discriminated between NASH groups in multivariate analyses

Through this analyse, we wanted to study if there was some organokine able to explain the grade of the disease.

Firstly, we performed a PLS-DA from where we obtained the VIP score (Figure 20). This showed which variables had more impact when it comes to distinguish between groups in the statistic model. We found out that FGF21 was the one with the most ability to distinguish between the groups.

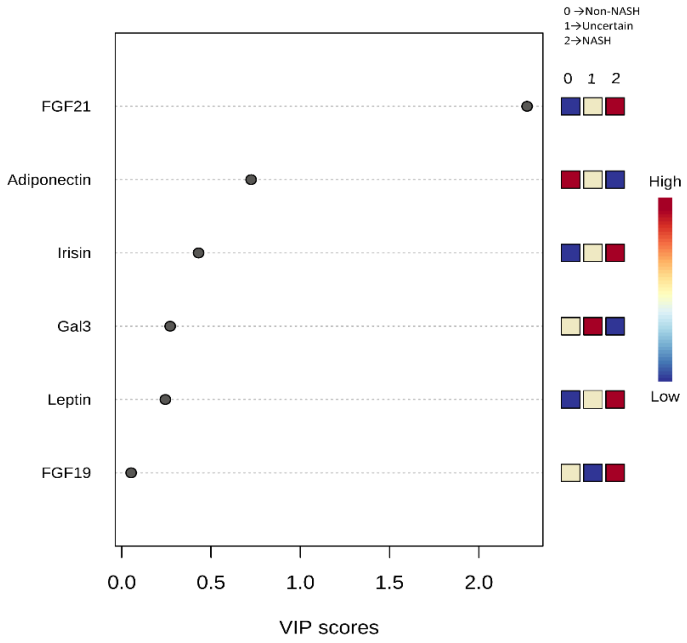

Figure 20: Variables Importance in Projection (VIP score) from Partial Least Square Discriminant Analysis showed that FGF21 had the most discriminant ability between groups.

Secondly, we performed a heat map (Figure 21) to determine if organokines concentrations were different depending on the degree of the NAFLD disease. We found out that organokines such as Adiponectin or Galectin3 were more present in non-NASH patients, whereas other organokines such as Leptin, Irisin or FGF21, were mostly increased in NASH patients. Finally, FGF19 clustered separately as its behaviour did not follow any pattern. Regarding the dendrogram, this showed Adiponectin and Galectin3, as well as Leptin and Irisin clustered together.

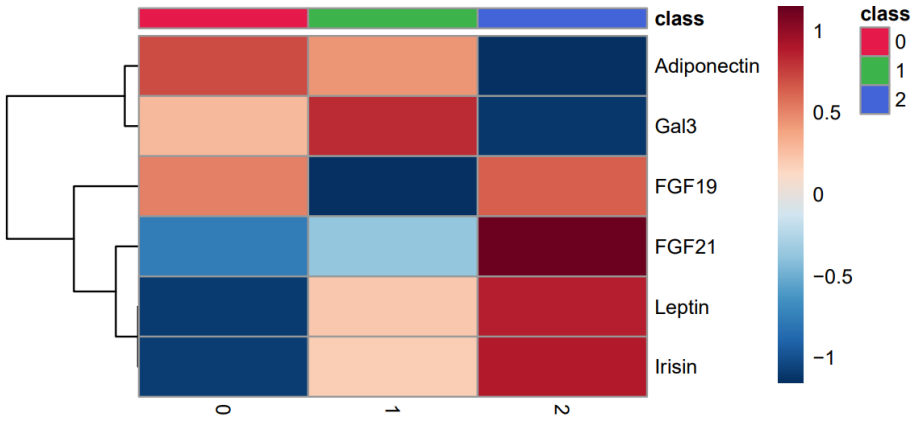

Figure 21: Heat map representing the different concentration of organokines depending on the NASH grade. It is also shown and endogram clustering together the organokines which shares more similitudes. 0: non-NASH, 1:uncertain, 2:NASH

We performed an ANOVA test to see if the differences seen in the heat map (Figure 21) and in the VIP score (Figure 20), were significative or not. Only the Adiponectin and the FGF21 comparisons between groups were statistically significant ($p= 0.016$ and $p<0.0001$, respectively).

Then, three different comparisons were carried out in order to better understand NASH-related changes. We compared two by two the different concentrations of each organokine using a t-test applying the following combinations: NASH vs non-NASH; non-NASH vs uncertain NASH; and uncertain NASH vs NASH. We found stadistically significant results when comparing NASH vs uncertain and NASH vs non-NASH in Adiponectin and FGF-21. (Table 6) We represent it in a Radar Chart. (Figure 22)

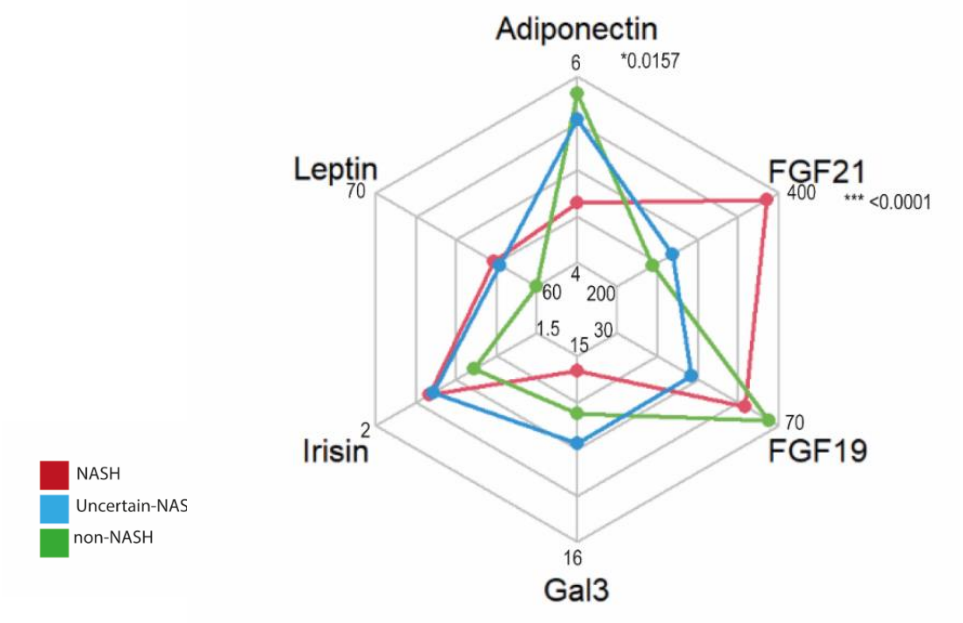

Figure 22: Radar chart representing the organokine concentration depending on the NASH, uncertain-NASH or non-NASH stage of the disease. Organokines ratio : 4-6 Adiponectin; 200-400 FGF21; 30-70 FGF19; 15-16 Gal3; 1.5-2 Irisin; 60-70 Leptin

Table 6: p values of the stadistically significant results obtained after the two by two comparisons.

	Comparation	p value
Adiponectin	NASH-nonNASH	0.0046
	NASH-uncertain	0.0182
FGF21	NASH-nonNASH	<0.0001
	NASH-uncertain	0.0005

We also performed a ROC curve of the Monte Carlo cross validation and we did not found any combination of organokines, enough powerful to be considered as biomarkers. (Data not shown)

3. NAFLD progression and histologic characteristics

3.1 NAFLD progression: role of organokines

To determine if there were different organokine concentrations levels among NAS score, a one-way one-way ANOVA was performed (Figure 23).

ANOVA test results in Adiponectin and Leptin showing stadistically significant results ($p<0.05$) explained bascally by the differences observed 0-5 and 0-6 score comparison in Adiponectin levels and 1-4 score comparison in Leptin levels.

Although not statistically significative, there was a tendency in FGF19 and FGF21 as NAS score was higher: the former decreased and the latter increased.

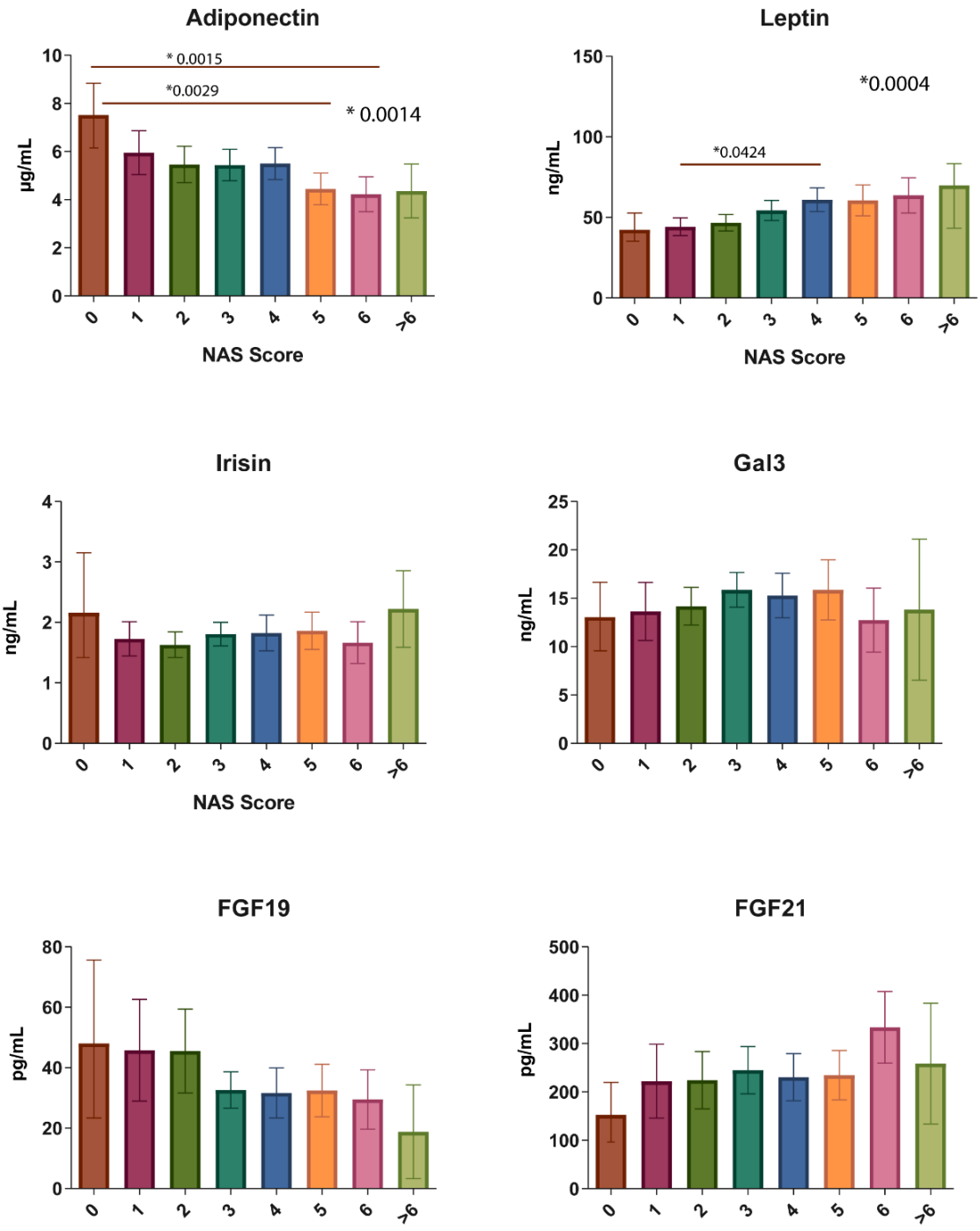

Figure 23: Bar chart representation of the different concentration of organokines among the NAFLD Activity Score (NAS).

3.2 Histologic characteristics were weakly associated with organokines levels

Steatosis, ballooning and lobular inflammation are the three characteristics used to evaluate NAFLD grading in patients. We performed an ANOVA test to see how these variables changed according to each organokine. We also realized a t-test comparing two by two each degree of the score (Figure 24).

We found out statistically significant results in Adiponectin, Leptin and FGF21, both when performing the ANOVA test and the t-test (Figure 24).

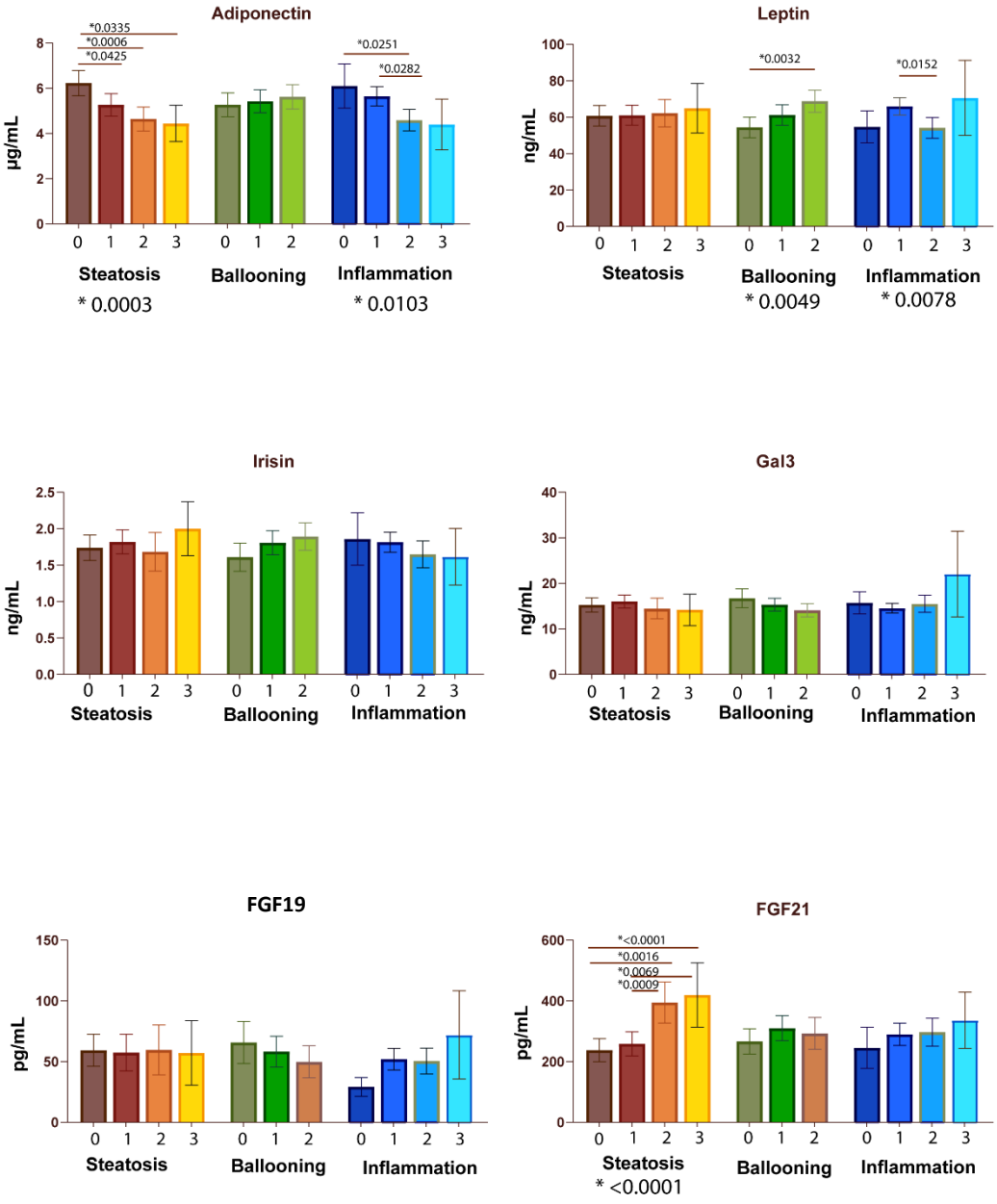

Figure 24 : Bar plot representation of the concentration of each organokine according to the degree of Steatosis, Ballooning and Inflammation. Steatosis: 0 = <5%, 1= 5%-33%, 2= 34%-66%, and 3= >66; Inflammation: 0= None, 1= <2 foci, 2=2-4 foci, and 3=>4 foci; Ballooning: 0=None, 1=Few Cells, 2=Many cells.

3.3 Organokines were different depending on hepatic fibrosis scores

We performed a Bar Plot for each organokine to determine if its concentration varied among the different fibrosis degree. We only found stadisticaly significant results in FGF21 (p =0.036), although a clear tendency was observed in Gal3: it decreased as fibrosis score increased, in contrast to Irisin, which had the opposite pattern. Also adiponectin and leptin had non-stadisticaly significant higher values in none fibrosis stage (Figure 25).

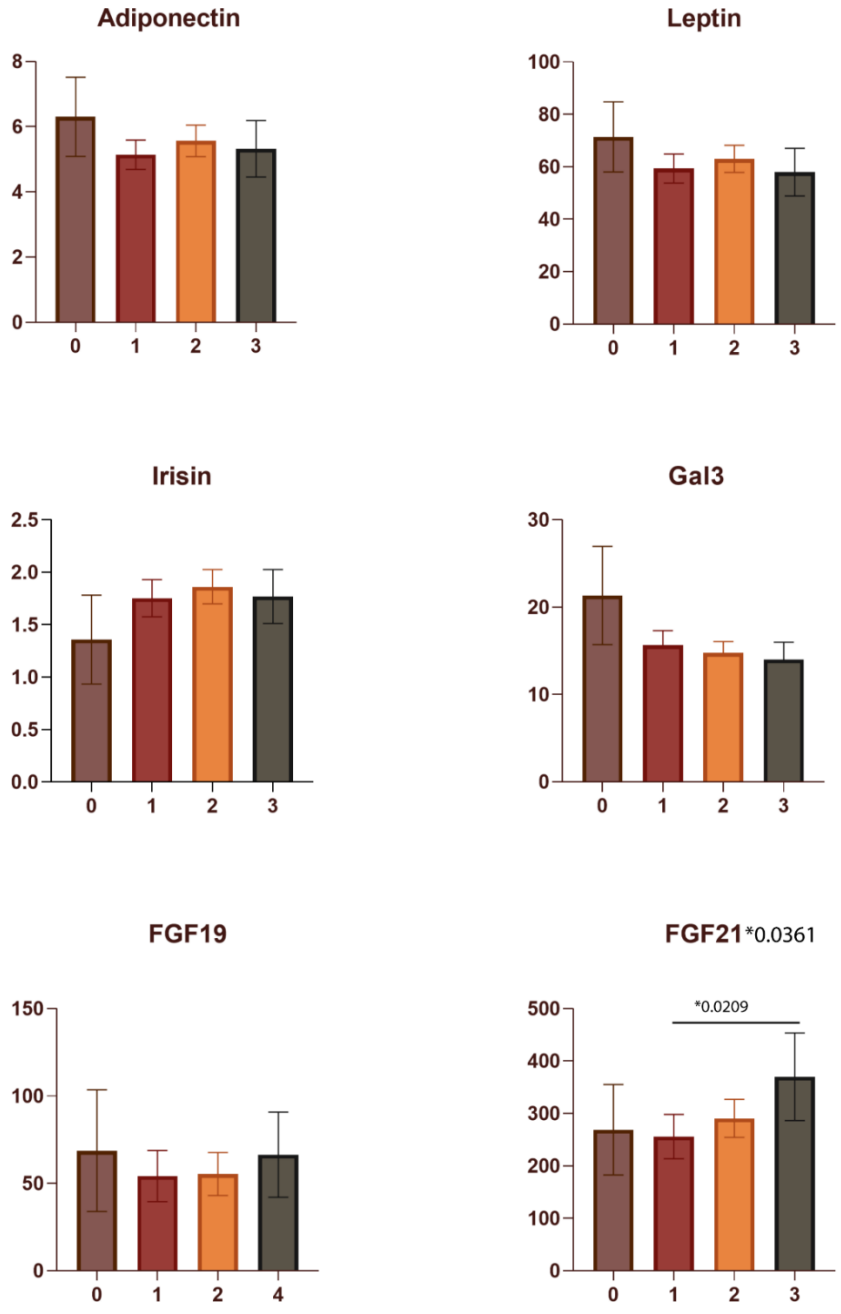

Figure 25: Bar plot representation of the concentration of each organokine according to the degree of Fibrosis: 0=None, 1=Perisinusoidal or periportal, 2=Perisinusoidal and portal/periportal, and 3=Bridging fibrosis.

3.4 Organokines are not suitable to be used a Fibrosis biomarkers

We performed a Monte Carlo cross validation model to test the organokine suitability to be biomarkers for fibrosis. From the analysis we obtained a ROC curve, which displays the specificity and the sensitivity of all generated models. We classified patients with obesity by their fibrosis score, considering the group of none fibrosis as control, and all the other stages as the cases. However, the resulting ROC did not have enough separation ability. (Figure 26)

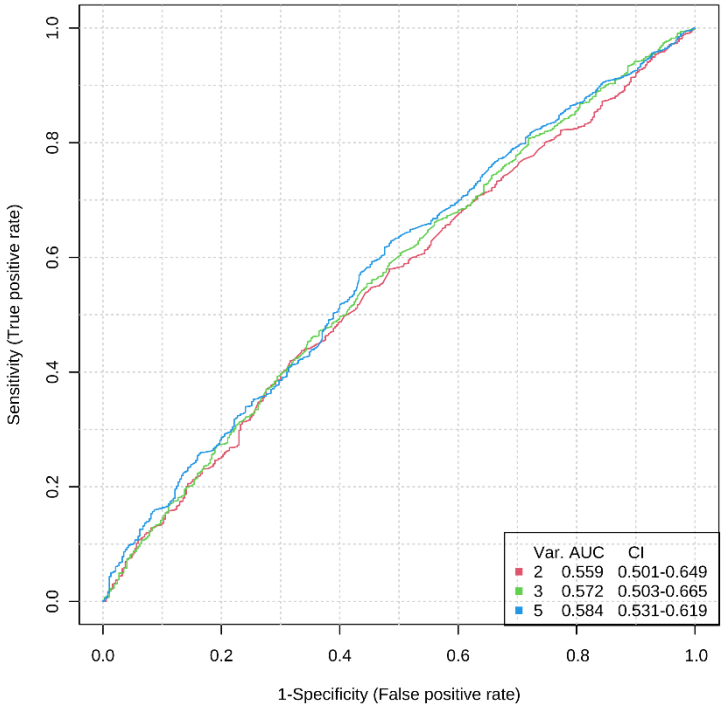

Figure 26: ROC curve of the Monte Carlo cross validation

DISCUSSION

As our goal was to determine the organokine profile along the obesity-associated NAFLD, we divided our results in three sections. The first one, dedicated to the comparison between healthy volunteers and NAFL patients aiming to study the effect of obesity. The second one, dedicated to review the variation in organokines concentration when studying the patients divided in three groups (NASH, uncertain and non-NASH). Finally, the third section was dedicated to the understanding of the NAFLD progression analysing the relationship between organokines and the histologic characteristics used to make the diagnostic.

Non-alcoholic fatty liver disease (NAFLD) is a spectrum of liver pathological states ranging from simple hepatic steatosis (NAFL) to non-alcoholic steatosis (NASH) with or without fibrosis or cirrhosis. Around 20% of patients with NAFLD have NASH and their metabolic deregulations are closely related to obesity. Both diseases are very related and are major health issue nowadays due to its increase in prevalence and mortality (36). Liver biopsy is the gold standard method to diagnose NAFLD and should only be performed after exclusion of other potential liver diseases. Moreover, each patient must be thoroughly aware of their diagnosis limitations as well as the lack of effective treatment(37).

Although there is no specific medication to cure NASH, patients suffering from it and who have treatment for other related disease such as T2DM or HT, seem to demonstrate a health improvement (38). However, we only found stadisticaly significative differences regarding to the medication administrated in the antihypertensive treatment.

Traditionally, obesity has been linked to a higher prevalence of non-communicable diseases, such as T2DM, hypertension and DLP, as well as being related to NAFLD and NASH (39). In fact, analysing the clinical and biochemical characteristics of the non-obese and obese patients in our population, we observed that the parameters associated with obesity such as BMI, blood pressure and obesity-related comorbidities were higher in the obese group, as expected.

As the principal objective was to determine the organokine profile levels among the progression of NAFLD, it was important to study whether the organokines were being affected by other comorbidities. We found out that leptin was correlated with the BMI, which means that Leptin levels might somehow depen on the BMI variable, or vice versa. By plotting a heatmap, we could see that, leptin values increased as the higher the BMI value was. These results are in line with other studies where had been also demonstrated that Leptin and BMI levels are correlated since Leptin is a dietary regulator hormone (40).

As many reviews claim, NAFL disease entails a liver dysfunction and, most of the time, it affects people who also suffer from numerous metabolic changes that commonly accompany the condition (obesity, IR, T2DM, and MS). This situation leads to a modified secretion of organokines which may contribute to the pathogenesis or progression of the disease (14). Analysing plasma organokines profile in healthy volunteers and NAFL patients, we noted that this profile was statistically significantly different between both groups, as expected. Moreover, since it does not exist a non-invasive method to determine if a patient has or not obesity-associated NAFLD, we claim that studying the organokines profile, specifically Adiponectin,

FGF21 and Irisin, may be a potential biomarker. Regarding to Leptine, a comparative among two groups with similar BMI should be done to clarify if it is a potential biomarker too.

Differences were not so clear when comparing NAFL patients divided in three groups (Non-NASH, uncertain NASH and NASH), and studying their different organokines profile. Considering those comorbidities related to metabolic syndrome (T2DM, hypertension, DLP and OSA), those were statistically significantly more frequent in NASH patients than in non-NASH and uncertain NASH, as expected due to their metabolic deregulation (36). The same happened with the hip and waist perimeters values. When looking for a link between hypothyroidism and NAFLD, and in agreement with the results obtained in previous studies (41), results demonstrate that there was no stadistically significant relationship between these pathologies.

Glucose, insulin and consequently, HOMA IR (a marker of IR) were also statistically significantly increased when NASH was developed. This correlated with the fact that several studies have shown that an increase in glucose is related to a malfunction of the liver. Due to this situation, insulin is also increased and HOMA IR do likewise, being both involved in the progression of disease conditions such as steatosis and NASH, as well as hepatic fibrosis progression (42).

Unexpected results were found when comparing the cholesterol related parameters. The liver is the major organ for cholesterol and lipoprotein metabolism (43) and, although we observed that total cholesterol and LDL-cholesterol concentrations were higher in uncertain and NASH than in non-NASH group, those differences were not statistically significant, and the same happened with HDL-cholesterol levels (44).

Regarding to the histologic characteristics used for the diagnostic of NASH, they were all statistically significantly increased as moving from non-NASH to NASH diagnosis (45).

Results were not so clear neither when comparing the organokines profile between the three groups of NASH, uncertain and non-NASH nor when studying their concentrations along the NAS score. As mentioned above, it is well known that NAFL disease involves metabolic changes in which many organs are affected and so do the organokines they produce. This presumably indicates that as worse the diagnostic of NAFLD is, the more affected those parameters will be. However, although we found visual differences along these three groups, only Adiponectin and FGF21 seemed to be significantly different when comparing NASH and non-NASH groups, and NASH and uncertain groups. The same pattern of stadistically significance was found when studying the organokine profile along the NAS score.

Nevertheless, it was shown that the liver may affect the lipids and glucose metabolism by hepatokines released into the blood and NAFLD seems to be associated with altered hepatokines production (46), and these results correlates with our results as FGF21, a hepatokine, showed higher concentrations when higher was the grade of NAS score. On the other hand, recent advances had shown that skeletal muscle produces myokines in response to exercise, which allow the crosstalk between the muscle and other organs, including brain, adipose tissue, bone, and liver. It has been identified that the biological roles of myokines include effects on, for example, cognition, lipid and glucose metabolism, browning of white fat, bone formation, endothelial cell function, hypertrophy, skin structure, and tumor growth (47). Our studied myokine, Irisin, showed no stadistically significant differences when studying NAFLD

progression but seem to be correlated with both VAT and SAT area in the three stages of NASH diagnosis, what make sense as myokines take part in adipose tissue crosstalk.

Adipose tissue, beyond its role as a fat storage depot, is an endocrine organ, capable of producing and releasing biologically active proteins, named 'adipokines', including leptin and adiponectin, the levels of which correlate with specific metabolic states and have the potential to alter metabolic homeostasis (40). Firstly, numerous epidemiological investigations have identified lower adiponectin level as an independent risk factor for NAFLDs and liver dysfunctions. Comparing the three groups, adiponectin levels were lower in NASH patients. Moreover, as the study conducted by Finelli C, et al (48) claims, its expression is decreased as more steatosis is present. NASH patients with lower levels of adiponectin showed higher grades of inflammation, suggesting that adiponectin deficiency is an important risk factor for the development of fatty liver, steatohepatitis and other forms of liver injurie. As our results were in line with those mentioned, we agree when the article (48) claims that adiponectin is a good predictor of steatosis grade and severity of NAFLD. Secondly, we noticed that leptin levels were increased as the disease got worse and so did the histological parameters. However, although different studies have proved that leptin levels correlate directly with the severity of hepatic steatosis, it is not with inflammation or fibrosis (49).

Hepatic fibrosis is a reversible wound-healing response characterized by the accumulation of extracellular matrix (ECM) due to an imbalance between production and dissolution of this matrix. It is important to detect fibrosis as its advanced stage may compromise liver function, however, we did not find any stadistically significant relationship between organokines concentration and fibrosis stages (50).

CONCLUSIONS

After the analyses performed in this study, several conclusions have been reached:

- Biochemical, clinical parameters and organokines' profile are far different between healthy volunteers and patients with obesity.
- Adiponectin, FGF21 and Irisin could serve as potential biomarkers to diagnose obesity-associated NAFLD.
- When comparing NASH, uncertain and non-NASH patients' organokines profile, some differences can be noted:
    - When FGF21 concentrations are high, Adiponecin ones are low, and vice-versa.
    - Results concerning leptin should be taken with caution. Further analysis should be carried out to find out how significant the correlation of organokine with BMI is.
    - It is still unclear if there is inter organ cross talk during the NAFLD development. However, it seems that Adiponectin and FGF21 concentrations vary during the NAFL progression.
- Refering to histologyc characteristics, we found a weak association with organokines concentration. However, statiticaly significant results were found in Adiponectin, FGF21 and Leptin.

Finally, is important to mention that there are some limitations in our study, mainly that a liver biopsy is necessary: the histologic evaluation of the biopsy sample is the gold standard method for the diagnosis, this technique has several disadvantages in relation to its routine and repeated use, as it is an expensive and invasive procedure. Moreover, a liver biopsy is not representative of the overall condition of the liver. Future perspectives include the realization of Western Blotting to study the proteins involved in cellular signaling pathways of these organokines. Last thing to mention is that control patients are considered healthy volunteers in our study. That is because, this group of patients suffered from non-communicable diseases in a normal proportion and we checked this patients did not suffer from associated liver pathologies.

To summarize, most objectives of this study have been achieved. However, it is needed further investigation in this field to find the answer to more questions that could not be found in this project.

ACKNOWLEDGEMENTS

En primer lloc m'agradaria agrair l'ajuda de la Dr. Helena Torrell, qui ha estat la meva tutora de TFG. Moltes gracies per les recomanacions i els consells per millorar el meu projecte.

També he d'agrair al Prof. Jorge Joven per fer-me un lloc a seu laboratori i permetrem aprendre i formar part dels seus projectes.

Gracies en general a tots els investigadors del grup, en especial a Helena Castañé. Gracies per guiar-me durant tot el procés de practiques i TFG, m'has sabut transmetre la teva dedicació i les ganes de seguir investigant. Espere seguir coincidint amb tu, poder compartir vivencies tant personals com profesionals i veure com aconsegueixes tot allò que t'il·lusiona.

Agrair també a Nuria Mimbrero, qui ha estat la meva companya de practiques i TFG. Ha sigut un plaer coneixert i poder compartir aquests mesos aprenent juntes, compartint preocupacions i rient-nos juntes dels errors.

Gracies als meus pares i la meva germana per estar sempre recolzant-me, escoltant-me durant hores al telèfon i fent-me sentir com a casa. Gracies als amics de sempre que, encara que els nostres camins es separin, troben la forma de fer-me sentir prop i amb ganes de riure junts de nou.

Per últim, donar-te les gracies a tu Manu. Gracies per sentar-te al meu costat cada vegada que em sentia bloquejada amb el treball, per escoltar-me durant hores explicar-te cada punt de l'estudi que anàvem aconseguint, per aguantar-me les llàgrimes i saber fer-me riure al instant. Gracies per fer de Tarragona una segona casa per a nosaltres. Sense tu, ni el TFG ni la carrera en general haguesin sigut el mateix.

BIBLIOGRAPHY

(1) Mundi MS, Velapati S, Patel J, Kellogg TA, Abu Dayyeh BK, Hurt RT. Evolution of NAFLD and Its Management. Nutr Clin Pract. 2020 February; 35(1):72-84.

(2) Raza S, Rajak S, Upadhyay A, Tewari A, Anthony Sinha R. Current treatment paradigms and emerging therapies for NAFLD/NASH. Front Biosci (Landmark Ed). 2021 January 1; 26(29):206-237

(3) Powell, Elizabeth E, Wong, Vincent Wai-Sun, Rinella, Mary. Non-alcoholic fatty liver disease. The Lancet. 2021 June 5; 397(10290):2212-2224

(4) Nakatsuka, T, Tateishi R, Koike K. Changing clinical management of NAFLD in Asia. Liver International. Liver Int. 2021 August 29

(5) Jennings J, Faselis C, Yao MD. NAFLD-NASH: An Under-Recognized Epidemic. Curr Vasc Pharmacol. 2018;16(3):209-213.

(6) Weir CB, Jan A. BMI Classification Percentile And Cut Off Points. StatPearls. 2022 Jan 29

(7 Polyzos SA, Kountouras J, Mantzoros CS.. Obesity and nonalcoholic fatty liver disease: From pathophysiology to therapeutics. Metabolism 2019 March;92:82-97.

(8) Xiaoyan Li & Hua Wang. Multiple organs involved in the pathogenesis of non-alcoholic fatty liver disease. Cell Biosci. 2020 December 07; 10(1):140

(9) Sanchez-Torrijos Y, Ampuero J. The Spectrum of NAFLD: From the Organ to the System. NAFLD and NASH. 2020 February 29.

(10) Rada P, González-Rodríguez Á, García-Monzón C, Valverde ÁM. Understanding lipotoxicity in NAFLD pathogenesis: is CD36 a key driver?. Cell Death Dis. 2020 September 25; 11(9):802.

(11 Pierantonelli I, Svegliati-Baroni G. Nonalcoholic Fatty Liver Disease: Basic Pathogenetic Mechanisms in the Progression From NAFLD to NASH. Transplantation. 2019 January;103(1)

(12) David E. Kleiner; Elizabeth M. Brunt; Mark Van Natta; Cynthia Behling; Melissa J. Contos; Oscar W. Cummings; Linda D. Ferrell; Yao-Chang Liu; Michael S. Torbenson; Aynur Unalp-Arida; Matthew Yeh; Arthur J. McCullough; Arun J. Sanyal. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology. 2005 June; 41(6), 1313–1321.

(13 Montgomery MK, De Nardo W, Watt MJ.. Impact of Lipotoxicity on Tissue “Cross Talk” and Metabolic Regulation. Physiology (Bethesda) 2019 March 1; 34(2):134-149.

(14) Santos JPMD, Maio MC, Lemes MA, Laurindo LF, Haber JFDS, Bechara MD,et al. Non-Alcoholic Steatohepatitis (NASH) and Organokines: What Is Now and What Will Be in the Future. Int J Mol Sci. 2022 January 2; 23(1):498.

(15) Zhang X, Ji X, Wang Q, Li JZ. New insight into inter-organ crosstalk contributing to the pathogenesis of non-alcoholic fatty liver disease (NAFLD). Protein Cell. 2018;9(2):164-177.

(16) Boutari C, Mantzoros CS. Adiponectin and leptin in the diagnosis and therapy of NAFLD. Metabolism 2020 February; 103:154028 .

(17) Polyzos SA, Anastasilakis AD, Efstathiadou ZA, Makras P, Perakakis N, Kountouras J, Mantzoros CS. Irisin in metabolic diseases. Endocrine. 2018 February;59(2):260-274

(18) Weaver MJ, McHenry SA, Sayuk GS, Gyawali CP, Davidson NO. Bile Acid Diarrhea and NAFLD: Shared Pathways for Distinct Phenotypes. Hepatol Commun. 2020 February 9; 4(4):493-503.

(19) Sumida Y, Yoneda M. Current and future pharmacological therapies for NAFLD/NASH. J Gastroenterol. 2018 March; 53(3):362-376.

(20) Talukdar S, Kharitonenkov A. FGF19 and FGF21: In NASH we trust. Mol Metab. 2021;46:101152. 2018 February; 59(2):260-274

(21) D R Matthews, J P Hosker, A S Rudenski, B A Naylor, D F Treacher, R C Turner. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985 July;28(7):412-9.

(22) Pierre Bel Lassen, Frederic Charlotte, Yuejun Liu, Pierre Bedossa , Gilles Le Naour, Joan Tordjman et all. The FAT Score, a Fibrosis Score of Adipose Tissue: Predicting Weight-Loss Outcome After Gastric Bypass. J Clin Endocrinol Metab. 2017 July 1;102(7):2443-2453.

(23) Chen HC, Farese RV Jr. Determination of adipocyte size by computer image analysis. J Lipid Res . 2002 Jun;43(6):986-9.

(24) Hadley Wickham and Jennifer Bryan (2019). readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl

(25) Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2022). dplyr: A Grammar of Data Manipulation. R package version 1.0.8. https://CRAN.R-project.org/package=dplyr

(26) Kazuki Yoshida and Alexander Bartel (2022). tableone: Create 'Table 1' to Describe Baseline Characteristics with or without Propensity Score Weights. R package version 0.13.2. https://CRAN.R-project.org/package=tableone

(27) Yihui Xie (2021). knitr: A General-Purpose Package for Dynamic Report. Generation in R. R package version 1.34.

(28) H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

(29) Taiyun Wei and Viliam Simko (2021). R package 'corrplot': Visualization of a Correlation Matrix (Version 0.92). Available from: https://github.com/taiyun/corrplot

(30) Alboukadel Kassambara (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0: https://CRAN.R-project.org/package=ggpubr

(31) Sacha Epskamp, Angelique O. J. Cramer, Lourens J. Waldorp, Verena D. Schmittmann, Denny Borsboom (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. Journal of Statistical Software, 48(4), 1-18. URL http://www.jstatsoft.org/v48/i04/

(32) Lüdecke D (2021). _sjPlot: Data Visualization for Statistics in Social Science_. R package version 2.8.10, <URL: https://CRAN.R-project.org/package=sjPlot>

(33) Julien Barnier, François Briatte and Joseph Larmarange (2022). questionr: Functions to Make Surveys Processing Easier. R package version 0.7.7. https://CRAN.R-project.org/package=questionr

(34) Lüdecke D (2021). _sjPlot: Data Visualization for Statistics in Social Science_. R package version 2.8.10, <URL: https://CRAN.R-project.org/package=sjPlot>

(35) Minato Nakazawa (2022). fmsb: Functions for Medical Statistics Book with some Demographic Data. R package version 0.7.3. https://CRAN.R-project.org/package=fmsb

(36) Sheka AC, Adeyi O, Thompson J, Hameed B, Crawford PA, Ikramuddin S. Nonalcoholic Steatohepatitis. JAMA, 2020 March 24;323(12):1175-1183

(37) Cortez-Pinto H, Camilo ME. Non-alcoholic fatty liver disease/non-alcoholic steatohepatitis (NAFLD/NASH): diagnosis and clinical course. Best Pract Res Clin Gastroenterol. 2004 December;18(6):1089-104.

(38) Singh S, Osna NA, Kharbanda KK. Treatment options for alcoholic and non-alcoholic fatty liver disease: A review. World J Gastroenterol. 2017 September 28;23(36):6549-6570.

(39) Litwin M, Kułaga Z. Obesity, metabolic syndrome, and primary hypertension. Pediatr Nephrol. 2021 April;36(4):825-837.

(40) Boutari C, Mantzoros CS. Adiponectin and leptin in the diagnosis and therapy of NAFLD. Metabolism. 2020 February; 103:154028.

(41) Martínez Escudé A, Pera G, Arteaga I, Expósito C, Rodríguez L, Torán P, Caballeria L. Relationship between hypothyroidism and non-alcoholic fatty liver disease in the Spanish population. Med Clin (Barc). 2020 January 10;154(1):1-6.

(42) Fujii H, Kawada N, Japan Study Group Of Nafld Jsg-Nafld. The Role of Insulin Resistance and Diabetes in Nonalcoholic Fatty Liver Disease. Int J Mol Sci. 2020 May 29;21(11):3863.

(43) Li H, Yu XH, Ou X, Ouyang XP, Tang CK. Hepatic cholesterol transport and its role in non-alcoholic fatty liver disease and atherosclerosis. Prog Lipid Res. 2021 July; 83:101109

(44) Kashyap SR, Diab DL, Baker AR, Yerian L, Bajaj H, Gray-Mcguire C, et al. Triglyceride levels and not adipokine concentrations are closely related to severity of nonalcoholic fatty liver disease in an obesity surgery cohort. Obesity. 2009;17(9):1696–701

(45) Enomoto H, Bando Y, Nakamura H, Nishiguchi S, Koga M. Liver fibrosis markers of nonalcoholic steatohepatitis. World J Gastroenterol. 2015;21(24):7427-7435.

(46) Lebensztejn DM, Flisiak-Jackiewicz M, Białokoz-Kalinowska I, Bobrus-Chociej A, Kowalska I. Hepatokines and non-alcoholic fatty liver disease. Acta Biochim Pol. 2016;63(3):459-67.

(47) Severinsen MCK, Pedersen BK. Muscle-Organ Crosstalk: The Emerging Roles of Myokines. Endocr Rev. 2020 August 1;41(4):594–609

(48) Finelli C, Tarantino G. What is the role of adiponectin in obesity related non-alcoholic fatty liver disease?. World J Gastroenterol. 2013;19(6):802-812.

(49) Procaccini C, Galgani M, De Rosa V, Carbone F, La Rocca C, Ranucci G, Iorio R, Matarese G. Leptin: the prototypic adipocytokine and its role in NAFLD. Curr Pharm Des. 2010 June;16(17):1902-12.

(50) Hernandez-Gea V, Friedman SL. Pathogenesis of liver fibrosis. Annu Rev Pathol. 2011;6:425-56.

ANNEX

Used R packages

List of the used packages and how to citate them

```
library(readxl)
citation("readxl")
library(dplyr)
citation("dplyr")
library(tableone)
citation("tableone")
library(knitr)
citation("knitr")
library(ggplot2)
citation("ggplot2")
library(corrplot)
citation("corrplot")
library(ggpubr)
citation("ggpubr")
library(qgraph)
citation("qgraph")
library(mice)
citation("mice")
library(questionr)
citation("questionr")
library(sjPlot)
citation("sjPlot")
library(fmsb)
citation("fmsb")
```

NAFL patients vs healthy volunteers statistics

EOM data base

```
library(readxl)
EOM_I_CONTROLSB_1_ <- read_excel("C:/Users/irene/Downloads/EOM I CONTROLS B (1).xlsx")
w<-EOM_I_CONTROLSB_1_
```

Variables to be included

```
As there were variables in the database that we did not want to include, we made a list of those that we wanted to be part of our study. With this list we did not have the content of the columns, that's why we had to create a data frame with the chosen variables.

variables<c("Sex","Age","BMI","Heart_rate","TAS","TAD","Pcadera","Pcintura","T2DM","HT","DLP","Glucose","Insulin2","HOMAIR","TG","CHOL", "LDL","HDL", "ALT","AST","GGT","Biguanides","Sulfonylureas","Statins", "Diuretics","ACE_ARBS", "Control_vs_EOM")

dfeo<-w[,variables]
```

Variables type

There were to different ways to explore our data type

```
str(dfeo)
lapply(dfeo,class)
```

From numeric to factor

```
Many of the variables we had were numerical variables of type 0 and 1. In order to made a difference between those that were numerical  such as height or weight, we passed these others to factor. Moreover, factor type variables assume a certain order, i.e. a category or a level. For example, in smoker: a 0 will be a NO, which has a better positive connotation than 1->YES.

Noves_factoreo<- c("T2DM","HT","DLP","Biguanides","Sulfonylureas", "Diuretics","ACE_ARBS","Statins")
dfeo[, Noves_factoreo] <- lapply(dfeo[, Noves_factoreo], as.factor)
```

Shapiro test

```
This test allowed us to determine which variables had a normal distribution and which did not. We could only apply the test to numerical variables, since the categorical variables were assumed to be non-normal.

numeriques<c("Age","BMI","Heart_rate","TAS","TAD","Pcadera","Pcintura","Glucose","Insulin2","HOMAIR","TG","CHOL", "LDL","HDL", "ALT","AST","GGT")
```

Correlation

```
library(readxl)
library(ggplot2)
library(corrplot)
require(Hmisc)
Proba.cor <- read_excel("C:/Users/irene/Desktop/TFG/Coses TFG/Base_defini
tiva.xlsx")
##Proba.cor<- cor(Proba.cor, method ="pearson")
vars_cor<-c("Adiponectin","Leptin","Irisin","Gal3", "FGF19","FGF21","BMI"
,"HT","DLP","OSA","CPAP","T2DM")
Proba.cor<-Proba.cor[,vars_cor]
Proba.cor<-rcorr(data.matrix(Proba.cor), type="pearson")
round(Proba.cor$r, digits=2)
corrplot(Proba.cor$r)
col<- colorRampPalette(c("#BB4444","#EE9988", "#FFFFFF","#77AADD", "#4477
AA"))
corrplot(Proba.cor$r, method="shade", shade.col =NA, tl.col= "black", tl.
srt =45, col= col(200), addCoef.col ="black", addcolorlabel= "no")
corrplot(Proba.cor$r, method="square", tl.col= "black", tl.srt= 45, col=c
ol(200), addCoef.col= "black", type ="upper", diag =F, addshade ="all")
```

Correlations

Leptin and BMI : graphics showed that there was some correlated variables. We performed a Pearson correlation and saw that Leptin and BMI were significantly correlated.

```
Proba.cor.BMI <- read_excel("C:/Users/irene/Desktop/TFG/Coses TFG/Base_de
finitiva.xlsx")

Leptines<-subset(Proba.cor.BMI, Proba.cor.BMI$Leptin>0)
Leptines2<-subset(Leptines, Leptines$BMI>30)
library(ggplot2)
library(ggpubr)

BMI <- ggscatter(Leptines2, x = "BMI",y = "Leptin",
   add = "reg.line",  # Add regressin line
   add.params = list(color = "blue", fill = "lightgray"), # Customize reg
. line
   conf.int = TRUE)
BMI+ stat_cor(method = "pearson")
```

Table one NAFL patients

Data base EOM

```
library(readxl)
Base_definitiva <- read_excel("C:/Users/irene/Desktop/TFG/Coses TFG/Base_definitiva.xlsx")

x<-Base_definitiva
```

Variables to be included

```
We make a list of the variables that we want to include in the study. This list does not have the content of the columns of each variable, for this reason, we will have to create a data frame with the chosen variables.

variables<c("Sex","Age","BMI","Heart_rate","TAS","TAD","Pcadera","Pcintura","T2DM","HT","DLP", "OSA", "CPAP","Hipotiroidism","Hipertiroidism","Glucose","Insulin2","HOMAIR","TG","CHOL", "LDL","HDL", "ALT","AST", "GGT", "Steatosis_score","Inflammation","Ballooning", "Fibrosis", "NAS","ACE_ARBS",  "Biguanides","Sulfonylureas","Statins", "Diuretics","Fibrates","Horm_tiroid","NAS_qual")

df<-x(,variables)
```

Variables type

Two different ways of studying which type of variables we had.

```
str(df)
lapply(x,class)
```

From numeric to factor

```
Many of the variables we have are numerical variables of type 0 and 1. In order to make a difference between those that are numerical such as height or weight, we pass these others to factor. Moreover, factor type variables assume a certain order, i.e. a category or a level. For example, in smoker: a 0 will be a NO, which has a more positive connotation than 1->YES.

Noves_factor<- c("T2DM","HT","DLP", "OSA","CPAP","Hipotiroidism","Hipertiroidism","Steatosis_score","Inflammation","Ballooning", "Fibrosis", "NAS", "Biguanides","Sulfonylureas","Statins", "Diuretics","ACE_ARBS","Fibrates","Horm_tiroid", "NAS_qual")
df(, Noves_factor) <- lapply(df(, Noves_factor), as.factor)
```

Shapiro test

```
This test allows us to determine which variables have a normal distrib
ution and which do not. We can only apply the test to numerical variab
les, since the categorical variables are assumed to be non-normal.

numeriques<c("Age","BMI","Heart_rate","TAS","TAD","Pcadera","Pcintura"
,"Glucose","Insulin2","HOMAIR","TG","CHOL", "LDL","HDL", "ALT","AST","
GGT")

shapiro<-lapply(df(, numeriques), shapiro.test)
shapiro
```

We saved the different variables in to list depending on whether they were normal or not normal

```
nonormales<-c("Age","BMI","Heart_rate","TAS","TAD","Pcadera","Pcintura
","Glucose","Insulin2","HOMAIR","TG","CHOL", "LDL","HDL", "ALT","AST",
"GGT")
```

Univariant statistics

```
We apply the tableone library that applies the appropriate test for ea
ch variable, depending on whether it has a normal distribution or not.
 And it indicates which are the p-values, thus being able to obtain wh
ich variables are significant. It is important to stratify using the N
AS_qual variable, which is the one we are interested

in.library(knitr)
library(tableone)
kableone <- function(x, ...) {
  capture.output(x_prima <- print(x,...))
  kable(x_prima, ...)
}
tab3 <- CreateTableOne(strata = "NAS_qual", data = df)
kableone(tab3, nonnormal = nonormales, formatOptions = list(big.mark =
 ","))
```

Correlation network

Libraries

```
library(qgraph)
library(graphicalVAR)
library(ggplot2)
library(bootnet)
library(plyr)
library(dplyr)
library(reshape2)
library(knitr)
library(magrittr)
library(lavaan)
library(lme4)
library(glmnet)
library(huge)
library(BayesFactor)
library(ltm)
library(depmixS4)
library(corpcor)
library(tidyverse)
```

Visualize all variables names from the database:

```
dput(names(EOM_120522))
```

Select variables we want to run the correlation analysis with:

```
selected_vars <- c("Age", "BMI", "Heart_rate", "TAS", "TAD",
"Pcadera", "Pcintura", "Steatosis_grade", "SAT_area", "VAT_area",
"Glucose", "Insulin", "HOMAIR", "TG", "CHOL", "HDL", "LDL", "VLDL",
"I_ATE", "ALT", "AST", "GGT", "LDH", "FA", "NEFAS", "Hematocrit",
"Hemoglobina", "Hematies", "VCM", "CHCM", "PLAQUETES", "VPM",
"Leucocits", "Limfocits", "Monocits", "Neutrofils", "Eosinofils",
"Basofils", "Albumina", "Adiponectin", "Leptin", "Irisin", "Gal3",
"FGF19", "FGF21")
```

Correlation network

Non-NASH patients

```
Data <- subset(EOM_120522, EOM_120522$NAS_qual == 0)
Data2 <- Data(c("Age", "BMI", "Heart_rate", "TAS", "TAD", "Pcadera",
"Pcintura", "Steatosis_grade", "SAT_area", "VAT_area", "Glucose",
"Insulin", "HOMAIR", "TG", "CHOL", "HDL", "LDL", "VLDL", "ALT", "AST",
"GGT", "LDH", "FA", "Albumina", "Adiponectin", "Leptin", "Irisin",
"Gal3", "FGF19", "FGF21"))
corMAT <- cor(Data2, use = 'pairwise.complete.obs', method =
"pearson")
#Variable classification in order to visualize them in different
colors:
Groups <- c(rep("Antropometrics", 7),
            rep("Histology", 3),
            rep("Biochemical", 14),
            rep("Organokines", 6))
#We optimize the model deleting nonsignificative correlations: FDR
(alpha) < 0.1
corMAT <- make.positive.definite(corMAT)
qgraph(corMAT, graph = "cor", layout = "spring", threshold = "BH",
```

```
       sampleSize = nrow(Data2), alpha = 0.1, groups = Groups,
       theme = "Borkulo", palette = 'pastel', minimum = "sig")
```

Uncertain NASH

```
Data <- subset(EOM_120522, EOM_120522$NAS_qual == 1)
Data2 <- Data(c("Age", "BMI", "Heart_rate", "TAS", "TAD", "Pcadera",
"Pcintura", "Steatosis_grade", "SAT_area", "VAT_area", "Glucose",
"Insulin", "HOMAIR", "TG", "CHOL", "HDL", "LDL", "VLDL", "ALT", "AST",
"GGT", "LDH", "FA", "Albumina", "Adiponectin", "Leptin", "Irisin",
"Gal3", "FGF19", "FGF21"))
corMAT <- cor(Data2, use = 'pairwise.complete.obs', method =
"pearson")
#Variable classification in order to visualize them in different
colors:
Groups <- c(rep("Antropometrics", 7),
            rep("Histology", 3),
            rep("Biochemical", 14),
            rep("Organokines", 6))
#We optimize the model deleting nonsignificative correlations: FDR
(alpha) < 0.1
corMAT <- make.positive.definite(corMAT)
qgraph(corMAT, graph = "cor", layout = "spring", threshold = "BH",
       sampleSize = nrow(Data2), alpha = 0.1, groups = Groups,
       theme = "Borkulo", palette = 'pastel', minimum = "sig")
```

NASH

```
Data <- subset(EOM_120522, EOM_120522$NAS_qual == 2)
Data2 <- Data(c("Age", "BMI", "Heart_rate", "TAS", "TAD", "Pcadera",
"Pcintura", "Steatosis_grade", "SAT_area", "VAT_area", "Glucose",
"Insulin", "HOMAIR", "TG", "CHOL", "HDL", "LDL", "VLDL", "ALT", "AST",
"GGT", "LDH", "FA", "Albumina", "Adiponectin", "Leptin", "Irisin",
"Gal3", "FGF19", "FGF21"))
corMAT <- cor(Data2, use = 'pairwise.complete.obs', method =
"pearson")
#Variable classification in order to visualize them in different
colors:
Groups <- c(rep("Antropometrics", 7),
            rep("Histology", 3),
            rep("Biochemical", 14),
            rep("Organokines", 6))
#We optimize the model deleting nonsignificative correlations: FDR
(alpha) < 0.1
corMAT <- make.positive.definite(corMAT)
qgraph(corMAT, graph = "cor", layout = "spring", threshold = "BH",
       sampleSize = nrow(Data2), alpha = 0.1, groups = Groups,
       theme = "Borkulo", palette = 'pastel', minimum = "sig")
```

Regression

Logistic Regression

```
We did a logistic regression to see if we can separate the Control group
from the EOMs only with the organokines.

library(readxl)
library(mice)
EOM_I_CONTROLS_2 <- read_excel("C:/Users/irene/Desktop/TFG/Coses TFG/EOM
I CONTROLS 2.xlsx")
EOM_I_CONTROLS_3<-mice(EOM_I_CONTROLS_2)
EOM_I_CONTROL_4<-complete(EOM_I_CONTROLS_3)
write.csv(EOM_I_CONTROL_4, "ROC.csv", row.names=FALSE)
library(dplyr)
numr<- c("Adiponectin", "Leptin", "Irisin", "Gal3", "FGF19","FGF21")
EOM_I_CONTROL_4 <- EOM_I_CONTROL_4%>% mutate_at ( (numr), ~ ( scale (.)%>
% as.vector ))
regresio<-glm(EOM_CONTROL~Adiponectin+Leptin+Irisin+Gal3+FGF19+FGF21, dat
a= EOM_I_CONTROL_4, family=binomial)
summary(regresio)
library(questionr)
odds.ratio(regresio)
```

Odds ratio visualization

```
library(sjPlot)
library(sjlabelled)
library(sjmisc)

plot_model(regresio, show.values = TRUE, value.offset = .2,  title = "Odd
s ratio for control vs EOM")
```

Explanation

0 is control 1 is EOM FGF21: raising one point of the des vest (because we have standardised) implies multiplying by 1.81 the probability of being EOM.

Radar chart

Radar chart

```
We downloaded the appropriate library and performed the median of each or
ganokine. We created a table where we had the median of each organokine i
n each state of the disease ( 0, 1, 2 ) and the maximum and minimum ratio
for each one that we wanted to represent. We  made the radar chart of NAS
H, non-NASH and uncertain NASH.library(readxl)
RADARCHART <- read_excel("C:/Users/irene/Desktop/TFG/Coses TFG/RADARCHART
.xlsx")

library(fmsb)

radarchart(RADARCHART,
           cglty = 1,       # line type of the grid
           cglcol = "gray", # Color grid
           pcol = 2:4,      # color for each line
           plwd = 2,        # width for each line
           plty = 1)        # line type
```