

# Dietary fatty acids and liver fat accumulation



Josefina María Orliacq  
Green Templeton College  
University of Oxford  
Michaelmas Term 2024

A thesis submitted for the Degree of  
*Doctor of Philosophy*

To my family

## **I. Abstract**

Metabolic dysfunction-associated steatotic liver disease (MASLD), defined as excessive liver fat in the presence of  $\geq 1$  cardiometabolic factors (obesity, diabetes, hypertension or dyslipidaemia), affects 25-30% of the global population. Epidemiological evidence on how different dietary fatty acids are associated with MASLD and with overall liver fat content is unclear. This thesis investigated the associations between saturated (SFAs), polyunsaturated (PUFAs), and monounsaturated (MUFAs) fatty acids, assessed by both dietary questionnaires and blood measurements, with liver fat outcomes. In addition, it looked at the associations between liver steatosis and very low-density lipoprotein (VLDL) triglycerides (TG) fatty acid composition.

In prospective analyses, associations between dietary fatty acids intake and liver fat content were explored among 9,268 participants from the UK Biobank. The associations between dietary fatty acids intake and MASLD were assessed in 12,031 participants from the same cohort. In observational analyses of 7,206 participants, associations between plasma fatty acids and liver fat were explored. In an additional cross-sectional analysis of 125 participants, the associations between liver steatosis and the fatty acid composition of plasma VLDL-TG were studied.

A 5% increase in SFAs dietary intake was associated with higher liver fat content (+5.6 % difference in liver fat geometric mean [CI: 3.3%-8.0%]), while a 5% increase in PUFAs dietary intake was associated with lower liver fat content (-4.7% [-8.3% to -0.9%]). MUFAs dietary intake showed non-significant associations. In subgroup analyses, associations between SFAs intake and liver fat were stronger in men, and in participants with higher visceral adipose tissue. Stearic and oleic acid exhibited the strongest associations with liver fat content.

Higher SFAs and higher MUFAs intake were associated higher odds of MASLD (OR: 1.25 [1.12–1.40];  $p < 0.001$ ) and OR: 1.25 [95% CI: 1.12–1.40];  $p < 0.001$ ), respectively). No significant associations were found between dietary PUFAs and MASLD. Stearic and palmitic acid exhibited the strongest associations with MASLD.

Higher plasma SFAs and MUFAs % were both associated with higher levels of liver fat content, and higher plasma PUFAs % were associated with lower levels of liver fat content. The associations observed remained similar when analyses were restricted to fasting participants. Liver steatosis was associated with high VLDL-TG SFAs% after adjusting for age and sex, but the association became non-significant after further adjustment for body mass index, and no significant associations were observed for PUFAs or MUFAs.

While the results for PUFAs and MUFAs were less consistent, SFAs presented positive associations with liver fat outcomes across analyses. The results suggest that SFAs, regardless of individual fatty acid, play a harmful role in relation to liver fat and MASLD. These findings reinforce current UK dietary guidelines, which recommend reducing SFAs intake (<10% of energy) due to risks of cardiovascular illness, further supporting these recommendations with the observed associations between SFAs and liver fat outcomes. Further research into the mechanisms linking dietary fatty acids and MASLD is needed to confirm the most effective dietary interventions that could reduce the prevalence and severity of this condition.

## **II. Abbreviations**

CVD: cardiovascular disease

FA: fatty acids

HDL: high-density lipoprotein

IHTG: intrahepatic triglyceride

LDL: low-density lipoprotein

MASLD: metabolic-dysfunction associated steatotic liver disease

MetALD: metabolic and alcohol associated steatotic liver disease

MUFAs: monounsaturated fatty acids

OCDEM: Oxford Centre of Diabetes Endocrinology and Metabolism

OR: odds ratio

PDFF: proton density fat fraction

PUFAs: polyunsaturated fatty acids

SFAs: saturated fatty acids

SLD: steatotic liver disease

TG: triglycerides

VLDL: very low-density lipoprotein

VLDL-TG: very low-density lipoprotein triglycerides

kJ: kilojoules

### **III. Declaration**

All the work presented in this thesis is my own unless otherwise stated. The dissertation I am submitting has not been submitted, either partially or in full, for any other qualification at the University of Oxford or at any other institution. My work was supervised by Dr Jennifer Carter, Dr Siôn Parry, Dr Aurora Pérez-Cornago, and Professor Leanne Hodson. With their comments and suggestions, I wrote all the chapters, conducted all literature reviews, carried out all statistical analyses, and produced all Tables and Figures.

For Chapters 3 and 4 I have adapted code developed by Dr Rebecca Kelly and Dr Heather Young in the Cancer Epidemiology Unit, who developed the Individual Fatty Acid dataset, and for calculating dietary exposures. I consulted Dr Isobel Barnes (senior statistician, NDPH Cancer Epidemiology Unit, Oxford) about methods for the analyses carried out in Chapter 3. The analyses carried out in Chapters 3 and 4 follow methods developed for a publication carried out in NDPH, for which I received suggestions from my supervisors, from Dr Dimitrios A Koutoukidis, and from the journal reviewers.

All data from the UK Biobank were collected and processed by the UK Biobank team. For the data obtained from the Oxford Centre of Diabetes, the NMR data collection and processing, VLDL assays, and liver fat magnetic resonance imaging analysed in Chapter 5 were conducted by researchers at the Clinical Research Unit at Oxford, and I received the final measurements. Results from this thesis have been published and there is one manuscript in preparation.

- Orliacq J, Pérez-Cornago A, Carter J Associations between dietary fatty acids and liver fat accumulation in the UK Biobank. *Proceedings of the Nutrition Society*. 2024;83(OCE2):E225.

## IV. Acknowledgements

It is very special to feel so much gratitude for so many people I have met in this process. The work I present would not have been possible without the guidance and expertise of my supervisors that led me through this invaluable experience. I am deeply grateful for the support I received from Dr Jennifer Carter, Dr Siôn Parry, Prof Aurora Perez Cornago and Prof Leanne Hodson. Their mentorship made me feel seen and encouraged as I navigated my way.

I am grateful to my academic advisors Dame Valerie Beral, and Dr Jack Satsangi for their help and supportive words. Thank you to NDPH and University of Oxford for believing in me and financing my studies, it has been an incredible privilege. I am grateful to my assessors during my DPhil milestones, Prof David Preiss, Prof Jeremy Tomlinson and Dr Huaidong Du who provided valuable feedback for my project.

Thank you to all the teachers I have had in Sarmiento (especially in English subjects), UCA and Oxford. Thank you to the colleagues and friends I made at NDPH and Green Templeton, especially at GTC Boat Club, where I was immensely happy. Thank you to the amazing people I met in Oxford. I was lucky to meet Trishna, Fabian, Sofia, Tin and Sandrene, fantastic colleagues and friends in this process. My friends, especially Sofi, Ayla, Cata, Sol, Meli and Mercedes, cheering from the distance, thank you for your support. Thank you to Paula M. for being part of my journey.

I am forever indebted to my family, for their unwavering support. My parents and my siblings, *gracias*. I am here today because of you.

I am super grateful to Jorin: it has been an honour and a wonderful gift to be by your side during this process, thank you for every lesson and joy we've experienced together.

I am grateful to God for my health and for this opportunity: *Dominus illuminatio mea*.

## Contents

<b>I. ABSTRACT .....</b>	<b>II</b>
<b>II. ABBREVIATIONS.....</b>	<b>IV</b>
<b>III. DECLARATION .....</b>	<b>V</b>
<b>IV. ACKNOWLEDGEMENTS .....</b>	<b>VI</b>
<b>V. LIST OF TABLES.....</b>	<b>XIII</b>
<b>VI. LIST OF FIGURES.....</b>	<b>XV</b>
<b>1 INTRODUCTION .....</b>	<b>1</b>
1.1 LIVER FAT ACCUMULATION: NAFLD, MASLD AND LIVER STEATOSIS .....	1
1.2 RISK FACTORS FOR EXCESSIVE LIVER FAT ACCUMULATION .....	2
1.3 DIETARY FATTY ACIDS AND SUBTYPES .....	5
1.3.1 <i>Fatty acids</i> .....	5
1.4 EVIDENCE ON DIETARY FATTY ACIDS AND LIVER FAT ACCUMULATION .....	7
1.4.1 <i>Evidence from observational studies</i> .....	7
1.4.2 <i>Evidence from interventional studies</i> .....	9
1.5 CRITICAL APPRAISAL OF EVIDENCE FROM LITERATURE REVIEW .....	10
1.5.1 <i>Inconsistency in current evidence</i> .....	10
1.5.2 <i>Sample sizes and study design</i> .....	10
1.5.3 <i>Dietary fatty acids assessment tools</i> .....	10
1.5.4 <i>Outcome ascertainment: imaging studies and different stages of SLD</i> .....	11
1.5.5 <i>Statistical models, adjustment for confounders and total energy intake</i> .....	11
1.5.6 <i>The role of body mass index and adiposity</i> .....	12
1.5.7 <i>Populations studied: pre-existent conditions, age, sex, genotype risk and ethnic background</i> .....	12



1.6	CURRENT NEED OF NEW EVIDENCE: GAPS IN THE KNOWLEDGE.....	13
<b>2</b>	<b>DATA SOURCES: UK BIOBANK AND OCDEM.....</b>	<b>16</b>
2.1	THE UK BIOBANK.....	16
2.1.1	<i>Recruitment.....</i>	<i>16</i>
2.1.2	<i>Informed consent and right to withdraw .....</i>	<i>17</i>
2.1.3	<i>Assessment visit .....</i>	<i>17</i>
2.1.4	<i>Dietary assessment in person and online .....</i>	<i>18</i>
2.1.5	<i>Validation of the Oxford WebQ.....</i>	<i>20</i>
2.1.6	<i>Dietary fatty acids intake calculation.....</i>	<i>20</i>
2.1.7	<i>Plasma fatty acids measured by Nuclear Magnetic Resonance (NMR).....</i>	<i>23</i>
2.1.8	<i>Liver fat assessment.....</i>	<i>23</i>
2.1.9	<i>Inclusion and exclusion criteria for data sources .....</i>	<i>25</i>
2.1.10	<i>Covariates.....</i>	<i>27</i>
2.2	OCDEM PARTICIPANTS .....	31
2.2.1	<i>Recruitment.....</i>	<i>31</i>
2.2.2	<i>Inclusion criteria .....</i>	<i>32</i>
2.2.3	<i>Liver fat measurement .....</i>	<i>33</i>
2.2.4	<i>Clinical research Unit assessments.....</i>	<i>33</i>
2.2.5	<i>VLDL-TG fatty acid composition measurement .....</i>	<i>33</i>
2.3	SUMMARY OF STRENGTHS AND LIMITATIONS. ....	34
2.4	ETHICAL APPROVAL .....	35
2.5	MISSING DATA .....	35
<b>3</b>	<b>DIETARY FATTY ACIDS AND LIVER FAT IN THE UK BIOBANK.....</b>	<b>36</b>
3.1	INTRODUCTION AND AIM .....	36

3.2	METHODS.....	36
3.2.1	<i>Study sample, inclusion and exclusion criteria</i> .....	36
3.2.2	<i>Dietary fatty acids and individual fatty acids assessment</i> .....	38
3.2.3	<i>Liver fat content</i> .....	39
3.2.4	<i>Covariates</i> .....	39
3.2.5	<i>Statistical analysis</i> .....	41
3.2.6	<i>Subgroup analyses by sex, visceral adiposity and menopausal status</i> .....	43
3.2.7	<i>Sensitivity analyses</i> .....	43
3.2.8	<i>Software for statistical analyses</i> .....	43
3.3	RESULTS.....	44
3.3.1	<i>Participants characteristics</i> .....	44
3.3.2	<i>Associations between dietary fatty acids and liver fat content</i> .....	46
3.3.3	<i>Associations between individual dietary fatty acids and liver fat content</i> .....	48
3.3.4	<i>Subgroup analyses</i> .....	51
3.3.5	<i>Sensitivity analyses</i> .....	54
3.4	DISCUSSION.....	57
3.4.1	<i>Main findings</i> .....	57
3.4.2	<i>Strengths and limitations</i> .....	58
3.5	CONCLUSIONS .....	60
<b>4</b>	<b>DIETARY FATTY ACIDS AND MASLD IN THE UK BIOBANK.....</b>	<b>62</b>
4.1	INTRODUCTION AND AIM .....	62
4.2	METHODS.....	62
4.2.1	<i>Study sample, inclusion, and exclusion criteria,</i> .....	62
4.2.2	<i>Steatotic liver disease (SLD) and categories</i> .....	63
4.2.3	<i>Dietary fatty acids and individual fatty acids assessment</i> .....	65

4.2.4	<i>Covariates</i> .....	66
4.2.5	<i>Statistical analysis</i> .....	66
4.2.6	<i>Subgroup analyses</i> .....	66
4.2.7	<i>Sensitivity analysis: <math>\geq 4</math> WebQ</i> .....	66
4.3	RESULTS.....	67
4.3.1	<i>SLD prevalence and final sample participant characteristics</i> .....	67
4.3.2	<i>Dietary fatty acids and MASLD</i> .....	70
4.3.3	<i>Individual dietary fatty acids and MASLD</i> .....	72
4.3.4	<i>Subgroup analysis</i> .....	74
4.3.5	<i>Sensitivity analysis</i> .....	74
4.4	DISCUSSION.....	77
4.4.1	<i>Main findings</i> .....	77
4.4.2	<i>Comparison with previous studies</i> .....	77
4.4.3	<i>Strengths and limitations</i> .....	79
4.5	CONCLUSIONS .....	81
<b>5</b>	<b>PLASMA FATTY ACIDS AND LIVER FAT CONTENT .....</b>	<b>82</b>
5.1	INTRODUCTION.....	82
5.1.1	<i>Previous studies</i> .....	82
5.1.2	<i>VLDL-TG fatty acids</i> .....	83
5.1.3	<i>Aims</i> .....	85
5.2	METHODS.....	85
5.2.1	<i>Study population</i> .....	85
5.2.2	<i>Liver fat assessment</i> .....	86
5.2.3	<i>NMR plasma fatty acids</i> .....	87
5.2.4	<i>VLDL-TG fatty acid composition measurement at OCDEM</i> .....	87

5.2.5	<i>Markers of cardiovascular risk at OCDEM</i>	88
5.2.6	<i>Covariates</i>	88
5.2.7	<i>Statistical analyses</i>	89
5.3	<b>RESULTS</b>	91
5.3.1	<i>Participants characteristics</i>	91
5.3.2	<i>NMR plasma fatty acids and liver fat in UK Biobank</i>	94
5.3.3	<i>Sensitivity analysis in fasting group</i>	96
5.3.4	<i>NMR plasma fatty acids and liver fat in OCDEM</i>	97
5.3.5	<i>Comparison between OCDEM and UK Biobank</i>	99
5.3.6	<i>Associations between liver steatosis and VLDL-TG fatty acid composition</i>	101
5.3.7	<i>Associations between liver steatosis and VLDL-TG individual fatty acid composition in OCDEM</i>	102
5.3.8	<i>Associations between steatosis and markers of cardiovascular risk in OCDEM</i>	103
5.4	<b>DISCUSSION</b>	104
5.4.1	<i>Main findings</i>	104
5.4.2	<i>Comparison with previous studies</i>	105
5.4.3	<i>Strengths</i>	107
5.4.4	<i>Limitations</i>	107
5.5	<b>CONCLUSIONS</b>	108
<b>6</b>	<b>DISCUSSION</b>	<b>110</b>
6.1	<b>SUMMARY OF FINDINGS</b>	110
6.2	<b>FINDINGS CONTRIBUTION TO CURRENT LITERATURE</b>	111
6.2.1	<i>Positive associations between dietary SFAs and plasma SFAs with liver outcomes across the thesis</i>	111

6.2.2	<i>Dietary and plasma PUFAs associations with liver fat outcomes across the analyses .....</i>	<i>112</i>
6.2.3	<i>Individual dietary fatty acids across analyses.....</i>	<i>113</i>
6.2.4	<i>Steatosis, VLDL-TG fatty acid composition .....</i>	<i>113</i>
6.2.5	<i>The role of menopause status .....</i>	<i>114</i>
6.2.6	<i>Sexual dimorphism .....</i>	<i>114</i>
6.2.7	<i>The role of BMI and adiposity across analyses.....</i>	<i>115</i>
6.2.8	<i>The role of alcohol intake across analyses.....</i>	<i>116</i>
6.3	METHODOLOGICAL CONSIDERATIONS .....	118
6.3.1	<i>Sample characteristics (comparison across samples, size, power, chance findings)</i>	<i>118</i>
6.3.2	<i>Ethnicity studied .....</i>	<i>118</i>
6.3.3	<i>Selection bias.....</i>	<i>119</i>
6.3.4	<i>Dietary intake assessment .....</i>	<i>120</i>
6.3.5	<i>Liver fat outcomes .....</i>	<i>121</i>
6.3.6	<i>Reverse causality.....</i>	<i>122</i>
6.4	NEW EVIDENCE FOR CURRENT DIETARY GUIDELINES.....	124
6.5	RECOMMENDATIONS FOR FUTURE RESEARCH.....	125
6.6	CONCLUSION .....	126
	<b>REFERENCES.....</b>	<b>128</b>

## V. List of Tables

Table 1.1 Evidence on dietary fatty acids and liver fat outcomes from observational studies..	8
Table 1.2 Evidence on dietary fatty acids and liver fat from interventional studies .....	9
Table 2.1 Individual fatty acids nomenclature and examples.....	22
Table 2.2 Exclusion criteria and data sources.....	26
Table 2.3 Data-fields from the UK Biobank dataset used for variables included in the analyses. .....	29
Table 2.4 Studies from which OCDEM participants sample were recruited.....	32
Table 3.1 Participant characteristics by lowest (Q1) and highest (Q5) quintiles of intake of fatty acids. N=9,268. ....	45
Table 3.2 Results obtained from sensitivity analyses restricted to participants who answered 4 to 5 WebQ (N=2,828) compared to results obtained from main sample .....	55
Table 3.3 Sensitivity analysis. Results obtained from fully adjusted models examining participants with lower alcohol consumption (N=5,477) <sup>1,2</sup> , compared with results obtained from main sample .....	56
Table 4.1 Variables used to define MASLD.....	65
Table 4.2. Participant characteristics by SLD categories. N=13,031. ....	68
Table 4.3 Participant characteristics by MASLD status. ....	69
Table 4.4 OR of MASLD per 5% of dietary fatty acids: subgroup analyses and tests for interaction .....	74
Table 5.1 Studies investigating plasma fatty acids and liver fat.....	83
Table 5.2. Characteristics of UK Biobank participants by steatosis.....	92
Table 5.3. OCDEM participant characteristics by steatosis. ....	93
Table 5.4 OR of high VLDL-TG individual fatty acids in participants with steatosis in OCDEM (N=125).....	102

Table 5.5 Odds ratio (OR) of high TG in participants with steatosis, by sex, adjusted for age, and BMI (N=125) .....	103
--	-----

Table 5.6 OR of high levels of markers of cardiovascular risk in participants with steatosis vs. controls with sequential adjustment for confounders (N=125).....	104
---	-----

## VI. List of Figures

Figure 1.1 Steatosis spectrum .....	2
Figure 1.2 Key intrahepatic pathways of fatty acid metabolism, intake and disposal.....	4
Figure 1.3 Examples of SFAs, MUFAs and PUFAs, from Tvrzicka et al.....	6
Figure 1.4 Summary of population, exposures, and outcomes studied in this thesis. ....	15
Figure 2.1 Section of the Oxford WebQ assessment. ....	19
Figure 2.2. Timeline of UK Biobank data collection relevant to this study's exposures and outcomes. ....	20
Figure 3.1 Flowchart of UK Biobank showing exclusion of participants included in the analyses.....	37
Figure 3.2 Liver fat geometric mean % difference per 5% increase in dietary fatty acids, in minimally adjusted models and by addition of confounders. ....	47
Figure 3.3 Liver fat geometric mean % difference per 5% increase in dietary fatty acids, in fully adjusted models. ....	48
Figure 3.4 Geometric mean % difference in liver fat, per 1SD increase in individual fatty acids intake.....	50
Figure 3.5 Subgroup analyses by adiposity. Linear regression models were used to calculate the geometric mean % difference in liver fat, by 5% increase in dietary fatty acids intake....	51
Figure 3.6 Subgroup analyses by sex. Linear regression models were used to calculate the geometric mean % difference in liver fat, by 5% increase in dietary fatty acids intake.....	53
Figure 3.7 Sensitivity analysis. Analysis was restricted to participants who answered 4 to 5 WebQ. ....	54
Figure 4.1 Flow chart with exclusion criteria. ....	63
Figure 4.2 Classification of steatotic liver disease.....	64



Figure 4.3 Sequential adjustment for confounders (n=12,084). Odds of MASLD per 5% increase in dietary fatty acids.....	71
Figure 4.4 OR of MASLD (N=12,084), per 5% increase in dietary intake. 1,388 cases. ....	72
Figure 4.5 OR of MASLD (N=12,084), per 1SD increase in individual fatty acid intake. 1,388 cases of MASLD. ....	73
Figure 4.6 Sensitivity analysis. OR of MASLD (N=3,649), restricted to those who answered $\geq 4$ WebQs. 424 cases. ....	75
Figure 4.7 Sensitivity analysis, restricted to those who answered $\geq 4$ WebQs. Odds ratio of MASLD (95%CI) per 1SD increase in dietary fatty acids (N=3,649).....	76
Figure 5.1 Very-low density lipoprotein (VLDL) triglyceride fatty acid composition. ....	84
Figure 5.2 Flowchart of UK Biobank showing exclusion of participants included .....	86
Figure 5.3 Summary of exposures and outcomes analysed in this chapter.. ....	88
Figure 5.4. Geometric mean % difference in liver fat, per 1% increase in plasma FAs%: sequential adjustment for confounders (N=7,206), UK Biobank. ....	95
Figure 5.5 Geometric mean % difference in liver fat per 1% increase in plasma fatty acids in the UK Biobank. Sensitivity analysis restricted to participants fasting (N=249). ....	97
Figure 5.6 Geometric mean % difference in liver fat (95%CI) per 1% increase in NMR plasma fatty acids in OCDEM (N=125).....	98
Figure 5.7. Geometric mean % difference in liver fat (95%CI) per 1% increase in NMR plasma fatty acids in UK Biobank (A) and OCDEM (B).....	100
Figure 5.8 Comparison of mean VLDL-TG fatty acid composition in OCDEM participants with and without steatosis, unadjusted N=125.. ....	101
Figure 5.9 OR of high FAs VLDL-TG in participants with steatosis, with sequential adjustment for confounders. N=125.....	102

Figure 5.10 Odds ratio of high levels of biomarkers in participants with steatosis in comparison to controls (N=125). .....	103
---	-----

# 1 Introduction

## 1.1 Liver fat accumulation: NAFLD, MASLD and liver steatosis

The liver acts as a coordinating centre for energy and metabolism, and plays a vital role as an endocrine organ, regulating blood glucose, lipids and albumin levels, and removing toxic substances from the body (1). While it usually contains fat, an excess of adipose tissue can lead to harmful accumulation, disrupting its structure and functions (2). This excess of >5% of liver fat is defined as liver steatosis (3). Since 1850, steatosis has been described and associated with many factors, such as physical activity and diet, but mainly with alcohol intake. In 1980 the term “non-alcoholic” was used by Ludwig for the first time, to refer to patients with steatosis that was not caused by excessive alcohol consumption (4, 5). NAFLD (non-alcoholic fatty liver disease) was, until very recently, used to describe the first step of a spectrum of steatotic liver disease in patients who did not demonstrate excessive drinking behaviours, and it is still used in recently published literature, and in guidelines from the National Institute for Health and Care Excellence (NICE) (3, 6). However, with the introduction of new definitions based on a Delphi global consensus in 2023, the term NAFLD was replaced by a set of definitions under the concept of steatotic liver disease (SLD)(7). These new definitions were included to reduce stigmatising patients, include comorbidities, and to be more specific than NAFLD, which was solely based on steatosis without excessive alcohol consumption, and therefore considered an “exclusion diagnosis”. As most of the literature and published epidemiological studies use the term NAFLD, the evidence on NAFLD will also be included in this thesis. SLD categorises patients with steatosis into five groups according to the main factor that drives excessive liver fat accumulation. MASLD (metabolic dysfunction-associated steatotic liver disease), defined as SLD in the presence of one of five cardiometabolic factors, has been adopted as a replacement of the term NAFLD, to include the role of comorbid conditions, and to improve patient acceptability while moving away from an exclusion diagnosis (7). It is estimated that

NAFLD (now MASLD) affects approximately one fourth of the global adult population and is currently the most rapidly growing contributor to liver mortality and morbidity globally (8, 9). It is estimated to have an economic burden of 35 billion euros each year in Germany, France Italy and the UK combined, which is in line with estimates of the economic burden of diabetes and heart disease (10).

MASLD can progress from simple steatosis to steatohepatitis (steatosis and inflammation of the organ) (Figure 1.1). Up to that point, this is reversible, but if liver fat accumulation persists, it can lead to irreversible conditions such as fibrosis (scars in the liver due to the death of liver cells), and finally to cirrhosis (patterns of complete fibrosis). It has been estimated that SLD will be the first cause of liver transplant by 2030 (11). There is a wide range of diagnostic tools, all of which have different specificity and sensitivity and are usually applied depending on the context of the patient and stage of SLD (6).

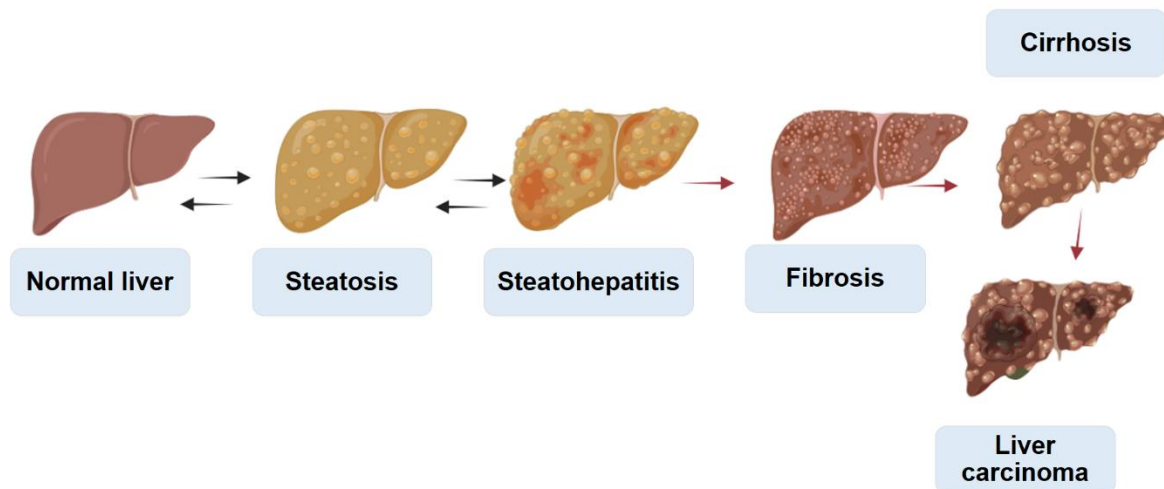


Figure 1.1 Steatosis spectrum Adapted from Wang et. al (12). Created with biorender.com

## 1.2 Risk Factors for excessive liver fat accumulation

Socioeconomic status, type 2 diabetes, and carrying certain genetic variants (e.g. patatin-like phospholipase domain-containing protein 3 (PNPLA-3) and Transmembrane 6 superfamily 2 human (TM6SF2)) are well-established risk factors for NAFLD (13). High total body fat is

also considered a key risk factor: obesity is associated with excessive liver fat accumulation, due to the increase in fat tissue and lipid accumulation within the liver cells (14). However, liver steatosis does not only occur in individuals living with obesity: NAFLD prevalence among non-obese, also known as ‘lean NAFLD’, is estimated to be 7–20% (15). There is currently no pharmacological treatment approved for reducing liver fat outcomes, and medical advice focuses on avoiding sedentarism and hypercaloric diets, which are modifiable risk factors (6).

The pathogenesis of liver steatosis follows a “two hit” model: insulin resistance promotes hepatic steatosis leading to oxidative stress and liver injury, which in turn promotes cell necrosis and inflammatory reactions (3). Imbalances between fatty acids (FA) input and disposal pathways drive fat accumulation, and when the liver exceeds its capacity to remove fatty acids, it stores them as triglycerides (TG) (16).

Fat accumulation in the liver appears to act through impairments in lipid storage and lipolysis (breaking down of adipose tissue into fatty acids into the bloodstream), and certain macronutrients may influence these changes differently (17-19). Excess calorie intake influences metabolic pathways that lead to liver fat accumulation, which explains why hypercaloric diets are associated with NAFLD (19). However, it still has not been fully understood how different macronutrients relate to liver fat accumulation, independently of calorie intake (18). Due to the public health relevance of this illness, and the need for clear evidence to guide dietary advice to prevent it, ongoing research is focusing on dietary macronutrients and whether they independently affect the risk of developing MASLD. Recent studies have focused on the role of different dietary carbohydrates, protein, energy restriction, and fat quality (20, 21). Currently, there is no consistent advice on the types of dietary fat recommended for MASLD prevention (22).

Dietary FA may relate in opposite directions to liver fat accumulation in the liver, possibly due to variations in liver metabolism, and the metabolic pathways that fatty acids are prone to go through (23). Fatty acids can enter the liver from several sources, including from diet and adipose tissue, as detailed in Figure 1.2 (adapted from Hodson et al.) (24). There, they can be esterified into TG and become part of the TG storage pool. Adipose tissue can break down (lipolysis) and release non-esterified fatty acids (NEFAs), which enter the liver. Fatty acids are removed from the liver by either secretion as VLDL-TG or by oxidation, as it can be observed in Figure 1.2 (24).

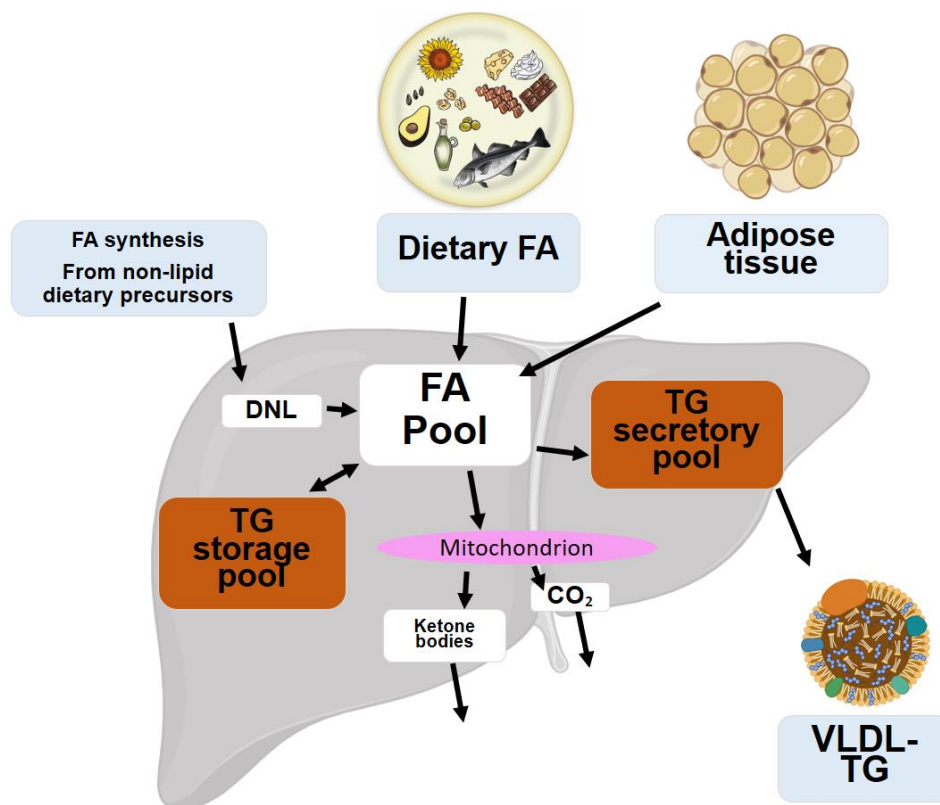


Figure 1.2 Key intrahepatic pathways of fatty acid metabolism, intake and disposal. DNL: de novo lipogenesis. FA: fatty acids. VLDL-TG: very-low density lipoprotein TG: triglycerides

Different fatty acids may relate differently to liver fat, due to differences in the partitioning of fatty acids into oxidation or storage pathways, or different associations with de novo

lipogenesis (DNL), leading to reductions in FA produced in the liver (23, 25-27). However, these observations have not been fully elucidated in human livers.

### 1.3 Dietary fatty acids and subtypes

#### 1.3.1 Fatty acids

Fatty acids are vital components of lipids and cell membranes in the human body. They play a key role in structure and take part in complex metabolic pathways with major biological roles. Fatty acids have a significant role in energy production and metabolism and can be produced endogenously or obtained only through diet, which are termed “essential fatty acids” (28).

Fatty acids are composed of a chain of carbon atoms with hydrogen and oxygen. They are called carboxylic compounds: they contain a carboxyl group ( $-\text{COOH}$ ), which consists of a carbonyl ( $\text{C}=\text{O}$ ) bonded to a hydroxyl ( $-\text{OH}$ ) group. They have a systematic name by International Union of Pure and Applied Chemistry (IUPAC) and a common name. Fatty acids can be classified due to their structure and based on their degree of unsaturation. Unsaturation is a double bond in cis-configuration, which means that those without this double bond are saturated fatty acids, those with one are monounsaturated fatty acids, and those with multiple double bonds are polyunsaturated fatty acids (28).

Figure 1.3 shows three fatty acids abundant in diet, as examples of the three different structure groups mentioned.

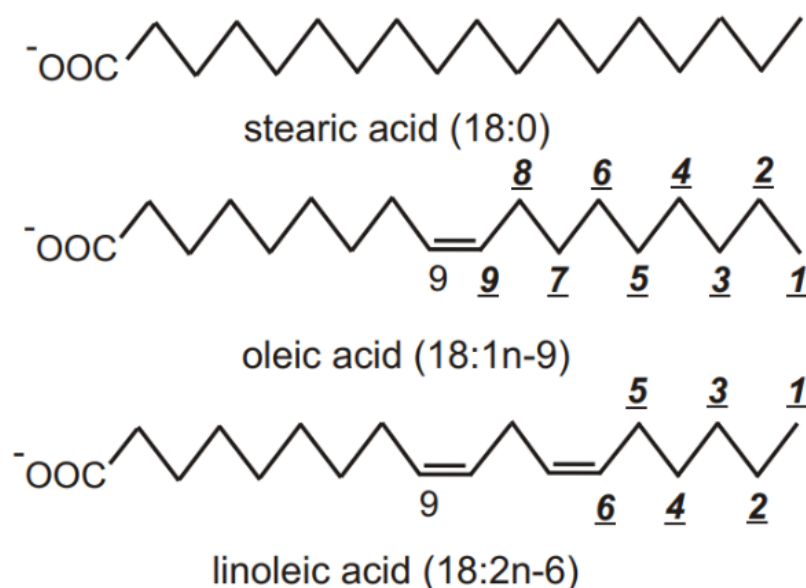


Figure 1.3 Examples of SFAs, MUFAs and PUFAs, from Tvrzicka et al. (28).

Stearic acid has no double bonds and is therefore an SFA. Oleic acid, a MUFA has only one double bond, located in carbon 9 and linoleic acid, a PUFA, contains more than one double bond (28). PUFAs can be sub-classified into n-3 (or omega-3) and n-6 (or omega-6) PUFAs based on the position of the first double bond from the methyl (omega) end of the fatty acid chain. Due to their different molecular structures, they present different fluidity characteristics that explain their different functions in human metabolism (28), and may also contribute to their differing relationship with liver fat. For example, omega-3s may reduce liver fat by enhancing oxidation while excess omega-6s, especially with a high omega-6/omega-3 ratio, may promote hepatic steatosis (28). SFAs can usually be found in fat from animal sources and tropical plant oils (e.g. coconut, palm). Unsaturated FAs can be found in vegetable oils (such as olive oil for MUFAs; sunflower oil and soybean oil for n-6 PUFAs; flaxseed oil for n-3 PUFAs) and marine sources (algae and fish oils).



## 1.4 Evidence on dietary fatty acids and liver fat accumulation

A literature search was carried out in the databases MedLine and Embase, to understand the available evidence in regards to dietary fatty acids and liver fat accumulation in adults. The searches were limited to English language and full text, and for research carried out in humans. Additionally, email alerts were configured to notify of any newly published literature that aligns with the specified search criteria.

### 1.4.1 Evidence from observational studies

Fourteen observational studies were found, with sample sizes that varied from 36 to 32,448 participants (Table 1.1). These studies were heterogeneous in regards to the population background, ethnicity, underlying health conditions and age. In addition, the covariates adjusted for varied, and control for energy intake, a key factor in these associations, was not carried out in six of them. In addition, in 2023 a systematic review and meta-analysis of 28 observational studies looked at dietary fats and NAFLD and found that excessive calorie intake was associated with NAFLD, but there were no significant associations between dietary SFAs, MUFAs, or PUFAs and NAFLD (29). Key limitations included high heterogeneity among included studies, variability in dietary assessment methods, and the predominance of cross-sectional data.

Table 1.1 Evidence on dietary fatty acids and liver fat outcomes from observational studies.

Study	Population	Dietary assessment	Liver fat outcome	Covariates	Findings
Nemer, 2024(30)	US, N=4,175	2 24-hour recall assessments	Steatosis (measured by elastography).	Age, sex, ethnicity, education BMI	SFAs was positively associated with steatosis.
Cheng, 2023(31)	UK, N=12,620	≥1 Oxford WebQ	1. Liver fat % by MRI 2. Steatosis(by MRI)	Age, sex, ethnicity, deprivation, BMI, physical activity, alcohol intake, smoking status, total energy intake	SFAs was positively associated with liver fat and steatosis. MUFAs and PUFAs were NS.
Friden, 2023(32)	UK, N=13,849	≥1 Oxford WebQ	Liver fat % by MRI.	Physical activity, smoking and age	SFAs was positively associated with liver fat. MUFAs and PUFAs were NS.
Tian, 2023 (33)	UK, N=10,623	≥2 Oxford WebQ	Steatosis (by MRI).	Age, sex, BMI, sedentary time, physical activity, income, and education	High PUFAs pattern was inversely associated with liver fat. Other patterns with different FAs were NS.
Cui, 2021 (34)	US, N=6,693	2 24-hour recall assessments	Fatty liver index.	Age, sex, ethnicity, education, smoking, physical activity, hypertension, diabetes, total energy intake, total cholesterol, triglycerides, HDL	n-3 PUFAs and n-6 PUFAs were inversely associated with liver fat.
Lopez Bautista, 2020 (35)	Mexico, N=299	FFQ	Steatosis by ultrasound.	NA	Stearic and linoleic acid intakes were associated with severity.
Noureddin, 2020 (36)	US, N=32,448	FFQ	Chronic liver disease excluding other causes.	BMI, total energy intake, physical activity, soda and coffee intake	No significant associations for SFAs, PUFAs or MUFAs.
Alferink, 2019(37)	Netherlands, N=3,882	FFQ	Steatosis by ultrasound.	BMI, age, sex, smoking, alcohol, physical activity, total energy intake, Dutch Healthy Diet Index, blood cholesterol, metabolic syndrome, diabetes	No significant associations for SFAs, PUFAs or MUFAs.
Vernekar, 2018 (38)	India, N=147	FFQ	Steatosis by ultrasound or biopsy.	NA	No significant associations for SFAs, PUFAs or MUFAs.
Cheng (39)	China, N=36	FFQ	Steatosis by MRI.	Age, BMI, and total energy intake,	SFAs was positively associated with steatosis.
Oya, 2010 (40)	Japan, N=796	FFQ	Steatosis by ultrasound.	BMI, Smoking status, physical activity, menopause status, eating quickly.	EPA and DHA presented inverse associations with liver fat, only in men.
Allard, 2008 (41)	Canada, N=73	Food diary	Steatosis by liver biopsy.	NA	No significant associations for SFAs, PUFAs or MUFAs.
Zelber, 2007 (42)	Israel, N=349	FFQ	Steatosis by ultrasound.	NA	No significant associations for SFAs, PUFAs or MUFAs.
Meisenger, 2019(43)	Germany N=283	FFQ and 24-h recall	Steatosis by MRI.	Age, sex, physical activity, smoking status, alcohol consumption, total energy intake	Total fat intake is associated with liver fat.

FFQ: food frequency questionnaire, MRI: magnetic resonance imaging, SFAs: saturated fatty acids, MUFAs: monounsaturated fatty acids, PUFAs: polyunsaturated fatty acids, NS: non-significant, NA: not applicable, US: United States, UK: United Kingdom.

### 1.4.2 Evidence from interventional studies.

The eight interventional studies presented below in Table 1.2 were performed by recruiting participants from different backgrounds and with different pre-existing conditions, and were carried out on a small scale due to the nature of their design. Researchers chose to compare different fatty acids diets: high SFAs vs. high PUFAs, high SFAs diet vs. high free sugars diet, high SFAs vs. low SFAs, or high MUFAs vs. control. In addition, the energy content of diets was high in some (hyperenergetic) and isoenergetic in others.

Table 1.2 Evidence on dietary fatty acids and liver fat from interventional studies

Author and year	Country	N	Duration (weeks)	Energy	Results in liver fat change (%)
<b>High PUFAs vs. high SFAs</b>					
Bjermo et al., 2012 (44)	Sweden	61	10	Isoenergetic	SFAs: ↑0.3. PUFAs: ↓1.1
Luukkonen et al. (33) (45)	Finland	26	3	Hyperenergetic	SFAs: ↑2.7* PUFAs: ↑0.7
Rosqvist et al., 2014(46)	Sweden	39	7	Hyperenergetic	SFAs: ↑0.5* PUFAs: ↑0.04, NSD
Rosqvist et al., 2019 (47)	Sweden	61	8	Hyperenergetic	SFAs: ↑1.5* PUFAs: ↓0.1
<b>High SFAs vs. high free sugars</b>					
Parry et al., 2020 (48)	UK	16	8	Isoenergetic	SFAs: ↑2 Free Sugars: NSD
<b>High SFAs vs. low SFAs</b>					
Utzschneider et al., (49)	USA	20		Isoenergetic	Low-SFAs: ↓2.2* High SFAs NSD
<b>High MUFAs vs. control</b>					
Bozzeto et al. (50)	Italy	17	8	Isoenergetic	MUFAs: ↓2.2* Control: ↓1.6.
<b>High PUFAs vs. control</b>					
Green et al., 2020(25)	UK	38	8	Isoenergetic	Omega3- PUFAs ↓ No controls for this outcome (secondary outcome)

IHTG: intrahepatic triglyceride \*statistically significant, SFAs: saturated fatty acids, MUFAs: monounsaturated fatty acids, PUFAs: polyunsaturated fatty acids, US: United States, UK: United Kingdom.

## 1.5 Critical appraisal of evidence from literature review

### 1.5.1 Inconsistency in current evidence

Evidence is not consistent across studies regarding the roles of different types and subtypes of fatty acids in liver fat accumulation. While the harmful role of total SFAs is consistent across many interventional and some observational studies (30, 32, 39, 45, 47), in many observational studies associations were not significant (38, 41, 42). Additionally, some omega 3 PUFAs show a protective role in the dietary trials, which is confirmed in some of the population-based observational studies but not in all of them.

### 1.5.2 Sample sizes and study design

Sample sizes ranged from 36 to 32,448 in observational studies. In the case of the smaller sample sizes, it is plausible that there was not enough power to detect small associations. Most of the observational studies were cross-sectional or case-control studies, and not prospective, which would have helped to understand associations between dietary fat quality and liver fat accumulation over a period of time. This would have reduced the impact of reverse causality and enabled the study of the long-term effect of dietary fat in liver fat. Interventional studies ranged from 16 to 61 participants, and designs varied in the choice of diets compared, including their total energy intake, methods of delivery, and fatty acids composition.

### 1.5.3 Dietary fatty acids assessment tools

Four of the observational studies investigated associations between NAFLD and individual dietary fatty acids or  $\text{cn3}$  and  $\text{cn6}$  PUFAs (34, 35, 38, 40), while the other ten focused only on total SFAs, MUFAs or PUFAs. The type and number of dietary assessments carried out varied between studies, which included food diaries, 24hr-recall assessments, or food frequency questionnaires (FFQ). Both Friden et al. and Cheng et al. only used one 24h-recall assessments, and in the case of Allard et al., the use of the food diary represented the most detailed dietary assessment across the studies included in this review (41). Interventional studies appeared to

be well controlled, with precise dietary interventions. However, it is difficult to interpret the results from dietary trials, as these may consist of experimental diets that may not reflect the population's real-life dietary habits, being aware that they are part of a study and adhering to healthier habits than they would usually have.

#### 1.5.4 Outcome ascertainment: imaging studies and different stages of SLD

Different methods have been used to measure liver fat content, which have very different sensitivity and specificity. The observational study with the largest sample size of 32,644 individuals, by Nouredin et al. chose a proxy measurement of NAFLD that is not recommended by current guidelines (based on chronic liver disease by hospital records, without any other evidence or diagnosis of NAFLD), which may have overestimated the outcome and therefore may provide biased results (36). Additionally, the dynamic nature of this disease makes it challenging to compare different stages of NAFLD. Some studies have focused on the early stages of NAFLD, while others studied more advanced conditions, in which fibrosis was already present, as Lopez-Bautista et al. did (35). Some studies (Table 1.1) have used ultrasound to measure steatosis, and this has threshold of around 20% in order to reliably detect steatosis, while magnetic resonance imaging (MRI) can detect very low liver fat levels (51, 52). MRI was used in four of the observational studies. In interventional studies, the outcome was measured precisely, using IHTG% measured by MRI and its changes after diet.

#### 1.5.5 Statistical models, adjustment for confounders and total energy intake.

Some studies presented multivariable logistic regression models, while four presented only univariable associations, without any further adjustment for confounders. However, most of the univariable associations were already non-significant, likely due to their small sample sizes (N=34-73) and their lack of power to identify modest associations. Total energy was considered in all the multivariable models, but it should be highlighted that in the case of Oya et al., an energy adjustment was carried out without including total energy intake, but rather body mass

index and physical activity as markers of energy expenditure (40). This may have introduced bias and altered the strength of the associations, as adjusting for total energy intake derived from the same assessment as the exposure was obtained from provides more accurate estimates compared to adjustments using only body mass index or physical activity (53).

#### 1.5.6 The role of body mass index and adiposity

Body Mass Index (BMI) is a widely used measure of body fat. It is a key variable in understanding the associations between dietary fatty acids and liver fat, as BMI is associated with high liver fat, and also associated with diets high in fat and high in calories (54).

In the case of Alferink et al., inverse associations between MUFAs and NAFLD lost significance after adjustment for BMI (37). In a recent meta-analysis, it was observed that BMI is moderately correlated to liver fat by MRI ( $r: 0.43 [0.41-0.44]$ ) (55). It has been suggested that visceral adipose tissue (VAT) may be more important in the development of MASLD, than subcutaneous adiposity. This is explained by one of the pathological mechanisms in liver steatosis: VAT secretes pro-inflammatory cytokines and adipokines, and releases FAs into circulation, causing dyslipidaemia and systemic insulin resistance (56). In addition, as previously mentioned, MASLD does not only occur in people who are obese and the prevalence of 'lean MASLD' is estimated to be 7–20% (15). Therefore, the addition of BMI in the statistical models is not consistent across the studies.

#### 1.5.7 Populations studied: pre-existent conditions, age, sex, genotype risk and ethnic background

Inclusion criteria for the studies presented encompassed a wide range of metabolic phenotypes, nationalities, diabetic status, metabolic syndrome, or obesity. Populations were very heterogeneous, and in some cases, they were chosen due to convenience sampling, such as an ageing population in the case of Alferink et al (37). While the studies reviewed in this thesis

focused adults overall, the metabolism may not be the same for the elderly population in comparison to the younger adults.

## 1.6 Current need of new evidence: gaps in the knowledge

The lack of consistency in the associations between dietary fatty acids and liver fat in many studies could be due to low power in the small sample sizes of some of the observational studies, to measurement error in diet and in the liver fat outcomes, the diversity of the populations studied, or a combination of these factors. Therefore, the role of different types of fatty acids, independent of calorie intake, remains unclear. As such, there is a need for prospective studies which may help reduce reverse causality and understand the long-term effect of fat quality in liver fat. Adequate measurements of MASLD are also required, as diverse methods in current evidence make it difficult to compare results between studies. Preferably, the use of MRI to measure liver fat content would be employed, as this method demonstrates the highest specificity (compared to ultrasound). Additionally, there is a challenge in properly measuring real-life diet in large populations, while including an adjustment for total energy intake. It would be helpful to study different types of fatty acids, with multiple dietary assessments in different opportunities, which would allow acceptable estimations of the usual intake of these exposures. This would be particularly important in the case of PUFAs, as their main sources are foods which in the UK may only be consumed approximately one portion per week. The use of BMI to estimate total adiposity should be carefully considered, as well as the current increase in lean MASLD, which makes it difficult to decide whether this variable should be adjusted for or should be considered. Different measurements of body fat, such as MRI or DEXA, could give a more precise estimate of overall adiposity, and further studies should examine potential effect modification between men and women. This would provide with more generalisable evidence about the associations between dietary fatty acids and liver steatosis.

Ensuring the use of at least two dietary assessments, as well as detailed exclusions, control for BMI and for alcohol intake can provide answers to this research question with precise methods. This may help provide a clear picture of the associations between these nutrients and liver fat accumulation, and if these associations are significant, attempt to estimate their strength.

Therefore, the aims of this thesis are:

- To study the associations between dietary fatty acids and liver fat in the UK Biobank
- To study the associations between dietary fatty acids and metabolic dysfunction-associated steatotic liver disease (MASLD) in the UK Biobank
- To study the associations between plasma fatty acids and liver fat in participants from the UK Biobank and from studies from the Oxford Centre of Diabetes Endocrinology and Metabolism (OCDEM)
- To study the associations between liver steatosis and very low-density lipoprotein-triglyceride (VLDL-TG) fatty acids composition in participants from studies from OCDEM
- To study the associations between liver steatosis and markers of cardiovascular risk in participants from studies from OCDEM.

In addition, this thesis will assess the potential effect modification of sex, visceral adipose tissue, and menopausal status on these associations.

The summary of the population, exposures and outcome studied in this thesis are outlined in

Figure 1.4.



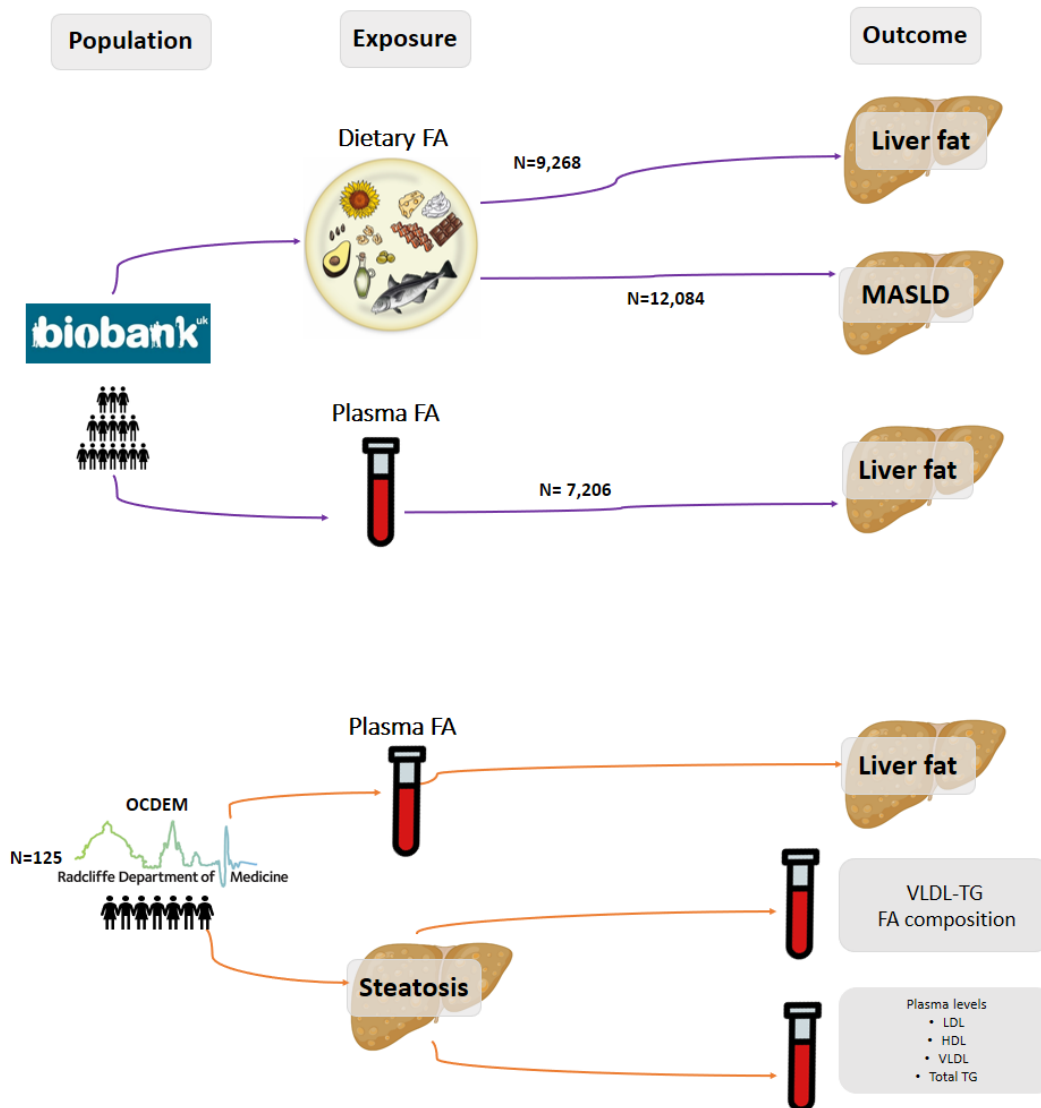


Figure 1.4 Summary of population, exposures, and outcomes studied in this thesis.

## **2 Data sources: UK Biobank and OCDEM**

### **2.1 The UK Biobank**

This study used data from the UK Biobank, a prospective cohort of more than 500,000 participants established to investigate how different exposures relate to various acute and chronic health outcomes. The main goal of the UK Biobank is to understand the determinants of common diseases, enabling scientific discoveries that improve population health (57, 58). This cohort was designed with record linkage to detect common conditions, supported by power calculations to ensure sufficient detection of incidence cases and conditions (59). The UK Biobank has approval from the Northwest Multi-Centre Research Ethics Committee (renewal reference number in November 2021: 21/NW/0157), and this study was conducted under the Nuffield Department of Population Health Cancer Epidemiology Unit application number 67506 (57, 59-61).

#### **2.1.1 Recruitment**

The UK Biobank recruited men and women aged between 40 and 69 years old, from 22 different assessment centres across England, Scotland and Wales between 2006 and 2010. In total, 9,238,453 people who were registered with the UK National Health Service received a postal invitation, with a potential date and recruitment centre that they could attend to if they decided to participate (57). From those, 5.45% responded and attended one of the recruitment centres, resulting in a sample size of 503,317 participants (57). Predicted occurrence of events based on simulations about common non-communicable events, such as diabetes mellitus, coronary heart disease, and breast cancer confirmed that 500,000 participants would be powered enough for analyses, considering more than 10 years of follow-up (59).

### 2.1.2 Informed consent and right to withdraw

Once participants attended their first assessment centre, they electronically signed an informed consent form, and were informed about the use of their data for storage and research, as well as their right to withdraw from the study at any point (62). They received a participant identifier number and their data was pseudo-anonymised. Any new withdrawals were informed by the UK Biobank team, so participants who no longer wished to be part of the research were removed from all analyses.

### 2.1.3 Assessment visit

In an initial assessment visit, after providing consent, participants completed a touch-screen questionnaire, and participated in a computer-assisted interview. Physical and functional measures were taken, and blood, urine and saliva samples were collected (57). A touchscreen questionnaire was included at the beginning of the visit, in which participants completed sections about medical history, socioeconomic status, education, smoking, alcohol consumption, medications, and physical activity (63, 64). These questions were designed taking into account key risk factors for common diseases, and responses were stored for further analyses. To ensure data accuracy, the program flagged implausible values in numeric responses, which made participants review their answers before proceeding to the next question.

After the touchscreen questionnaire, a verbal interview was conducted by a trained interviewer, who considered the responses recorded from the touch screen in case they had to ask more about any diagnoses. The interviewer asked participants about their place of birth, personal and family medical history, and answers were recorded in a computer. Questions addressed any cancer diagnoses as well as an extensive range of non-cancer illness, operations, prescriptions and use of over-the-counter medications (65).

Physical measurements were collected after completing the touch screen questionnaires, interviews and blood pressure measurement, along with ocular measurements. Sitting and standing height were measured, and hip and waist measurements were also recorded. In addition, bio-impedance measurements were assessed (66).

#### 2.1.4 Dietary assessment in person and online

Participants' dietary data were obtained using the Oxford WebQ, a tool developed by the Cancer Epidemiology Unit at the University of Oxford. This dietary recall assessment is a validated web-based 24-hour tool, which was developed for repeated administration in large prospective studies. Unlike typical interviewer-administered 24-hour recall, it is quick to complete (it takes 10-20 minutes) and it can be done completely online (67, 68). It was chosen because of its reduced administrative costs, the convenience for participants, and the presence of an algorithm that automatically calculates nutrients based on the food products and amounts chosen. Data about dietary intake of fat and fatty acids was obtained from these 24-hr dietary assessments, in which participants reported the products consumed the previous 24 hours from a list of 206 foods and 32 beverages (67, 69). As an example, in Figure 2.1, it can be observed that each question had a detailed description and options to choose the amount consumed. In this section, the questionnaire included a question designed to quantify the amount of bread consumed on the previous day. If a participant selected a sandwich, only the bread component was initially recorded. Subsequent questions gathered additional details about the sandwich, including the presence and type of spread, its thickness, whether it was store-bought or purchased from a café, and whether it contained ingredients such as egg, cheese, fish, meat, or vegetables. This structured approach ensured that the dietary assessment did not treat a sandwich as a single item but instead considered its individual components for a more precise evaluation of dietary intake. However, it is important to note that the food items in the WebQ are limited, and the estimation of fatty acid content is based on the data from the responses.



Did you eat any bread or crackers yesterday? ☐ No ☒ Yes [Show Help](#)

*E.g. toast, sandwiches, rice cakes, bread rolls, hotdog roll, crumpets, tortilla wraps.*

Bread	Amount	None	½	1	2	3	4	5	6+
 Sliced bread	Slice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
 Sandwich baguette, ciabatta, panini, or sub	Bread item	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
 Large sandwich bap, stotty, pitta bread	Bread item	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
 Bread roll, bap, burger bun, hotdog roll, bagel	Bread item	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other bread and crackers	Amount	None	½	1	2	3	4	5	6+
 Naan bread	Bread item	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
 Garlic bread	Slice	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
 Crackers, crispbread, rice cakes, corn cakes (e.g. Ryvita)	Biscuit / item	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
 Oat cakes	Biscuit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
 Other bread (e.g. crumpets, tortilla wraps, breadsticks)	Slice / item	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 2.1 Section of the Oxford WebQ assessment.

Participants were invited to complete the Oxford WebQ during the initial assessment visit, and virtually via a link they received in their email in other four different opportunities. Figure 2.2 describes the different points at which dietary data was collected. Participants who attended their first assessment between April 2009 and September 2010 were asked to complete a dietary questionnaire in person (68). Approximately 70,000 participants completed it in clinic during the baseline assessment. In addition, all UK Biobank participants who had provided an email address were invited on four separate occasions. Invitations were sent to participants who provided a valid email, (approximately 306,500 participants) and responses were collected according to the timeline described below. The proportion of participants that responded each email was as follows: 32.8% for the first e-mail invitation, 26.2% for the second, 32.6% for the third done, and 30.5% for the fourth. The four different invitations were sent in different days of the week and months of the year to attempt to capture different diet intakes across the year and that may vary between week and weekdays. For any round of invitations, the respond rate was significantly lower in the summer months ( $p < 0.001$ )(70).

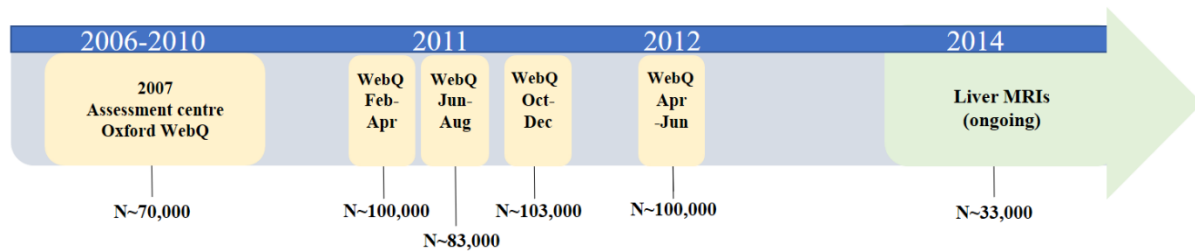


Figure 2.2. Timeline of UK Biobank data collection relevant to this study's exposures and outcomes.

### 2.1.5 Validation of the Oxford WebQ

The Oxford WebQ was validated against blood recovery biomarkers and another 24-hr recall interviewer administered questionnaire, administered over telephone call (71). It was observed that it performs very well to estimate total energy intake, as well as protein, potassium, and total sugar. There were no blood biomarkers for dietary fatty acids, as no recovery biomarker has yet been identified to represent fat intake.

### 2.1.6 Dietary fatty acids intake calculation

Once participants select the options that correspond to what they eat the previous day, replies are coded automatically to obtain estimated nutrient intake, considering the product and amount consumed. From these responses, intakes of total fat, SFAs, MUFAs, PUFAs, n3-PUFAs and n6-PUFAs were automatically calculated and stored as variables in grams. In order to code all dietary variables as percentage of total energy intake, the mean intake of exposures in grams was multiplied by 37.7, to calculate the amount of total kJ provided. This was then calculated as a percentage of each participant's daily total energy intake.

To understand if the associations are driven by specific fatty acids from each group, individual fatty acids dietary intakes were obtained from the UK Biobank Individual Fatty Acids dataset. These were calculated by researchers from the Cancer Epidemiology Unit at the University of Oxford, composed with data from UK McCance and Widdowson's *The Composition of Foods*, and the US Department of Agriculture food composition tables. This provided information about intakes of 21 individual FAs: 10 SFAs, 4 MUFAs and 7 PUFAs (72), that are described

in Table 2.1. McCance and Widdowson's composition of foods integrated dataset 2021 was used to calculate individual fatty acids, as it is the only UK-specific food composition table with individual fatty acid dataset (72). In addition, the US National Nutrient Database (USDA) for standard references, was used as a secondary food composition table, containing more than 8,000 food codes, with a version compiled close to the time frame completion of Oxford WebQ, between 2011 and 2012. The individual fatty acid dataset consisted of 437 unique food codes from the food composition tables: 93% were from McCance and Widdowson, (406 codes), while 7% corresponded to USDA codes (31 codes) (72).

Table 2.1 Individual fatty acids nomenclature and examples.

Notation	Systematic name by IUPAC	Common name	Fatty acid	Food product rich in nutrient (28, 73)
4:0	Butanoic	Butyric	SFAs	Butter
6:0	Hexanoic	Caproic	SFAs	Butter, coconut
8:0	Octanoic	Caprylic	SFAs	Coconut, palm kernel, butter
10:0	Decanoic	Capric	SFAs	Coconut, palm kernel, butter
12:0	Dodecanoic	Lauric	SFAs	Coconut, palm kernel, butter
14:0	Tetradecanoic	Myristic	SFAs	Coconut, palm kernel, butter
15:0	Pentadecanoic	Pentadecylic	SFAs	Red meat, dairy
16:0	Hexadecanoic	Palmitic	SFAs	Red meat, dairy, palm oil
16:1n-7	Hexadecenoic	Palmitoleic	MUFAs	Beef, lamb, pork, salmon, mackerel, Macadamia oil
17:0	Heptadecanoic	Margaric	SFAs	Red meat, dairy
18:0	Octadecanoic	Stearic	SFAs	Chocolate, meat, dairy
18:1n-9	Octadecenoic	Oleic	MUFAs	Olive oil, sunflower seed, pumpkin, sesame, corn, soya, walnut, flaxseed, hemp seed, wheat germ
18:2n-6	Octadecadienoic	Linoleic*	n-6 PUFAs	Olive oil, cultivated rapeseed oil, canola oil, peanut oil
18:3n-3	Octadecatrienoic	Alpha-linoleic*	n-3 PUFAs	Canola oil, cultivated rapeseed oil
18:4n-3	Octadecatetraenoic	Stearidonic	n-3 PUFAs	Hemp oil, fish
20:1n-9	Eicosenoic	Gondoic	MUFAs	Sesame seed oil, walnut, pumpkin seed oil, flaxseed oil
20:4n-6	Eicosatetraenoic	Arachidonic	n-6 PUFAs	Eggs, poultry, beef and dairy
20:5n-3	Eicosapentaenoic	EPA	n-3 PUFAs	Salmon, canned sardines
22:1n-9	Docosenoic	Erucic	MUFAs	Rapeseed and mustard oil
22:5n-3	Docosapentaenoic	DPA	n-3 PUFAs	Fish oil, red salmon
22:6n-3	Docosahexaenoic	DHA	n-3 PUFAs	Grass-fed meat, dairy, fatty fish, shellfish

\*Essential fatty acids.

SFAs: saturated fatty acids, MUFAs: monounsaturated fatty acids, PUFAs: polyunsaturated fatty acids, IUPAC: International Union of Pure and Applied Chemistry.



### 2.1.7 Plasma fatty acids measured by Nuclear Magnetic Resonance (NMR)

Plasma fatty acids were obtained from participant's blood samples collected at baseline. While there were no fasting requirements for participants, the estimated hours of fasting were recorded for each of them. Samples were collected and Nightingale Health Ltd Services were used to analyse biomarker data (74). Nightingale Health is based in Finland and performed analyses based on high-throughput nuclear magnetic resonance (NMR).

Plasma samples collected were prepared on 96-well plates, with DTA (anticoagulant), and shipped to Nightingale Laboratories. Buffers were added to samples, and six 500 MHz proton NMR spectrometers were used to analyse the samples (75). To quantify biomarkers (249 in total, which include lipoproteins, fatty acids, amino acids, ketones and glycolysis metabolites), automated spectral processing software was used. Data were then cleaned and provided to the UK Biobank (76). Quality control flags were identified and recorded if the sample presented values below limit of quantification, were contaminated, degraded, or presented any unidentified small molecules, and those samples were excluded, by using Data-Field 23755 for SFAs, Data-Field 23754 for MUFAs, and Data-Field 23753 for PUFAs (74). The NMR platform demonstrated high reproducibility, with many biomarkers showing coefficients of variation below 5%, real-time monitoring ensuring consistency across spectrometers, and high correlation ( $>0.9$ ) with clinical biochemistry (77).

### 2.1.8 Liver fat assessment

A subsample of this cohort participated in the multimodal imaging studies, which started in 2014 (78). Initially, 5,000 participants were invited to take part in a pilot study in 2014, and as it was feasible, funding for imaging additional 95,000 was confirmed. Then, all participants from the UK Biobank, regardless of their area of residence, were invited to participate in an imaging visit (78). Those who had not responded to the invitation were sent emails 2 weeks, 4 weeks, 6 months and 1 and 2 years after the initial invitation. The goal of 100,000 imaging was

set up based on power calculations, and it is estimated to be achieved as more than 31% of the full cohort responded that they were interested in participating, and by November 2024, 90,000 image studies were completed. The imaging assessments were conducted in centres located in Stockport (referred to as Central), Newcastle-upon-Tyne (North), Reading (South-East), and Bristol (South-West). These locations were chosen to reduce travel time for most participants, considering driving distances and public transportation options. This focus on minimizing travel was essential, as it was identified as a key factor influencing participants' likelihood of attending (78). In order to participate in the study, participants underwent a screening via telephone call to assess if they were eligible, which meant that they did not have any metal implant, pacemaker, or claustrophobia.

Participants in the multimodal imaging study at the UK Biobank Imaging Assessment Centres were studied via Magnetic Resonance Imaging (MRI). All measurements were performed using Siemens 1.5T MAGNETOM Aera scanners, which provided body imaging data from the abdomen, the brain, and the heart (79). MRI is the most accurate non-invasive method to study liver fat (80), and liver fat proton density fat fraction (PDFF) was derived from these multi-slice MRI abdominal scans, which represented liver fat content (%). Proton density refers to the abundance of hydrogen atoms found in the tissue and is estimated as follows:

$$\text{Proton density fat fraction} = \frac{\text{Fat}}{\text{Fat} + \text{Water}}$$

The liver imaging variables were obtained following the IDEAL (iterative decomposition of water and fat with echo asymmetry and least-squares estimation) protocol, which allows the separation of water and fat signals (81). The variables were processed and provided by AMRA (advanced magnetic resonance analytics), using the AMRA® Profiles developed by AMRA Medical AB in Linköping, Sweden (82).

To cover a representative sample of the liver tissue, three 15mm diameter regions of interest (ROIs) were chosen to prevent signals from vessels, bile ducts and other organs. Then the mean of all of the pixels within the 3 ROIs was calculated to obtain the proton density fat fraction.(83). A single transverse slice scan was measured at the porta hepatis to represent the liver, and two sequences were necessary to obtain the data to derive the liver fat variables.

To validate the measurements, phantom studies (studying objects designed to mimic the properties of human tissue instead of participants) were carried out on a different Siemens scanner (1.5T Avanto) at the University of Oxford centre of Clinical Magnetic Resonance Research, and the results collected were similar in both scanners (79).

#### 2.1.9 Inclusion and exclusion criteria for data sources

Participants were excluded from analyses if certain conditions were identified. Participants may have recently been diagnosed with a condition that motivates them to change their diet significantly. Therefore, to reduce reverse causality, participants with prevalent cardiovascular disease, diabetes, and chronic liver disease who may have already received dietary advice were excluded from the analyses. Participants with conditions that can alter liver metabolism significantly, or using cholesterol lowering medication, or drugs that can lead to liver inflammation, were identified and excluded from the final sample. Hospital Episode Statistics were checked in June 2022 to perform exclusions, and checked again in August 2024 to identify any incident cardiovascular and liver disease case that required hospitalisation. Data sources for the exclusion criteria can be found in Table 2.2. For proportions of exclusions, each chapter identifies the number of participants excluded per category.

Table 2.2 Exclusion criteria and data sources.

<b>Exclusion</b>	<b>Description</b>	<b>ICD-10<sup>1</sup></b>	<b>Self-reported Assessment Centre Data</b>
<b>Liver disease</b>	Any alcoholic liver disease	K70	Verbal Interview <sup>2</sup>
	Any toxic liver disease	K71	
	Hepatic failure, not elsewhere classified	K72	
	Chronic hepatitis, not elsewhere classified	K73	
	Autoimmune hepatitis	K75.4	
	Inflammatory liver disease, unspecified	K75.9	
	Fatty liver not elsewhere classified	K 76.0	
	Liver failure/ cirrhosis		
	Infective/viral hepatitis		
	Non-infective hepatitis		
<b>Cardiovascular disease</b>	Hepatitis		Verbal Interview <sup>2</sup>
	Acute myocardial infarction	I21	
	Subsequent myocardial infarction	I22	
	Complications of myocardial infarction	I23	
	Old myocardial infarction	I25.2	
	Dressler's syndrome	I24.1	
	Subarachnoid haemorrhage or infarction	I60	
	Intracerebral haemorrhage	I61	
	Cerebral infarction	I63	
	Stroke, not classified as haemorrhage	I64	
	Stroke		
	Heart attack/myocardial infarction		
	Angina		
	Peripheral vascular disease		
<b>Diabetes</b>	Transient ischaemic attack		Verbal interview
	Subdural haemorrhage		
	Aortic aneurysm rupture		
<b>On cholesterol lowering drugs</b>	Diagnosed (or insulin user) <sup>4</sup>		Touch screen <sup>3</sup> /verbal interview
<b>Medication</b>	Dexamethasone		Verbal interview
	Hydroprednisone		
	Rifampicin		
<b>Pregnancy</b>			Verbal interview
<b>Endocrine disease</b>	Thyrotoxicosis		Verbal interview
	Cushing syndrome		

<sup>1</sup>ICD-10: International classification of diseases and related health problems, by the world health organisation. Obtained from Data field 41270, from Hospital Episode Statistics

<sup>2</sup>All verbal interview diagnoses were obtained from Data-Field 20002: Non-cancer illness, self-reported

<sup>3</sup>From Data-Field 6150: vascular/heart problems diagnosed by doctor

<sup>4</sup>From Data-field 6177 and 6153: medication for cholesterol, blood pressure or diabetes

#### 2.1.10 Covariates

Covariates were chosen after reviewing current literature to identify factors that may contribute to both liver fat accumulation and dietary habits, and the data fields used to code them can be found in Table 2.3. The decision to exclude directed acyclic graphs was based on the complexity of dietary exposures and confounders, and the lack of consensus on a definitive causal structure for liver fat accumulation and MASLD.

Ethnicity was collected from the touchscreen data at baseline described in the assessment centre visit. This was coded as ‘White’ for those who reported British, Irish, or any other white background, ‘Mixed’ for those who selected the option White and Black Caribbean, White and Black African, White and Asian, or any other mixed background, ‘Asian or Asian British’, for those who indicated Indian, Pakistani, Bangladeshi or any other Asian background, and ‘Black or Black British’ for those who selected Caribbean, African, or any other Black background). Other or missing categories were also included in the analyses (84).

To estimate deprivation, the Townsend index was calculated for each participant from responses collected from the touch screen, and based on preceding national census output areas. The census variables utilised in the Townsend index include ‘economic activity’, to estimate the percentage of people who are economically active and unemployed, ‘persons per room’, to estimate the percentage of households that are overcrowded, with more occupants than there are rooms, ‘car or van availability’ to estimate the percentage of households that do not own a car or van, and ‘tenure’, to estimate the percentage of households that are not owner-occupied (85).

Participants region was determined based on the assessment centre each participant attended, resulting in 10 categories. St Bartholomew’s Hospital, Hounslow, and Croydon were grouped as “London”; Swansea, Wrexham, and Cardiff as “Wales”; Stockport, Manchester, Liverpool,

and Bury as “North-West England”; Newcastle and Middlesbrough as “North-East England”; Leeds and Sheffield as “Yorkshire and the Humber”; Stoke and Birmingham as “West Midlands”; Nottingham was renamed “East Midlands”; Reading and Oxford as “South-East”; Bristol as “South-West”; and Glasgow and Edinburgh as “Scotland.”

Only data from participants who attended the MRI assessment centres were included in this study. Consequently, the regional category was restricted to centres where the imaging study was conducted, including only North-West England, North-East England, Yorkshire and the Humber, and East Midlands, thereby excluding participants from Wales and other regions of England. To classify participant’s physical activity, data from the touchscreen questionnaire at baseline was used, which described time spent daily walking, doing moderate physical activity, and vigorous physical activity. Using these variables, metabolic equivalent (MET)-hours per week were calculated. MET values are 3.3 for walking, 4.0 for moderate physical activity, and 8.0 for vigorous physical activity. Participants were then categorised in physical activity groups according to total MET hours value per week (low, [ $<10$  MET hours/week], moderate [ $10\text{--}50$  MET hrs/wk.], high [ $>50$  MET hrs/wk.]), following the guidelines for data processing and analysis of the International Physical Activity Questionnaire (86).

Smoking status was categorised using participant self-reported information at baseline, into ‘Never’, ‘Current’, or ‘Previous’ tobacco smokers. Education was categorised according to the highest degree obtained from those specified in the touchscreen questionnaire at baseline. The categories were ‘College or University degree or other professional qualifications’, ‘Exams at ages 17 to 18 years (A levels/AS levels or equivalent)’, ‘Exams at 16 years (O levels/GCSEs/CSEs)’, ‘none of the above’ or ‘Unknown’.

Table 2.3 Data-fields from the UK Biobank dataset used for variables included in the analyses.

	<b>Data-field</b>	<b>Additional information</b>
<b>Outcome</b>		
Liver fat fraction	24352	Processed by AMRA.
Diabetes	6177 and 6153	Medication for cholesterol, blood pressure or diabetes
Cholesterol lowering medication	20002 6177 and 6153	Verbal interview: hypertension Medication for cholesterol, blood pressure or diabetes
Alcohol intake		Touch screen questionnaire: 'white wine', 'red wine', 'beer', 'spirits'
	4407	Average monthly red wine intake
	4418	Average monthly champagne plus white wine intake
	4429	Average monthly beer plus cider intake
	4440	Average monthly spirits intake
	4462	Average monthly intake of other alcoholic drinks
	1568	Average weekly red wine intake
	1578	Average weekly champagne plus white wine intake
	1588	Average weekly beer plus cider intake
	1598	Average weekly spirits intake
	1608	Average weekly fortified wine intake
	4462	Average monthly intake of other alcoholic drinks
	1568	Average weekly red wine intake
<b>Confounders</b>		
Sex	31	
Age at recruitment	21022	
Ethnic background	21000	
Qualifications	6138	If participants selected more than one, the highest was used to estimate qualification
Physical activity	864 874 884 894 904 914	Number of days/week walked 10+ minutes Duration of walks Number of days/week of moderate physical activity 10+ minutes Duration of moderate activity Number of days/week of vigorous physical activity 10+ minutes Duration of vigorous activity
Townsend deprivation index	189	
Region	54	UK Biobank assessment centre
Smoking status	20116	
Body mass index (BMI)	23104	Imaging assessment centre
Fruits and vegetable intake	26123 26065 26115 26090 26089 26098 26125 26146 26093 26091	Raw salad Allium vegetables Peas and corn Berries Apples and pears Green leafy vegetables Root vegetables Other vegetables Other fruits Citrus fruits

AMRA: advanced magnetic resonance analytics,

Vegetable and fruit intake was calculated by combining the total intake in grams of raw salad, leafy greens, root vegetables, tomatoes, allium vegetables, other vegetables, peas, corn, citrus fruit, berries, apples, pears, and other fruit variables from the WebQ data, and calculating a mean in grams/day. This was obtained to adjust for a variable that could be a proxy marker of a healthy diet.

Body mass index (BMI) was calculated by ratio of participant's weight in kilograms divided by the square of the height in meters ( $\text{kg/m}^2$ ). Following the international classification for adults aged over 20 years (87), participants were categorised in groups as <25, 25-29.9, 30-34.9, and  $\geq 35 \text{ kg/m}^2$ . Due to the small number of participants with a BMI <18, defined as 'underweight', this category was collapsed into the normal weight category. The same BMI categories were applied to all participants in the analysis, regardless of ethnicity. BMI was measured at the time of MRI to ensure that adjustments for BMI were performed concurrently with liver fat assessment. As BMI may be part of the causal pathway but is not yet confirmed, it was added as the last confounder in the models, to allow assessment of results before and after its addition.

Alcohol intake was measured using data from the touchscreen questionnaire, in which there was information about the frequency and type of alcoholic drinks consumed. This enabled quantification of total alcohol grams per drink, and this was transformed into Units to meet the standard measurements and recommendations from NICE guidelines in the UK (88).

The rationale for including menopause in the analyses that will be carried out in the UK Biobank was to consider the metabolic changes caused by oestrogen level changes that occur during that period, as previous research has observed more cases of liver steatosis in post-menopausal women (89, 90). The definition of menopause is 12 months without a menstrual period for women older than 45, or 24 months and blood hormone tests for women younger



than 45 (91). In this thesis, menopause status was defined as ‘Yes’ if women reported it in the assessment visit, had bilateral oophorectomy or hysterectomy, or were more than 55 at the time of the liver imaging.

Visceral adipose tissue (VAT) volume was measured with MRI at the same time as liver fat and coded as litres. It was defined as the adipose tissue within the abdominal cavity, excluding adipose tissue outside the abdominal skeletal muscles and adipose tissue and lipids within the cavity and posterior of the spine and back muscles. This was included as a potential interaction due to its association with pathological mechanisms in liver steatosis. VAT secretes pro-inflammatory cytokines and adipokines, and releases FAs into circulation, causing dyslipidaemia and systemic insulin resistance (56, 92). Therefore it may be more detrimental to metabolic health than other fat depots such as subcutaneous adipose tissue.

## 2.2 OCDEM participants

In Chapter 6, the study population consisted of a set of participants from seven interventional studies conducted at the clinical research unit of the Oxford Centre of Diabetes, Endocrinology, and Metabolism (OCDEM). These have been integrated into a unified dataset, and will be referred to as the OCDEM sample for the purposes of this study. All data analysed in this thesis was obtained from measurements that were carried out at baseline, before any intervention.

### 2.2.1 Recruitment

Participants who joined these studies were primarily, but not exclusively recruited from the Oxford Biobank, by advertisement via posters in hepatology clinics at the Oxford University Hospital Trust, social media and by word of mouth. The Oxford Biobank is a project in which participants aged 30 to 50 years old were recruited, which was established to study people with specific genetic traits and to collect detailed information about their metabolism and body measurements. It consists of biological material and health-related information from around

8,000 participants, which began recruitment in 1999. In addition, participants had the opportunity to take part in further clinical studies in Oxford (93). Those participants who provided consent to be re-approached for future projects were invited to these studies (93).

The studies from which participants were recruited are outlined in Table 2.4. While most of the studies focused on participants free from any known metabolic disease, both SMASH (Study 3) and Study 7 included the recruitment of participants with diagnosis of steatohepatitis.

Table 2.4 Studies from which OCDEM participants sample were recruited.

<b>Study Number</b>	<b>Study description</b>	<b>N</b>
1	Sugar overfeeding study	24
2	Hepatic fatty acid partitioning in pre-and post-menopausal women: risk factors for cardiovascular disease	43
3	SGLT2 inhibitors and metformin on metabolism and non-alcoholic steatohepatitis (SMASH)	20
4	Fatty Acid Metabolism in individuals Undergoing Sulforaphane supplementation	5
5	Liver Fat case-control	18
6	Fat overfeeding study	5
7	MRI and Stable Isotope Tracer Studies for Detecting the Progression of Non-Alcoholic Steatohepatitis (NASH)	10
<b>Total</b>		<b>125</b>

### 2.2.2 Inclusion criteria

Participants were included in the study if they provided informed consent, were between 18 and 65 years old, had had a stable weight for the previous 3 months, their BMI was between 19 and 35 kg/m<sup>2</sup>, and they were not taking any lipid or glucose lowering medication. Participants were excluded if they did not provide informed consent, if they were smokers, and/or had a history of alcoholism. In addition, women drinking >20 g of alcohol daily and men drinking >20 g of alcohol daily were excluded. Those who were pregnant or presented elevated levels of cholesterol or triglycerides, or glucose after screening were also excluded from the studies 1-2 and 4-6. If participants presented low haemoglobin (<135 mg/dl for men and <120 mg/dl for women), or donated or lost more than 250 ml of blood in the previous months, they were excluded from the study. To ensure participants' safety during the MRI, the

presence of pacemakers, metallic implants, piercings or history of claustrophobia made participants ineligible for this study.

### 2.2.3 Liver fat measurement

Approximately within a week of the first assessment visit, MRI measurements were carried out to obtain data about liver fat content, using spectroscopy. A single voxel (20x 20 x 20 mm<sup>3</sup>) was positioned in the posterior part of the left liver lobe, and both water-suppressed and non-water-suppressed stimulated acquisition mode (STEAM) measurements were performed. The proportion of intrahepatic triacylglycerol (IHTG) in the liver tissue was determined using the OXSA toolbox (94).

### 2.2.4 Clinical research Unit assessments

During study visits, participants would attend in the morning after an overnight fast, and would provide a venous blood sample, which was collected into heparinised tubes. Plasma was immediately separated for analysis by centrifugation at 4°C. Then, concentrations of plasma glucose and lipoproteins were analysed on a semi-automatic analyser. In addition, samples were sent to Nightingale Health for metabolomics analysis, which assessed multiple lipoproteins in plasma. In addition, anthropometric measurements were taken, and body mass index was recorded.

### 2.2.5 VLDL-TG fatty acid composition measurement

Chylomicron- and VLDL-rich fractions were separated from plasma samples by sequential flotation, using density gradient ultracentrifugation (Svedberg flotation rate (Sf) >400 and 20-400, for chylomicrons and VLDL, respectively). Ultracentrifugation was performed in a SW40Ti swinging bucket rotor (Beckman Instruments, Palo Alto, CA) at 40,000 rpm at 15°C. The gradients were run for 32 min to float Sf >400 lipoproteins and for a further 16 h to float Sf 20-400 lipoproteins (95). Following isolation, the Sf 20-400 fraction was further separated

by immunoaffinity chromatography to isolate particles containing apolipoprotein (apo) B-100 from those with apo-B48; leaving the fraction completely devoid of apoB-48; as such, this fraction can be deemed to contain the VLDL-rich lipoprotein specifically (96).

To determine the specific FA composition of VLDL triglyceride (TG) the isolated VLDL-rich fraction was separated by solid-phase extraction. Initially, a chloroform: methanol solution of 2:1 (v/v) was added to samples, as described by Folch et al. (97). The chloroform solubilized lipids, while the methanol aided the separation of proteins, facilitating the lipid extraction (98). Samples were centrifuged and the lipids were collected. Next, the samples were added to a set of solid phase extraction columns, for lipid purification and isolation of TG. TG were eluted with a 2 x 1 ml hexane: chloroform: ethyl acetate solution under vacuum, as described by Burdge et al (99). The samples were further dried in an evaporator. The dried samples were methylated, to convert them into fatty acid methyl esters, which is necessary for gas chromatography (GC) analysis, as they must be converted into derivatives with lower boiling points (100). Therefore, the samples were treated with a solution of 1.5% H<sub>2</sub>SO<sub>4</sub> in methanol, and placed in a water bath at 80°C for one hour. After adding a neutralising solution and centrifuging for 10 minutes, the samples were dried in an evaporator.

Finally, the samples were injected into a GC machine. The GC analysed the fatty acid methyl esters and produced a chromatogram, which displayed peaks at specific retention times. The peaks were used to identify the different fatty acids, by comparing sample retention times to a known standard, and results were expressed as mol%.

### 2.3 Summary of strengths and limitations.

Using both the UK Biobank and OCDEM datasets offers complementary strengths: while the UK Biobank provides large-scale, population-based data with high statistical power, the OCDEM dataset offers more detailed clinical and metabolic measures. In particular, OCDEM

measurements were collected within the same week, ensuring temporal consistency, whereas in the UK Biobank, these measurements were taken years apart. Although dietary data were unavailable for OCDEM participants and confounder details were more limited, the dataset remains a valuable resource for comparing NMR results, and for studying VLDL-TG fatty acid composition, a complex measurement that is challenging to obtain.

## 2.4 Ethical approval

The UK Biobank study has complied with all necessary research ethics protocols, according to the Declaration of Helsinki, and approved by the Northwest Multi-Centre Research Ethics Committee (reference number 21/NW/0157)(61).

The studies carried out at OCDEM were reviewed and approved by the Oxfordshire, Portsmouth, North of Scotland, North West- Lancaster, North East- Tyne & Wear and South Clinical Research Ethics Committees and all research complied with the Declaration of Helsinki. All participants gave written informed consent.

## 2.5 Missing data

Only participants with complete data on exposure and outcome were included in the study, for both the UK Biobank and the OCDEM sample. For covariates, multiple imputation was not used as missing values represented <5%, and they were included in a “Missing” category.

### **3 Dietary fatty acids and liver fat in the UK Biobank**

#### **3.1 Introduction and aim**

While there is no approved pharmacological treatment for reducing excessive liver fat accumulation, and there is a need to prevent it from developing to chronic irreversible stages, the focus is on lifestyle and avoiding hypercaloric diets (6, 22). Dietary guidelines recommend moderate calorie intake, and a reduction in SFAs, a Mediterranean diet, but this is not consistent across societies and institutions, and there is no consensus on advice on PUFAs or MUFAs consumption (101-103).

There is still no agreement about the optimal dietary pattern to prevent excessive liver fat accumulation (18, 104, 105), and the aim of this study is to explore the associations between dietary fatty acids and liver fat in the UK Biobank, in a metabolically healthy population. The need of large prospective cohorts that study dietary fatty acids and liver fat has been pointed out in previous research (106), and presented in Chapter 1, and the UK Biobank is an example of an appropriate cohort to look at this research question, as it has detailed data on dietary assessments and liver fat content.

#### **3.2 Methods**

##### **3.2.1 Study sample, inclusion and exclusion criteria**

The population studied was part of the UK Biobank prospective cohort, described in detail in Chapter 2, Section 2.1. Participants included had data on liver MRI taken in the first imaging visit. Approximately 20% of the UK Biobank expressed interest, and by 2024, 90,000 (18%) have completed an imaging assessment. At the time that these analyses were carried out, data about MRIs was available for 31,334 participants, taken between 2014 and 2021. Dietary questionnaire data was excluded from analyses if diet was not of a typical day: if participants reported being ill, having fasted the previous day, or if the total energy intake was considered

implausible (defined as less than 2510 kJ or more than 14644 kJ in women, or as less than 3,347 kJ or more than 17,572 kJ in men (107)). In total, 1793 participants were excluded from the analyses if their diet was atypical or their energy intake was considered implausible. Only those who had data from at least two plausible 24-hr dietary questionnaires (42% of participants who had data available on liver fat) were included in the study and the exclusions carried out to obtain the final sample are detailed in Figure 3.1.

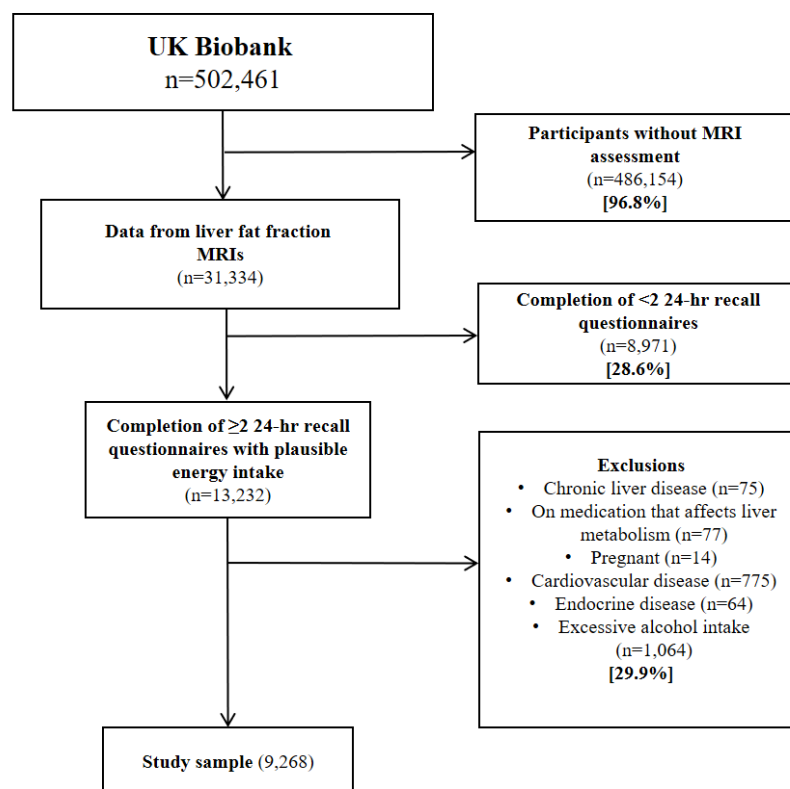


Figure 3.1 Flowchart of UK Biobank showing exclusion of participants included in the analyses.

The exclusion criteria were determined based on health conditions that could alter the associations between dietary fatty acids and liver fat. To reduce reverse causality, participants with prevalent cardiovascular disease, diabetes, and chronic liver disease at recruitment, who may have received dietary advice for their conditions, were excluded from the analyses. If participants were admitted to the hospital during follow up due to cardiovascular or liver

causes, they were also excluded. Participants using cholesterol lowering medication, or drugs that can lead to liver inflammation, were identified and excluded from the final sample.

Alcohol intake can promote liver fat accumulation, and to prevent assessing liver fat in participants in which alcohol related steatotic liver disease may be present, participants were excluded if they were in the highest category of alcohol consumption according to NICE alcohol use disorders guidelines (harmful drinking, defined as >50 alcohol units per week in men, and >30 units per week in women)(88). In addition, alcohol intake was added as a confounder in the model. Alcohol intake was obtained by considering type and frequency of alcohol drinks, data recorded in touchscreen in the first visit, and details on the variables used to code alcohol intake are in Table 2.3 in Chapter 2.

In total, exclusions based on chronic disease and excessive alcohol intake represented 3,964 participants, 29% of the sample with MRI data and 2 WebQs. For more detail on the codes used for this set of exclusions, see Table 2.2 in Chapter 2, Section 2.1.12.

### 3.2.2 Dietary fatty acids and individual fatty acids assessment

Fatty acids intake was measured using the Oxford WebQ, responded in at least two opportunities. Participants completed this questionnaire by selecting the products consumed the previous 24 hours from a list of 206 foods and 32 beverages. From each answer, the intakes of total fat, SFAs, MUFAs, total PUFAs,  $\Omega$  3 PUFAs (n-3 PUFAs) and  $\Omega$  6 PUFAs (n-6 PUFAs) were automatically calculated by the software and presented as variables in grams.

To investigate whether the observed associations differ among individual fatty acids within each group, the UK Biobank Individual Fatty Acids dataset was analysed. This dataset provides detailed information on the intake of 21 individual fatty acids, including 10 SFAs (butyric acid, caproic acid, caprylic acid, capric acid, lauric acid, myristic acid, pentadecylic acid, palmitic acid, margaric acid, and stearic acid), 4 MUFAs (palmitoleic acid, oleic acid, gondoic acid, and



erucic acid), and 7 PUFAs (linoleic acid, alpha-linoleic acid, stearidonic acid, arachidonic acid, eicosapentaenoic acid, docosapentaenoic acid, and docosahexaenoic acid).

The mean of all completed questionnaires was calculated to estimate average daily intake for different nutrients. The exposures of interest were percentage of total energy intake. To calculate this, the mean intake of dietary fatty acids in grams was multiplied by 37.7, which is the kJ provided by one gram of fat, to calculate the amount of total kJ provided. This was then calculated as a percentage of each participant's daily total energy intake.

### 3.2.3 Liver fat content

All participants were assessed with abdominal MRI, to estimate their liver fat PDFF, which represents liver fat content (%). Details on the calculations and methods to measure liver fat can be found in Chapter 2, Section 2.1.8.

### 3.2.4 Covariates

All UK Biobank data fields used to obtain these covariates can be seen in more detail in Chapter 2, Section 2.1.10. A set of potential confounders were chosen *a priori*, based on previous research about variables that are associated with diet and liver fat (11, 108, 109).

Total energy intake, calculated as the average from 24-hr dietary assessments, was included as a confounder in the models, as a continuous variable, due to the evidence of its association with liver fat accumulation (16). This approach enabled the adjustment of total energy intake in the coding of the exposure (as a % of total energy), and by the inclusion of total energy intake as a confounder in the model. This nutrient density model was chosen because it considers nutrient intakes relative to total energy intake, which aligns with public health dietary guidelines, and further adjusting for total energy intake as an additional confounder improves the accuracy of the estimates obtained, as suggested by Tomova et al. (110). Ethnicity was self-reported during the assessment visit at recruitment, through the touch-screen questionnaire, and coded as

‘White’, ‘Mixed’, ‘Asian or Asian British’, ‘Black or Black British’, ‘Other’ or ‘Missing’ (84). The Townsend index was used to estimate deprivation. This index, based on unemployment, overcrowding, and lack of house or car ownership was used, was categorised into quintiles from most deprived to least deprived (85). To adjust for region, data on the assessment centre attended was used, and participants were categorised into 10 regions based on the location of the assessment centre they attended. Physical activity was included as a confounder, as it plays a significant role in determining the variation in energy expenditure among participants (111). Data from the baseline touchscreen questionnaire was used to categorise participants according to total MET hours value per week (low, [ $<10$  MET hours/week], moderate [ $10-50$  MET hrs/wk.], high [ $>50$  MET hrs/wk.]) (86). Smoking status was self-reported and participants were categorised into ‘Never’, ‘Current’, or ‘Previous’ tobacco smokers. Education was categorised according to the highest degree obtained from the ones specified in the touchscreen questionnaire at baseline. Alcohol intake was included in the models as a categorical variable, using data from the touchscreen questionnaire, and categorised based on groups defined by NICE guidelines: none (0 units/week), lower-risk/moderate drinking ( $>1$ ,  $<14$  units/week) and hazardous drinking ( $>14$ ,  $<35$  in women,  $<50$  in men units/week) (88). As described in the exclusions section, those in the harmful drinking category ( $>35$  in women,  $>50$  in men units/week) were excluded from the sample, due to the impact excessive alcohol consumption has on liver fat and overall liver metabolism.

Vegetable and fruit intake was included as a covariate as a marker of healthy diet. It was calculated by combining the total intake in grams of fruits and vegetables automatically calculated in variables from each Oxford WebQ, and by estimating the mean of all questionnaires. BMI was available for all participants, and the value used was the one calculated at the time of liver MRI assessment. Following the international classification for adults aged over 20 years (87), participants were categorised in groups as  $<25$  (normal weight),

25-29.9 (overweight, 30-34.9 (obese), and  $\geq 35$  (obese II). Only 26 participants had a BMI  $< 18$ , defined as 'underweight', and they were collapsed into the  $< 25$  category (normal weight).

To explore a potential interaction between menopause status and the association between dietary fatty acids and liver fat, data on menopause status was included. This was based on findings that suggest that oestrogen slows the progression of chronic liver diseases, which could be because it suppresses inflammation, improves mitochondrial function and alleviates oxidative stress and insulin resistance (112). Women were categorised in the postmenopausal group if they reported it in the assessment visit, had bilateral oophorectomy or hysterectomy, or were more than 55 years old at the time of the liver MRI. Visceral adipose tissue (VAT) was included as a potential interaction, due to its association with pathological mechanisms in liver steatosis, linked with its secretion of pro-inflammatory cytokines and adipokines. In addition, VAT releases FAs into circulation, promoting insulin resistance and dyslipidaemia (56, 92). VAT was assessed with MRI at the same time as the liver fat, as part of the abdominal MRI, and a continuous variable consisting of volume of liver fat (litres) was calculated for each participant by the software developed by AMRA. Participants were divided into groups depending on whether their measurements were lower or higher than the median VAT, which was 2.6 litres for this sample.

### 3.2.5 Statistical analysis

Due to its skewed distribution, liver fat content (%) was logarithmically transformed. Multivariable linear regression models were built to estimate the difference in log liver fat content, per 5% increase in dietary exposure. Adjusted geometric mean percentage difference of liver fat% were calculated from the resulting beta coefficients. Beta coefficients obtained were calculated as percentage, to provide the geometric mean percentage difference in liver fat, per 5% increase of dietary exposure (113).

$$\textbf{\textit{Geometric mean \% difference}} = 100 (e^{\beta} - 1)$$

For individual FAs, multivariable linear regression models were built for each of the 21 exposures, to estimate the geometric mean % difference in liver fat, per 1 SD increase in FA percentage of energy intake. To address the issue of multiple comparisons across many statistical tests, adjustment of the level of significance for multiple testing was made by implementing Bonferroni correction. This was carried out as 0.05/21 exposures, and thus the threshold for statistical significance was 0.0024 (114). Bonferroni correction was not applied to the broader FAs categories (SFAs, PUFAs, and MUFAs), as they involved fewer comparisons and were already aggregated. Only fatty acids contributing >1% to total energy intake were discussed to focus on those with meaningful dietary relevance. To illustrate differences between high and low nutrient intake groups, population characteristics were presented by lowest and highest quintile of intake of fatty acids.

In minimally adjusted models (M1), adjustments were made for sex and age at the time of MRI. Models were further adjusted for ethnicity (M2); region, education, and deprivation as markers of socioeconomic status (M3); smoking status (M4); physical activity (M5); alcohol intake (M6); total energy intake (M7); fruit and vegetable intake (M8); and BMI (M9). The covariates chosen were added sequentially, and their contribution to the models was assessed by using likelihood ratio tests (LRTs). To decide the coding of the nutrient variables, tests for departures from linearity were carried out using LRTs comparing a model with dietary fatty acid as a continuous variable with a model with dietary fatty acid coded as categories. Tests confirmed that linear variables would be a better fit and therefore exposures were coded as continuous % of total energy intake. Linear association assumptions were tested: model residuals were assessed using Q-Q plots and histograms, to ensure their normal distribution and mean of zero,

and homoscedasticity was confirmed by using residual versus fitted plots. No violations of the linear association assumptions were observed.

### 3.2.6 Subgroup analyses by sex, visceral adiposity and menopausal status

Tests for heterogeneity across sex, menopause status, and VAT groups were carried out for each of the dietary exposures. This was calculated by using likelihood ratio tests, comparing a model with the subgroup category as a confounder with another model that includes the subgroup as an interaction term in the association between exposure and liver fat content.

VAT subgroup analyses were initially conducted in models excluding BMI. Additionally, results from VAT subgroup analyses in models including BMI were presented.

### 3.2.7 Sensitivity analyses

A subsample consisting of participants who responded to at least 4 WebQs was analysed to explore the associations between dietary fatty acids and liver fat content. This allowed for the study of these associations in participants whose dietary assessments provided a better estimate of usual intake and offered more opportunities to capture food groups that are not consumed daily (115, 116).

An additional sensitivity analysis was included, analysing only those participants with lower alcohol intake according to NICE guidelines regarding alcohol consumption (men and women who reported <14 units/week), which represented 59% of the sample (N= 5,477). This facilitated the investigation of the associations between moderate alcohol intake and dietary FA intake (88).

### 3.2.8 Software for statistical analyses

All analyses were done using STATA S.E 16 (StataCorp. 2019. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC). Plots were made using the R package Jasper version 2 (117).

### 3.3 Results

#### 3.3.1 Participants characteristics

Participant characteristics (N=9,268) are presented across lowest (Q1) and highest (Q5) quintiles of dietary intake of fatty acids in Table 3.1. Median liver fat was 2.6% (IQR: 1.9%-4.0%), and 1,416 participants (15.2%) had  $\geq 5.6\%$  of liver fat, the cut-off value to define liver steatosis by MRI. The mean age of the participants was 62.8 (SD: 7.3) years at the time of MRI measurement, and a larger proportion were women (59.3%). Individuals who reported the highest total fat, SFAs, MUFAs and PUFAs intake had a lower mean alcohol intake and higher energy intake than those in the lowest quintiles of intake. Participants were more likely to be current smokers in Q5 of MUFAs and SFAs consumption than in Q1. There was a higher proportion of participants in the most deprived category in the highest intake of all four dietary fat exposures. Fruits and vegetable median intake was lower in participants who reported the highest consumption of total fat, SFAs, and MUFAs, while it was higher in those who reported the highest intake of PUFAs.

The individual fatty acids that represented  $>1\%$  of energy intake in this study were oleic acid, linoleic acid, myristic acid, palmitic acid and stearic acid. This is consistent with the overall distribution of these fatty acids intake in the whole UK Biobank: linoleic and oleic acid are the major contributors to PUFAs and MUFAs respectively, while palmitic and stearic are the main SFAs (72). In this sample, the median time between last WebQ and liver MRI was 7 years and 3 months (IQR: years 5 months-7 years 10 months). The median difference in BMI between baseline and time of liver MRI was +4% (IQR: 1.7- 7.1).

Table 3.1 Participant characteristics by lowest (Q1) and highest (Q5) quintiles of intake of fatty acids. N=9,268.

	Total fat		SFAs		PUFAs		MUFAs		Overall
	Q1	Q5	Q1	Q5	Q1	Q5	Q1	Q5	N=9,268
<b>Liver fat content (%) *</b>	2.6 (1.8,4.0)	2.7 (1.9,4.3)	2.4 (1.8,3.7)	2.7 (1.9,4.3)	2.6 (1.9,4.2)	2.4 (1.8,3.7)	2.6 (1.8,3.8)	2.7 (1.9,4.2)	2.6 (1.9,4.0)
<b>Age at MRI (years)</b>	63.1 (7)	62.4 (7)	62.9 (7)	63.2 (7)	63.4 (7)	62.1 (7)	63.1 (7)	62.3 (7.3)	62.8 (7.3)
<b>Men^</b>	755 (40.7)	737 (39.8)	695 (37.5)	768 (41.4)	786 (42.4)	682 (36.8)	737 (39.8)	747 (40.3)	3,738 (40.3)
<b>BMI (kg/m<sup>2</sup>)</b>	25.8 (3.8)	25.9 (4.2)	25.6 (3.9)	25.7 (4.0)	25.8 (3.9)	25.5 (4.1)	25.6 (3.8)	26.0 (4.1)	25.7 (3.9)
<b>White ethnicity ^</b>	1,787 (96.4)	1,794 (96.8)	1,766 (95.3)	1,815 (97.9)	1,814 (97.8)	1,765 (95.3)	1,802 (97.2)	1,789 (96.5)	8,984 (96.9)
<b>Most deprived ^</b>	361 (19.5)	399 (21.6)	372 (20.1)	389 (21.0)	350 (18.9)	434 (23.5)	332 (17.9)	406 (22.0)	1,849 (20.0)
<b>University degree ^</b>	1,471 (79.3)	1,514 (81.7)	1,472 (79.4)	1,505 (81.2)	1,480 (79.8)	1,493 (80.6)	1,453 (78.4)	1,491 (80.5)	7,455 (80.4)
<b>Current smoker ^</b>	79 (4.3)	112 (6.0)	79 (4.3)	116 (6.3)	98 (5.3)	75 (4.0)	78 (4.2)	118 (6.4)	436 (4.7)
<b>Low physical activity^</b>	341 (18.4)	418 (22.6)	323 (17.4)	412 (22.2)	400 (21.6)	380 (20.5)	328 (17.7)	421 (22.7)	1,837 (19.8)
<b>Hazardous alcohol intake ^</b>	852 (46.0)	658 (35.5)	832 (44.9)	671 (36.2)	840 (45.3)	627 (33.8)	805 (43.4)	693 (37.4)	3,780 (40.8)
<b>Alcohol intake (units/week)*</b>	16 (8,25)	12(6,21)	15(8,24)	12 (6,22)	15 (8,25)	12(6,20)	14 (8,24)	13(6,22)	14(8,23)
<b>Total energy intake (kJ)</b>	7573 (1696)	8938(1973)	7619 (1735)	8974 (1962)	8001 (1847)	8417 (1945)	7620 (1689)	8860 (1986)	8400 (1894)
<b>Fruits and vegetables (grams/day) *</b>	413 (283, 594)	307 (191,448)	428 (293,616)	292 (187,418)	356 (232,505)	367 (237,533)	403 (276,575)	314 (199,460)	352 (232,502)
<b>Total fat (% of energy intake)</b>	23.9 (2.7)	39.6 (3.2)	25.5 (4.4)	37.6 (4.2)	27.1 (5.0)	36.1 (5.2)	24.4 (3.2)	38.9 (3.8)	31.7 (5.6)
<b>SFAs (% of energy intake)</b>	8.4 (1.8)	14.5 (2.5)	7.7 (1.2)	15.6 (1.6)	11.3 (3.1)	11.2 (2.7)	9.1 (2.3)	13.3 (2.7)	11.5 (2.8)
<b>MUFAs (% of energy intake)</b>	8.8 (1.3)	15.2 (1.9)	10.0 (2.3)	13.6 (2.3)	9.6 (1.8)	14.2 (2.5)	8.6 (1.1)	15.6 (1.6)	12.0 (2.5)
<b>PUFAs (% of energy intake)</b>	4.5 (1.1)	6.8 (1.6)	5.6 (1.6)	5.5 (1.3)	3.8 (0.5)	7.8 (1.0)	4.4 (1.1)	7.0 (1.4)	5.6 (1.5)

Values are mean (SD), unless otherwise specified. ^N (proportion %) \*Values are median (IQR) SFAs: saturated fatty acids PUFAs: polyunsaturated fatty acids MUFAs: monounsaturated fatty acids kJ: kilojoules, BMI: body mass index.

### 3.3.2 Associations between dietary fatty acids and liver fat content

The sequential adjustment for confounders presents the impact of each covariate added to the models in Figure 3.2. In minimally adjusted models, a 5% in both SFAs and MUFAs was positively associated with liver fat: +7.7% difference in liver fat geometric mean (95% Confidence Interval: 5.2, 10.3) for SFAs, and +5.1% (2.3, 7.9), for MUFAs. In contrast, PUFAs presented an inverse association with liver fat: -7.4% (-11.3, -3.3). It can be observed that the addition of most confounders had a moderate impact on the estimates. In particular, the inclusion of body mass index notably decreased the strength of the associations between all fatty acids and liver fat. Body mass index had a substantial impact in the statistical significance of the associations between MUFAs and liver fat, and its addition changed the estimates from +3.5% (0.7, 6.4) to +1.9% (-0.6, 4.4). Adjustment for most of the other covariates did not have substantial impact, except for fruits and vegetables, which modestly attenuated the associations for total fat, SFAs and MUFAs.



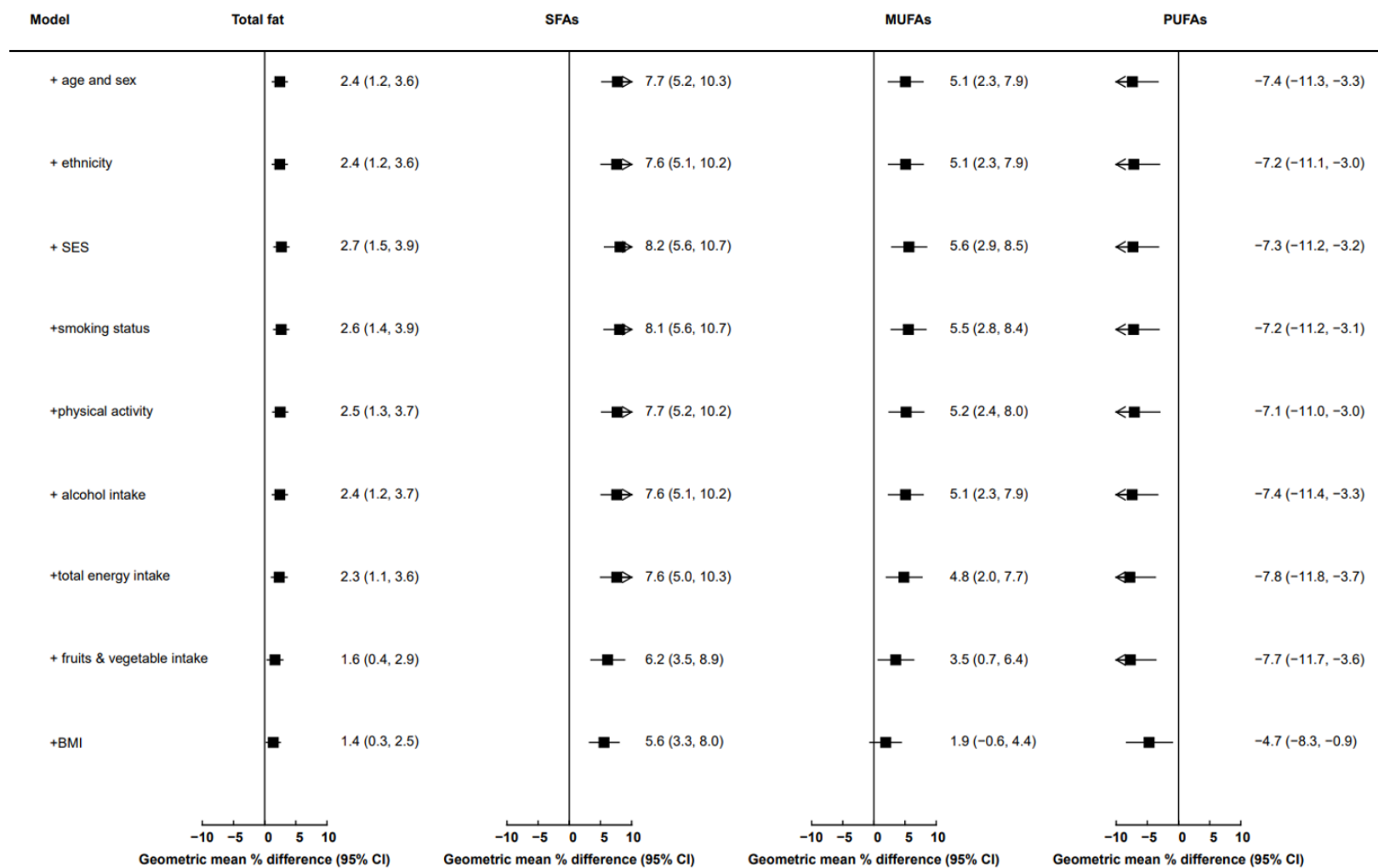


Figure 3.2 Liver fat geometric mean % difference per 5% increase in dietary fatty acids, in minimally adjusted models and by addition of confounders.

Fully adjusted associations between a 5% increase in dietary fatty acids and liver fat % are presented in Figure 3.3. In this sample, a 5% increase in total fat intake was associated with a liver fat geometric mean percentage difference of +1.4% (0.3, 2.5; p value=0.02). A 5% increase in SFAs was associated with a +5.6% (3.3, 8.0; p <0.001) liver fat geometric mean percentage difference. Conversely, an increase of 5% intake of PUFAs was inversely associated with liver fat, with -4.7% (-8.3, -0.9; p=0.01) liver fat geometric mean. MUFAs presented a non-significant positive association with liver fat: +1.9% (-0.6, 4.4; p=0.15).

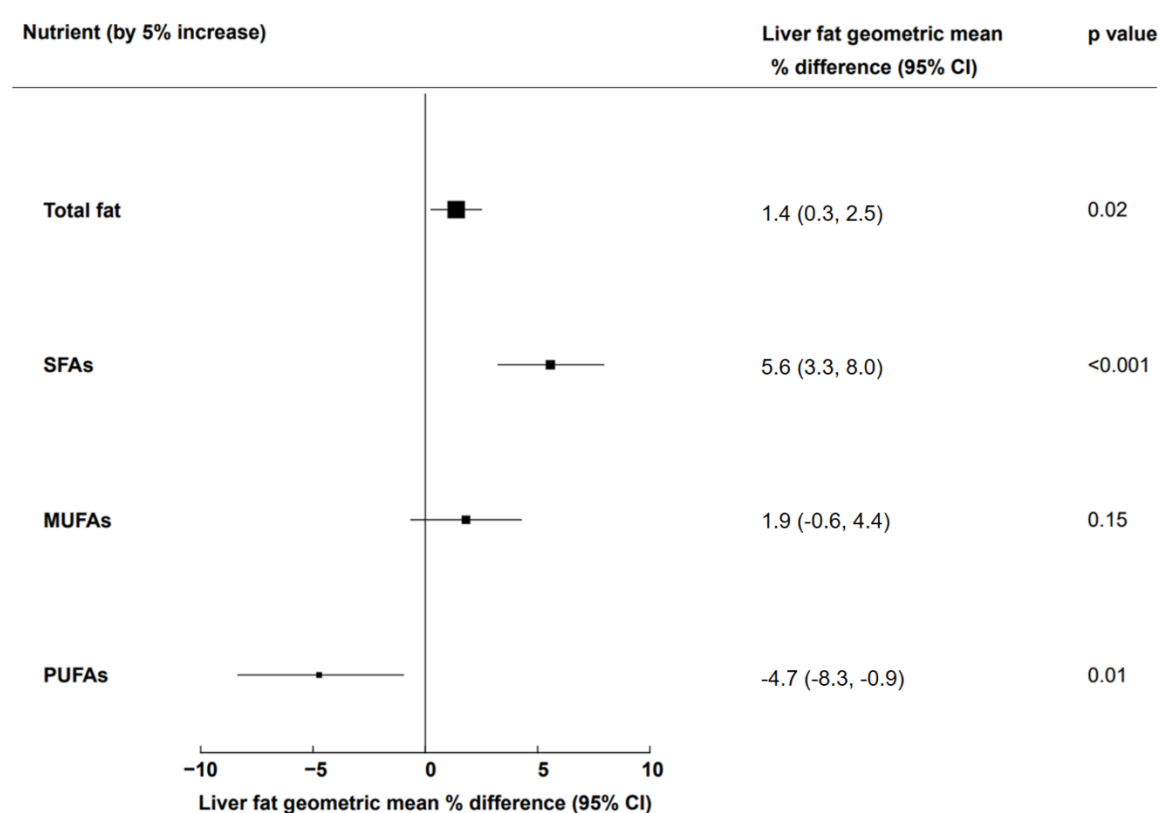


Figure 3.3 Liver fat geometric mean % difference per 5% increase in dietary fatty acids, in fully adjusted models. Models were adjusted for age, ethnicity, Townsend deprivation index, education, region, smoking status, physical activity, alcohol intake, total energy intake, fruits and vegetables intake and body mass index. N=9,268.

### 3.3.3 Associations between individual dietary fatty acids and liver fat content

As many of the individual FAs represented a very low proportion of total energy intake, the reporting of findings will only focus on individual FAs which represented at least 1% of total energy intake. After fully adjusting for confounders, the largest association was observed

between stearic acid and liver fat: per 1SD increase in stearic acid, there was a +3.8 % (2.5, 5.51) geometric mean % difference in liver fat (Figure 3.4). Palmitic acid was also associated with higher liver fat, with geometric mean difference of +2.9% (1.6, 4.2) per 1SD increase in intake. In the case of PUFAs, total omega-3 fatty acids and omega-6 fatty acids both had inverse associations with liver fat: per 1 SD increase there was a geometric mean difference of -1.5% (-2.6, -0.4) and -1.6 (-2.7, -0.5), respectively. However, the significance threshold established to correct for multiple testing was not met ( $p=0.006$  and  $0.004$ , respectively).

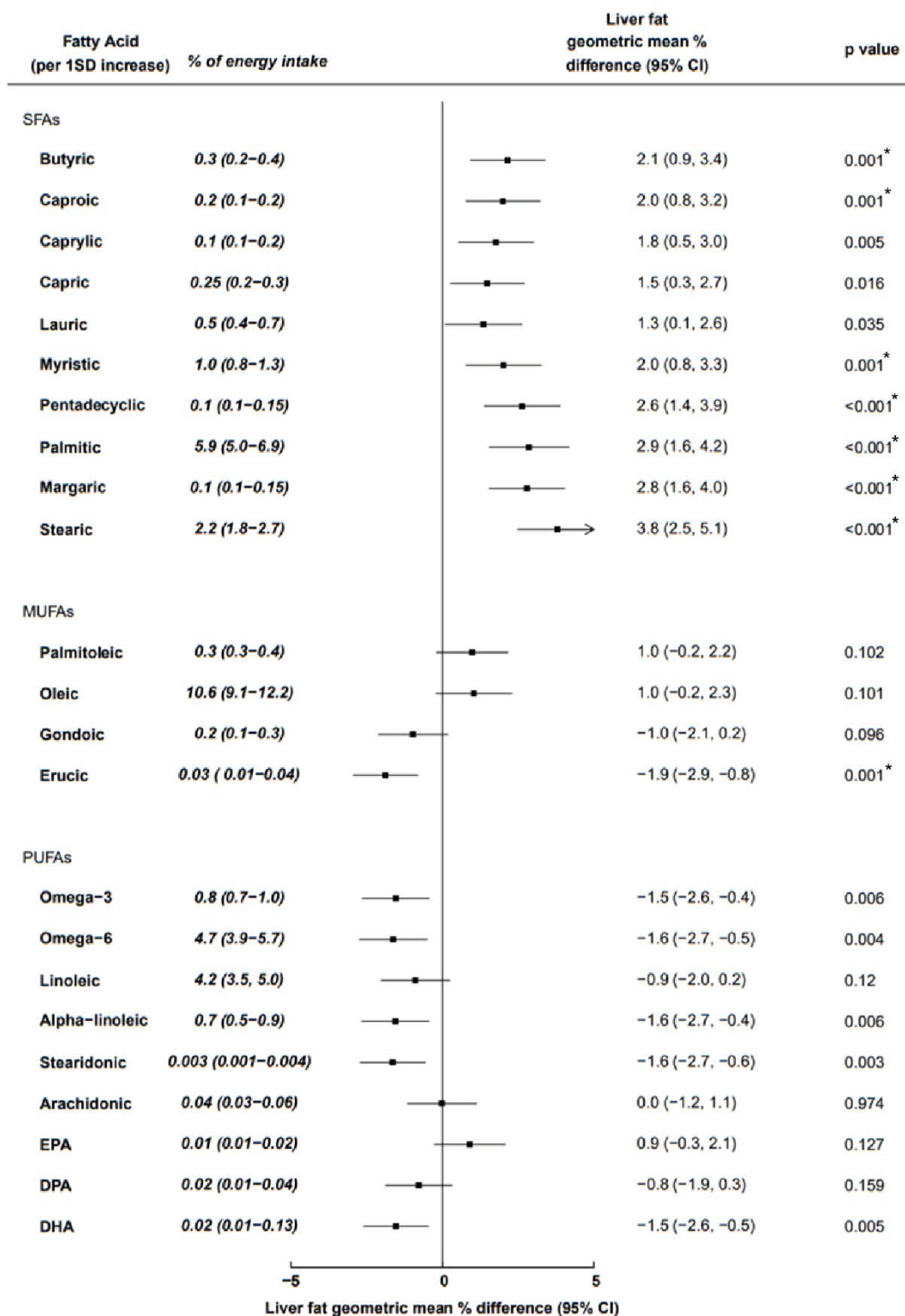


Figure 3.4 Geometric mean % difference in liver fat, per 1SD increase in individual fatty acids intake. Models were adjusted for age, ethnicity, Townsend deprivation index, education, region, smoking status, physical activity, alcohol intake, total energy intake, fruits and vegetable intake and BMI. \*statistically significant: <0.0024.

### 3.3.4 Subgroup analyses

An interaction between visceral adiposity and SFAs was observed, but was not present for PUFAs or MUFAs (Figure 3.5).

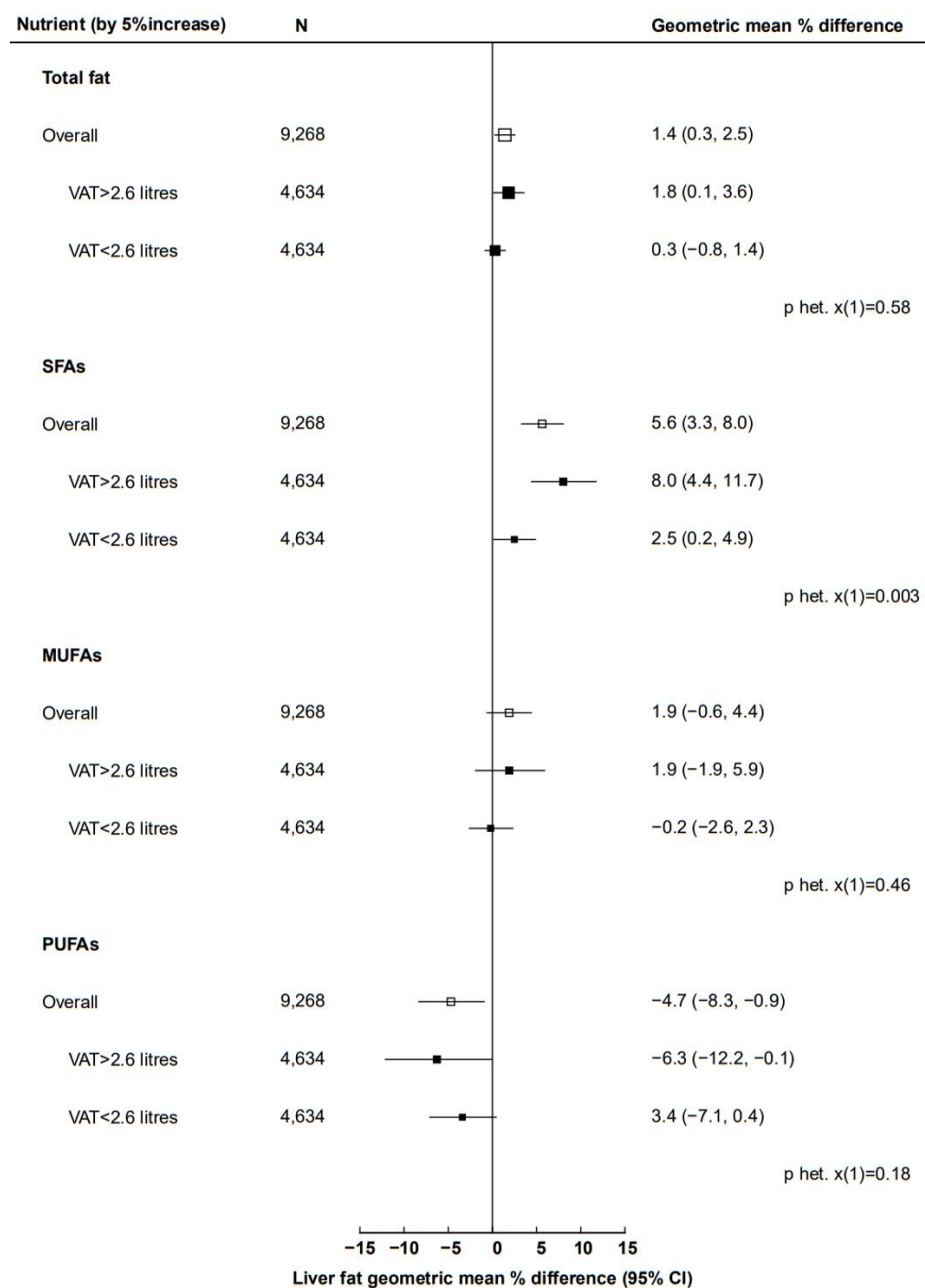


Figure 3.5 Subgroup analyses by adiposity. Linear regression models were used to calculate the geometric mean % difference in liver fat, by 5% increase in dietary fatty acids intake. Models were adjusted for age, sex, ethnicity, Townsend deprivation index, education, region, smoking status, physical activity, alcohol intake, total energy intake, fruits and vegetables intake, and BMI.

For SFAs, those in the group with VAT higher than the median (2.6 litres) had a liver fat geometric mean difference of +8.0% (4.4, 11.7) per 5% increase in SFAs intake. On the opposite, per 5% in SFAs intake, those on the lower VAT group presented a liver fat geometric mean difference of 2.5% (0.2, 4.9, p for heterogeneity=0.003).

For the associations between SFAs and liver fat, men presented a stronger association between SFAs and liver fat than women (+7.5% [3.9, 11.3] liver fat geometric mean compared to +4.4% [1.6, 7.2], per 5% increase in SFAs intake), which can be observed in Figure 3.6.

No significant interactions were found between menopausal status and SFAs (p heterogeneity=0.60), MUFAs (p heterogeneity =0.86), or PUFAs (p heterogeneity=0.39).

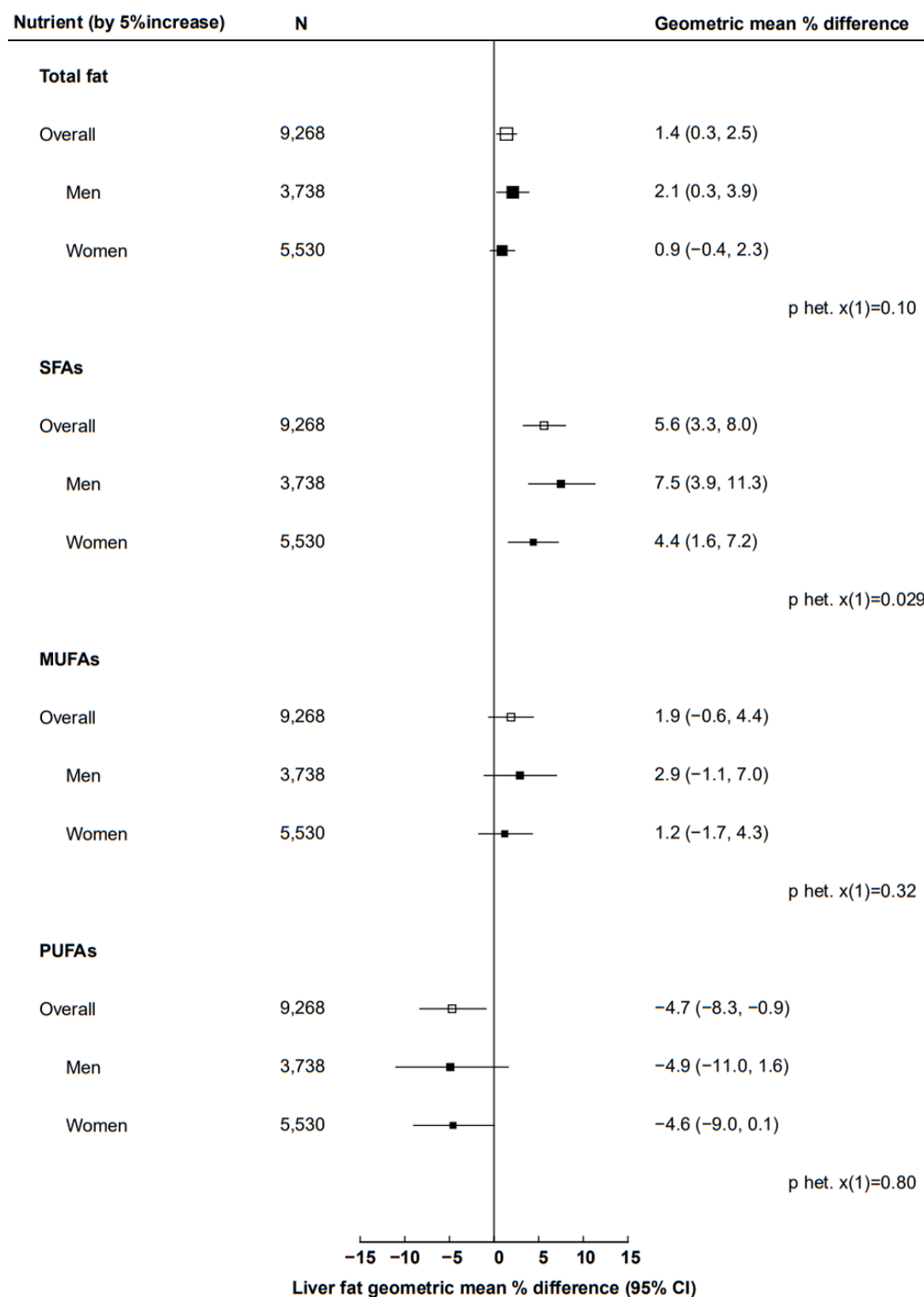


Figure 3.6 Subgroup analyses by sex. Linear regression models were used to calculate the geometric mean % difference in liver fat, by 5% increase in dietary fatty acids intake. Models were adjusted for age, ethnicity, Townsend deprivation index, education, region, smoking status, physical activity, alcohol intake, total energy intake, fruits and vegetables intake and BMI.

### 3.3.5 Sensitivity analyses

Participants who responded to four to five WebQs represented 30.5% of the sample (N=2,828). The direction of the associations and the level of significance was similar to the results from the main sample, with the exception of total fat, which became non-significant. Associations between PUFAs and liver fat became stronger, as the geometric mean % difference in liver fat per 5% increase in PUFAs was -10.7% (-17.6, -3.3) (Figure 3.7). This was substantially stronger than the results from the main sample: -4.7% (-8.3, -0.9) liver fat geometric mean % difference per 5% increase in PUFAs.

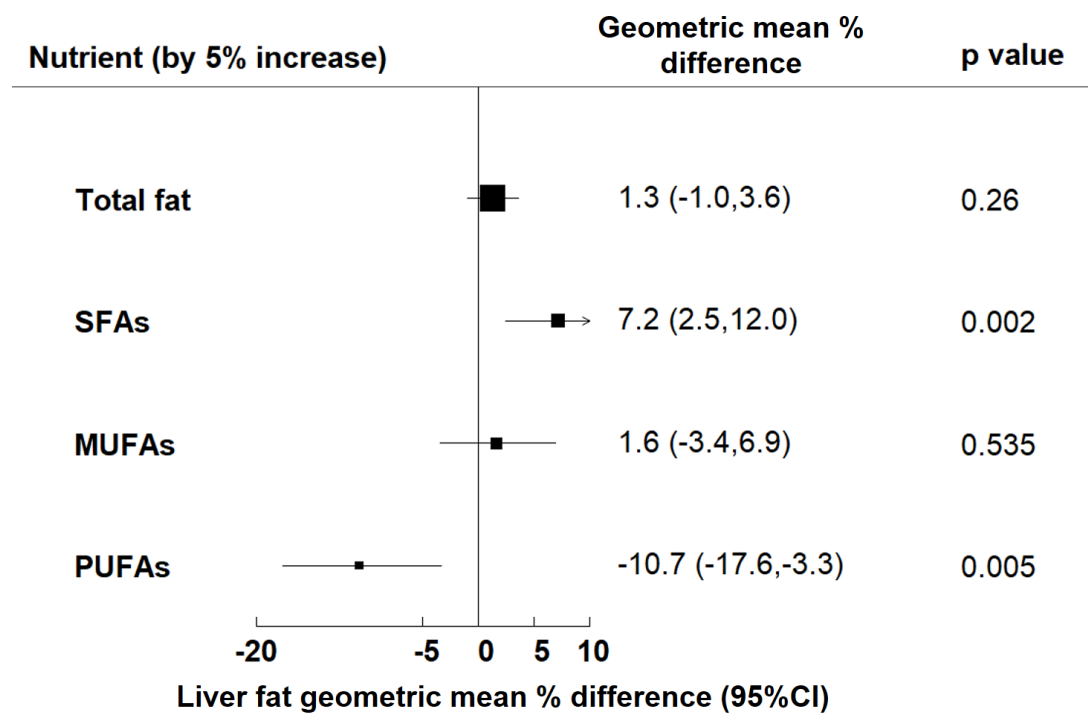


Figure 3.7 Sensitivity analysis. Analysis was restricted to participants who answered 4 to 5 WebQ. Models were adjusted for age, ethnicity, Townsend deprivation index, education, region, smoking status, physical activity, alcohol intake, total energy intake, fruit and vegetable intake, and body mass index (N=2,828).

For individual fatty acids, restricting the sample to those who answered  $\geq 4$  WebQs did not substantially change the associations presented in the main sample (Table 3.2).



Table 3.2 Results obtained from sensitivity analyses restricted to participants who answered 4 to 5 WebQ (N=2,828) compared to results obtained from main sample.

<b>Individual fatty acids (per 1SD increase)</b>	<b>Geometric mean %difference in liver fat (95%CI) (N=2,828)</b>	<b>Geometric mean % difference in liver fat (95%CI) (Main sample: N=9,268)</b>
<b>SFAs</b>		
Butyric	3.0 (0.6, 5.5)	2.1 (0.9, 3.4)
Caproic	2.8 (0.3, 5.2)	2.0 (0.8, 3.2)
Caprylic	2.2 (-0.4, 4.8)	1.8 (0.5, 3.0)
Capric	1.9 (-0.5, 4.4)	1.5 (0.3, 2.7)
Lauric	1.6 (-1.2, 4.5)	1.3 (0.1, 2.6)
Myristic	2.7 (0.3, 5.2)	2.0 (0.8, 3.3)
Pentadecyclic	3.1 (0.7, 5.5)	2.6 (1.4, 3.9)
Palmitic	3.6 (1.0, 6.3)	2.9 (1.6, 4.0)
Margaric	3.4 (0.9, 6.0)	2.8 (1.6, 4.0)
Stearic	5.2 (2.5, 7.9)	3.8 (2.5, 5.1)
<b>MUFAs</b>		
Palmitoleic	1.1 (-1.3, 3.6)	1.0 (-0.2, 2.2)
Oleic	0.8 (-1.8, 3.4)	1.0 (-0.2, 2.3)
Erucic	-3.1 (-5.7, -0.6)	-1.9 (-2.9, -0.8)
Gondoic	-0.2 (-2.9, -2.5)	-1.0 (-2.1, 0.2)
<b>PUFAs</b>		
Total omega 3	-3.8 (-6.2, -1.2)	-1.5 -2.6, -0.4)
Total omega 6	-3.4 (-5.6, -1.1)	-1.6 (-2.7, -0.5)
Linoleic	-2.4 (-4.7, -0.1)	-0.9 (-2.0, 0.2)
Alpha-linoleic	-3.6 (-5.8, -1.4)	-1.6 (-2.7, -0.4)
Stearidonic	-3.0 (-5.6, -0.5)	-1.6 (-2.7, -0.6)
Arachidonic	-0.1 (-2.6, 2.5)	0.0 (-1.2, 1.1)
EPA	1.0 (-1.5, 3.6)	0.9 (-0.3, 2.1)
DPA	-1.6 (-4.2, 1.0)	-0.8 (-1.9, 0.3)
DHA	-3.1 (-5.6, -0.5)	-1.5 (-2.6, -0.5)

SFAs: saturated fatty acids PUFAs: polyunsaturated fatty acids MUFAs: monounsaturated fatty acids, EPA: eicosapentanoic acid DPA: docosapentanoic acid, DHA: docosahexaenoic acid

Limiting the sample to participants whose alcohol consumption was lower than 14 units per week resulted in a sample of N=5,477 (Table 3.3). When fully adjusting for confounders, the positive associations between liver fat and dietary SFAs and total fat remained similar to the overall sample. However, PUFAs associations were not significant in this sensitivity analysis.

Table 3.3 Sensitivity analysis. Results obtained from fully adjusted models examining participants with lower alcohol consumption (N=5,477)<sup>1,2</sup>, compared with results obtained from main sample

	<b>Geometric mean % difference in liver fat (95%CI) (N=5,477)</b>	<b>Geometric mean % difference in liver fat (95%CI) (Main sample:N=9,268)</b>
<b>Fatty acid (per 5% increase)</b>		
Total fat	1.8 (0.2,3.5)	1.4 (0.3, 2.5)
PUFAs	-4.1 (-9.2,1.2)	-4.7 (-8.3, -0.9)
SFAs	6.0 (2.7,9.3)	5.6 (3.3, 8.0)
MUFAs	3.5 (-0.1,7.1)	1.9 (-0.6, 4.4)
<b>Individual fatty acids ( per 1SD increase)</b>		
<b>SFAs</b>		
Butyric	1.7 (0.1,3.4)	2.1 (0.9, 3.4)
Caproic	1.6 (-0.1,3.2)	2.0 (0.8, 3.2)
Caprylic	1.5 (-0.2,3.2)	1.8 (0.5, 3.0)
Capric	0.8 (-0.8,2.5)	1.5 (0.3, 2.7)
Lauric	1.0 (-0.7,2.8)	1.3 (0.1, 2.6)
Myristic	1.6 (-0.1,3.3)	2.0 (0.8, 3.3)
Pentadecyclic	2.2 (0.5,3.9)	2.6 (1.4, 3.9)
Palmitic	3.3 (1.5,5.2)	2.9 (1.6, 4.0)
Margaric	2.7 (1.0,4.4)	2.8 (1.6, 4.0)
Stearic	4.4 (2.5,6.3)	3.8 (2.5, 5.1)
<b>MUFAs</b>		
Palmitoleic	1.6 (-0.1,3.3)	1.0 (-0.2, 2.2)
Oleic	1.9 (0.1,3.7)	1.0 (-0.2, 2.3)
Gondoic	-0.7 (-2.2,0.9)	-1.9 (-2.9, -0.8)
Erucic	-2.2 (-3.7,-0.7)	-1.0 (-2.1, 0.2)
<b>PUFAs</b>		
Total omega 3 PUFAs	-1.9 (-3.4,-0.4)	-1.5 -2.6, -0.4)
Total omega 6 PUFAs	-1.5 (-3.1,0.1)	-1.6 (-2.7, -0.5)
Linoleic	-0.5 (-2.0,1.1)	-0.9 (-2.0, 0.2)
Alpha-linoleic	-2.3 (-3.9,-0.8)	-1.6 (-2.7, -0.4)
Stearidonic	-2.0 (-3.5,-0.5)	-1.6 (-2.7, -0.6)
Arachidonic	0.5 (-1.1,2.1)	0.0 (-1.2, 1.1)
EPA	1.2 (-0.5,2.9)	0.9 (-0.3, 2.1)
DPA	-0.9 (-2.4,0.6)	-0.8 (-1.9, 0.3)
DHA	-1.9 (-3.3,-0.4)	-1.5 (-2.6, -0.5)

SFAs: saturated fatty acids PUFAs: polyunsaturated fatty acids MUFAs: monounsaturated fatty acids, EPA: eicosapentanoic acid DPA: docosapentanoic acid, DHA: docosaheaxenoic acid.

### 3.4 Discussion

#### 3.4.1 Main findings

This observational study of 9,268 participants from the UK Biobank was the first to investigate the associations between individual fatty acids and liver fat content in the UK Biobank, and found positive associations between dietary SFAs intake, and inverse associations between dietary PUFAs intake and liver fat content, whilst no relationship was observed between dietary MUFAs intake and liver fat content, after fully adjusting for confounders. Individual fatty acids observed followed the group they belonged to, and those that were significant represented the major sources of energy in this sample. In subgroup analyses, the associations between SFAs and liver fat content were stronger for men, and for people in the high VAT category. No significant interactions by menopausal status were observed in any analyses performed.

A 5% increase in SFAs was positively associated with liver fat content, which is consistent with many observational and interventional studies investigating the role of SFAs in liver steatosis across different settings (31, 32, 44, 45, 48). This study reinforces the positive associations between SFAs and liver fat published in studies on the UK Biobank in recent years (31, 32) and adds to the published literature by studying a larger sample, with the use of at least 2 dietary assessments, and careful consideration of alcohol intake, VAT and BMI. A potential explanation is that SFAs would promote accumulation of liver fat, because SFAs are associated with higher production of ceramides (23, 47). Ceramides promote insulin resistance in the liver, inflammation and mitochondrial dysfunction, which would explain the higher liver fat content (23). The higher association between SFAs and liver fat content in participants with high VAT may be explained by the fact that a higher VAT depot contributes to higher flux of FAs, as VAT secretes FAs to the bloodstream, increasing the availability of SFAs, and the excess SFAs would promote accumulation of liver fat.

Previous studies have found that a high PUFAs dietary pattern is associated with a lower risk of NAFLD (33, 105), consistent with my results for PUFAs and their inverse association with liver fat content. This could be explained by the fact that whole-body oxidation of dietary PUFAs, may be higher than other fatty acids: an interventional study found that linoleic acid (the major PUFAs contributor to energy intake in this study) was partitioned into oxidation pathways at much greater extent than palmitic acid (the major contributor to energy intake in this sample), therefore reducing its availability to be metabolised towards fat deposition in the liver (118). It could be argued that PUFAs may increase oxidation by the activation of peroxisome proliferator activated receptors in the liver (119).

However, when restricting the sample to participants with lower alcohol intake in a sensitivity analysis, the associations between PUFAs and liver fat content became non-significant. As this group consisted of less than 60% of the total main sample, a potential explanation of the lack of significance may be a substantial reduction in power, lowering the ability to detect associations. No individual PUFAs presented significant associations in main analyses or sensitivity analyses.

### 3.4.2 Strengths and limitations

The analyses presented had many strengths, including outcome ascertainment, as the use of MRIs provided a non-invasive and accurate measurement of liver fat content (80). The use of MRI for quantifying liver fat has recently been compared to biopsies, and a meta-analysis concluded that the area under the curve consistently exceeds 90% for detecting all grades of steatosis, while it slightly drops at very advanced levels of steatosis (120). In addition, there is no need to carry out biopsies which are invasive and present more risks and discomfort for participants. This sample consisted of participants without a known diagnosis of diabetes. The detailed data available from the recruitment assessment, along with linkage to hospital admissions data, enabled the identification and exclusion of relevant cases. This is a strength

because it ensured the inclusion of participants without metabolic impairments caused by diabetes or insulin resistance, conditions that substantially influence liver metabolism and may confound the associations between fatty acids and liver fat content.

Another strength of this study is that it investigated potential differences between men and women, and levels of visceral adiposity, independently of BMI. This provides results not only considering participants total BMI, but the VAT component, which has a metabolic role and is linked to inflammation and overall metabolic impairment.

The data obtained from the Oxford WebQ provided detailed information about dietary fatty acids and individual fatty acids. In contrast with two similar studies carried out by Chen et al. and Friden et al. in the UK Biobank, with similar research questions, a minimum of two WebQs instead of a minimum of one was required for inclusion in this study (31, 32). This provided more precise measurements of diet, as one dietary assessment is considered a poor estimate of true long-term diet (31, 32, 121). The use of sensitivity analyses restricted to  $\geq 4$  dietary assessments was particularly relevant in the case of PUFAs, as their main sources are certain types of food products which in the UK may be consumed approximately around one portion per week (107, 122). This repeated assessment attempted to account for variation in nutrient intake across time by administering it throughout different times of the week and year, in a way that prevents response fatigue (70). However, it is important to note that self-reported dietary intake may introduce measurement error, and therefore information bias. For example, it is likely that self-report biases the associations towards the null in this case, as food products considered unhealthy, rich in saturated fats, may have been underreported (123). Therefore, the strength of the associations observed here may be lower than in real conditions.

Due to the age at recruitment of the UK Biobank cohort, most of the women in this sample were in the post menopause group (women included were 62 years old on average). In addition,

the measurement of menopausal status was imperfect as it relied on self-report and age at the time of MRI. The potential interaction between menopause status and the association between dietary fatty acids and liver fat could be studied more precisely in cohorts that recruit women at younger age, and with more precise measurements of menopausal status.

While this cohort attempted to recruit participants who are representative of the UK population, only ~5% of those who were invited actually participated. Plausibly, the small amount of those who participated compared to those invited may have introduced a healthy (and wealthy) volunteer bias, which could especially affect the willingness to respond multiple dietary questionnaires (70). Still, while acknowledging the tendency to be more health-conscious, analyses can be carried out and comparisons between low and high intakes can inform different associations between dietary fats and liver fat. In addition, it has been shown that results from studies using the UK Biobank could be generalisable, as they have been consistent with studies that were adequately representative of the population of interest (124, 125). Finally, the UK Biobank population is >90% of white ethnicity, and the results shown here may not be generalisable to more diverse populations.

### 3.5 Conclusions

This was the first study to assess the associations between individual fatty acids and liver fat content, and the first one to assess dietary fatty acids and liver fat using at least 2 dietary questionnaires in the UK Biobank. It was observed that dietary SFAs intake exhibits a positive association with liver fat content, with stronger associations for men, and for participants in the highest category of VAT. The associations of PUFAs and MUFAs are less clear and should be considered carefully, with new studies that allow ideal control for confounders, especially alcohol intake. More research about fatty acids and liver fat could be enhanced by looking at more diverse populations, larger groups of premenopausal women, using data of individual fatty acids, and more precise measurements of alcohol intake.

The results from this chapter highlight the potentially harmful role of SFAs in regards to liver fat accumulation, and confirms the current recommendations in the UK of low dietary SFAs intake (<10% of total energy intake).

## **4 Dietary fatty acids and MASLD in the UK Biobank**

### **4.1 Introduction and aim**

Until 2023, liver steatosis that was not attributable to excessive alcohol consumption or other specific diseases was categorised as NAFLD (non-alcoholic fatty liver disease). MASLD (metabolic dysfunction-associated steatotic liver disease), defined as SLD in the presence of one cardiometabolic factor, has been adopted as a replacement of the term NAFLD (7).

The guidelines from EASL have been updated and developed recommendations based on these new definitions, while some others, like the UK NICE guidelines, yet have to incorporate the new definitions, as they still refer to NAFLD (6, 101).

Studies have observed that data from NAFLD can be extrapolated to MASLD, and studies have observed that MASLD has very similar mortality rates and clinical profiles (126-129). However, they found that some patients who would have been diagnosed as NAFLD were not diagnosed with MASLD, and they were younger, leaner and with less comorbidities. This suggests that NAFLD and MASLD can still be considered two different entities, and dietary habits may relate to them differently. Therefore, the aim of this study was to examine the associations between dietary fatty acids and MASLD, to understand if any of the expected associations are different when the outcome is slightly different.

### **4.2 Methods**

#### **4.2.1 Study sample, inclusion, and exclusion criteria,**

Participants were included if they had liver MRI data available and had completed at least 2 dietary assessments using the Oxford WebQ. To study MASLD as a binary outcome, the exclusion criteria and data sources used were the same as the previous study looking at liver fat content in Chapter 3, with the difference that cholesterol lowering medication users and participants living with diabetes were retained for this analysis, as they consist of



cardiometabolic factors that are part of the definition of MASLD (Figure 4.1). The cardiometabolic factors considered for the MASLD definition are high BMI, high waist circumference, diabetes, hypertension, and dyslipidaemia.

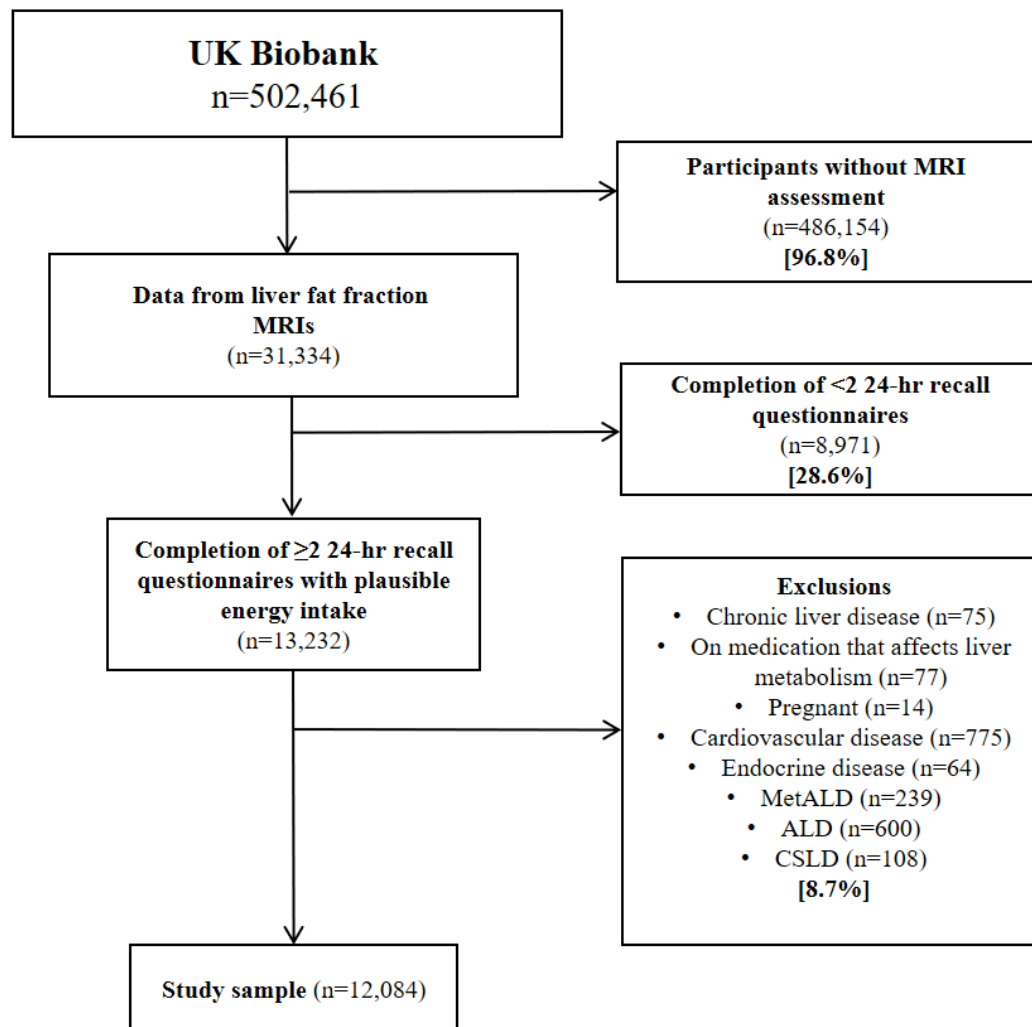


Figure 4.1 Flow chart with exclusion criteria.

#### 4.2.2 Steatotic liver disease (SLD) and categories

Participants were classified following the new SLD definitions proposed by the global consensus presented in 2023 which can be seen in Figure 4.2 (7). Other than MASLD, the rest of the SLD categories were cryptogenic (CSLD), MetALD (metabolic and alcohol associated steatotic liver disease), ALD (alcohol associated steatotic liver disease) and SLD due to specific causes such as drug induced liver injury. Steatosis was defined as >5.6% of liver fat content

measured by MRI. This was obtained from data from the first imaging visit in the UK Biobank, which in this sample were performed between 2014 and 2020. The cardiometabolic factors were defined as BMI $\geq$ 25 measured at the time of MRI assessment, waist circumference $>$ 80 cm in women or  $>$ 90 cm in men, diabetes, blood pressure lowering medication, or use of cholesterol lowering medication.

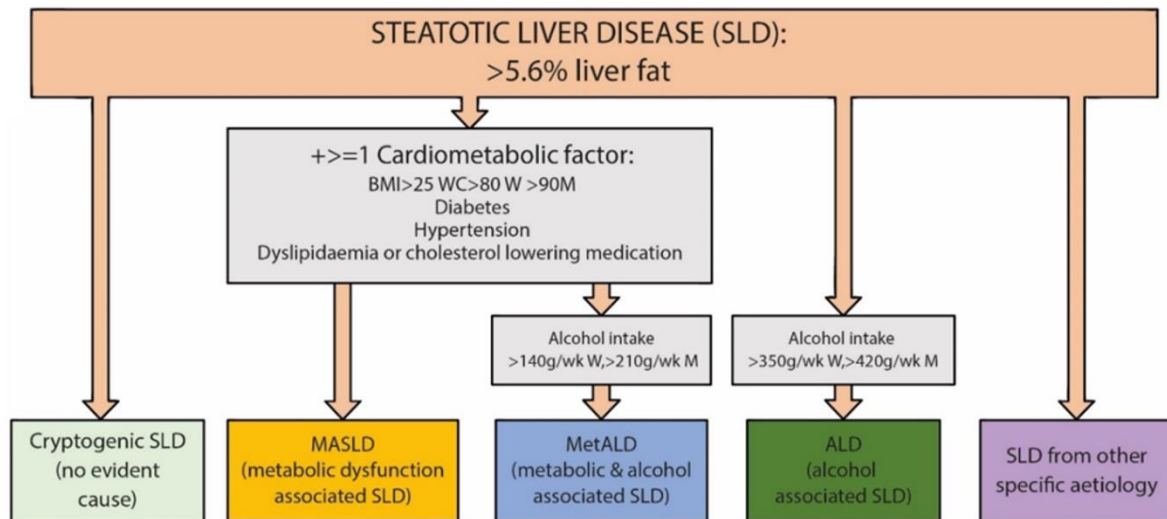


Figure 4.2 Classification of steatotic liver disease. W: women, M: men, WC: waist circumference.

The variables used to define steatosis and cardiometabolic factors can be found in Table 4.1. Furthermore, participants with alcohol intake surpassing the thresholds for MASLD were categorized under either MetALD (metabolic and alcohol associated steatotic liver disease) or ALD (alcohol related steatotic liver disease). MetALD, ALD, and CSLD cases were excluded from the final sample.

Table 4.1 Variables used to define MASLD.

	<b>Data-field</b>	<b>Additional information</b>
<b>Steatosis</b>		
Liver PDFF (fat %)	24352	Liver proton density fat fraction
<b>Cardiometabolic factors</b>		
Body mass index (BMI)	23104	Body mass index measured at time of MRI
Waist circumference	67	Waist circumference
High blood pressure	6177 and 6153	Medication for cholesterol, blood pressure or diabetes
	20002	Verbal interview: hypertension
Diabetes	6177 and 6153	Medication for cholesterol, blood pressure or diabetes
	20002	Verbal interview: diabetes diagnosed by a doctor

#### 4.2.3 Dietary fatty acids and individual fatty acids assessment

The dietary exposures considered were total fat, and different fatty acids, according to their structure: SFAs, MUFAs and PUFAs. Dietary fatty acids data was obtained from the Oxford WebQ questionnaires, which were carried out in at least two of five opportunities. The periods in which each were taken can be seen in Chapter 2: Figure 2.2. Fatty acid intake was automatically calculated from the participants' answers, and the mean of completed dietary questionnaires was obtained for each participant to calculate their mean intake per day. In addition, the individual fatty acids dataset from the UK Biobank provided information about 21 fatty acids (72). This facilitated the investigation into the relationship between 10 SFAs, 4 MUFAs, and 7 PUFAs and MASLD. More information about the calculation of nutrients obtained from the participants answers can be found in Chapter 2, Section 2.1.6. All exposures were coded as percentage of total energy intake.

#### 4.2.4 Covariates

Due to their relationship with liver fat and their relevance in previous literature, age, sex, ethnicity, education, physical activity, smoking status, total energy intake and fruits and vegetables intake were considered covariates to include in the analyses. These confounders followed the same coding as those in Chapter 3. More information about the coding and categories of covariates is presented in Chapter 2, Section 2.1.10.

#### 4.2.5 Statistical analysis

Multivariable logistic regression models were built to calculate the odds ratio of MASLD, per 5 % increase in dietary fatty acids. In minimally adjusted models (M1), adjustments were made for sex and age at recruitment. Models were further adjusted for ethnicity (M2); region, education and deprivation as markers of socioeconomic status (M3); smoking status (M4); physical activity (M5); total energy intake (M6) and fruits and vegetables intake (M7). Models were not adjusted by BMI as this was considered one of the cardiometabolic factors to define the outcome MASLD, or by alcohol intake as this data was used to classify participants into SLD categories. For individual fatty acids, multivariable logistic regression models assessed ORs of MASLD per 1SD increase of each of the 21 individual fatty acids.

#### 4.2.6 Subgroup analyses

For each of the dietary exposures, tests for heterogeneity across sex, VAT groups and menopause status were carried out. This was done by using LRTs that compare models with the variable as a confounder with a more complex one including it as an effect modifier.

#### 4.2.7 Sensitivity analysis: $\geq 4$ WebQ

As a sensitivity analysis, the participants who answered at least 4 Oxford WebQs were included, to allow for a better chance to capture consumption of food items that are usually eaten only once or a few times a week and prevent potential misclassification.

## 4.3 Results

### 4.3.1 SLD prevalence and final sample participant characteristics

Following the SLD definitions, 1,388 participants (10.7%) were categorised as MASLD in this sample (Table 4.2). Characteristics across all SLD categories were summarised to provide a comprehensive overview of the study population using the new SLD definitions, acknowledging their recent inclusion and the opportunity to explore these categories in the UK Biobank. The proportion of SLD cases other than MASLD (CSLD, MetALD, and ALD) combined was 7%. Subsequently, the sample was restricted to participants with MASLD and with no SLD for all analyses.

Table 4.2. Participant characteristics by SLD categories. N=13,031.

	No steatotic liver disease	CSLD	MASLD	MetALD	ALD	Total	p-value
	10,696 (82.1%)	108 (0.8%)	1,388 (10.7%)	N=239 (1.8%)	N=600 (4.6%)	N=13,031	
<b>Liver fat content (%)*</b>	2.4 (1.8,3.3)	7.2 (6.1,9.1)	8.8 (6.8,12.5)	8.7 (6.8,12.5)	9.6 (7.0,13.4)	2.7 (1.9,4.4)	<0.001
<b>Age (years at MRI)</b>	54.8 (7.5)	54.9 (7.2)	54.9 (7.2)	55.5 (7.2)	54.5 (7.1)	54.8 (7.4)	0.53
<b>Men^</b>	4,693 (43.9%)	73 (67.6%)	765 (55.1%)	195 (81.6%)	312 (52.0%)	6,038 (46.3%)	<0.001
<b>Body mass index*</b>	24.8 (22.7,27.2)	23.8 (23.1,24.5)	29.7 (27.3,32.6)	29.5 (27.5,31.6)	30.1 (27.3,33.4)	25.5 (23.1,28.3)	<0.001
<b>White ethnicity^</b>	10,377 (97.0%)	103 (95.4%)	1,354 (97.6%)	235 (98.3%)	576 (96.0%)	12,645 (97.0%)	0.055
<b>Most deprived^</b>	2,107 (19.7%)	14 (13.0%)	280 (20.2%)	46 (19.2%)	154 (25.8%)	2,601 (20.0%)	<0.001
<b>Professional qualification^</b>	1,122 (10.5%)	12 (11.1%)	176 (12.7%)	24 (10.0%)	82 (13.7%)	1,416 (10.9%)	<0.001
<b>London region^</b>	1,460 (13.6%)	7 (6.5%)	167 (12.0%)	24 (10.0%)	49 (8.2%)	1,707 (13.1%)	<0.001
<b>Current smoker^</b>	545 (5.1%)	7 (6.5%)	78 (5.6%)	23 (9.6%)	53 (8.8%)	706 (5.4%)	<0.001
<b>Low physical activity^</b>	2,100 (19.6%)	19 (17.6%)	343 (24.7%)	61 (25.5%)	174 (29.0%)	2,697 (20.7%)	<0.001
<b>Alcohol intake (g/wk.)*</b>	128.0 (64.0,216.0)	128.0 (66.0,200.0)	112.0 (52.0,180.0)	332.0 (304.0,368.0)	528.0 (448.0,644.0)	128.0 (64.0,224.0)	<0.001
<b>Total energy intake (kJ)</b>	8463.0 (1911.3)	8827.1 (1927.4)	8505.0 (1964.0)	9212.5 (1954.5)	8786.7 (2149.6)	8499.1 (1933.1)	<0.001
<b>Total fat (% total energy)</b>	31.3 (5.7)	31.7 (5.5)	31.9 (5.9)	30.5 (5.8)	31.6 (6.1)	31.4 (5.7)	<0.001
<b>SFAs (% total energy)</b>	11.3 (2.8)	12.0 (3.2)	11.7 (2.9)	11.2 (2.8)	11.7 (3.1)	11.4 (2.8)	<0.001
<b>PUFAs (% total energy)</b>	5.6 (1.5)	5.3 (1.3)	5.5 (1.4)	5.2 (1.2)	5.4 (1.4)	5.6 (1.4)	<0.001
<b>MUFAs (% total energy)</b>	11.8 (2.5)	11.8 (2.2)	12.1 (2.6)	11.6 (2.5)	12.0 (2.6)	11.9 (2.5)	<0.001

Values are mean (SD), unless otherwise specified. ^N (proportion %) \*Values are median (IQR) SFAs: saturated fatty acids PUFAs: polyunsaturated fatty acids MUFAs: monounsaturated fatty acids kJ: kilojoules Low physical activity: <10 metabolic equivalent hours/week

After excluding non-MASLD cases of SLD, the final sample consisted of 12,084 participants (Table 4.3). MASLD participants with 1 cardiometabolic factor represented 56% of the cases (n=782), participants with 2 cardiometabolic factors represented 28 % (n=391), 13% had 3 (n=176), and 3% had 4 (N=39).

Table 4.3 Participant characteristics by MASLD status.

	No steatosis	MASLD	Total	p-value
	N=10,696	N=1,388	N=12,084	
<b>Liver fat content (%)</b>	2.4 (1.8,3.3)	8.8 (6.8,12.5)	2.6 (1.9,3.9)	<0.001
<b>Age at MRI (years)</b>	54.8 (7.5)	54.9 (7.2)	54.8 (7.4)	0.77
<b>Men</b>	4,693 (43.9%)	765 (55.1%)	5,458 (45.2%)	<0.001
<b>Body mass index (kg/m<sup>2</sup>)</b>	24.8 (22.7,27.2)	29.7 (27.3,32.6)	25.2 (22.9,27.9)	<0.001
<b>Current smoker</b>	545 (5.1%)	78 (5.6%)	623 (5.2%)	0.30
<b>Low physical activity*</b>	2,100 (19.6%)	343 (24.7%)	2,443 (20.2%)	<0.001
<b>Alcohol intake (g/wk.)</b>	128.0 (64.0,216.0)	112.0 (52.0,180.0)	128.0 (64.0,208.0)	<0.001
<b>Diabetes</b>	164 (1.5%)	88 (6.3%)	252 (2.1%)	<0.001
<b>Total energy intake</b>	8463.0 (1911.3)	8505.0 (1964.0)	8467.8 (1917)	0.44
<b>Total fat (% of energy intake)</b>	31.3 (5.7)	31.9 (5.9)	31.4 (5.7)	<0.001
<b>SFAs (% of energy intake)</b>	11.3 (2.8)	11.7 (2.9)	11.3 (2.8)	<0.001
<b>PUFAs (% of energy intake)</b>	5.6 (1.5)	5.5 (1.4)	5.6 (1.4)	0.16
<b>MUFAs (% of energy intake)</b>	11.8 (2.5)	12.1 (2.6)	11.9 (2.5)	<0.001

\*<10 Metabolic-equivalent hour/week

Values are N (proportion %), or mean (SD)

^Values are median (IQR) SFAs: saturated fatty acids PUFAs: polyunsaturated fatty acids MUFAs: monounsaturated fatty acids  
kJ: kilojoules

The most prevalent cardiometabolic factor was high BMI or high waist circumference, present in 1,346 of the MASLD cases (97%), while diabetes was the least prevalent, present in 88 cases (6%). Those in the MASLD group reported a higher intake of total fat, SFAs, and MUFAs. There was no significant difference in PUFAs intake between cases and participants without steatosis.

#### 4.3.2 Dietary fatty acids and MASLD

The results obtained in every model after adjusting for confounders can be observed in Figure 4.3. There were minimal changes in the associations across models as confounders were included. In fully adjusted models, (Figure 4.4), a 5% increase in total fat presented a positive association with MASLD: OR 1.10 (1.05-1.16). A 5% increase in SFAs and MUFAs was associated with higher odds of MASLD, with OR of 1.25 (1.12-1.39) and 1.25 (1.12-1.40) respectively. Dietary PUFAs did not show any significant association with MASLD, in crude or adjusted models (OR: 0.92 [0.76-1.11]).



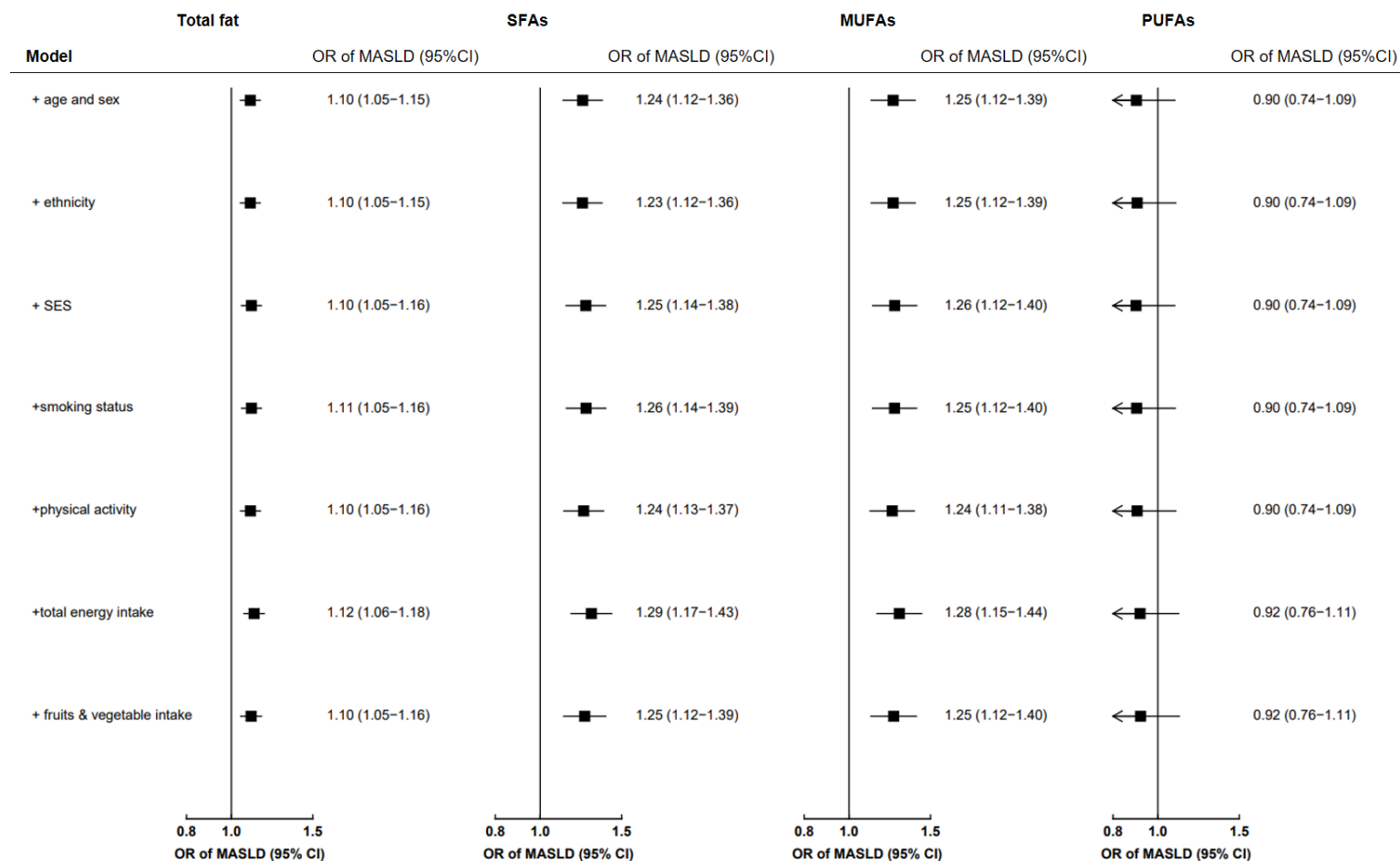


Figure 4.3 Sequential adjustment for confounders (n=12,084). Odds of MASLD per 5% increase in dietary fatty acids.

SFAs: saturated fatty acids PUFAs: polyunsaturated fatty acids MUFAs: monounsaturated fatty acids

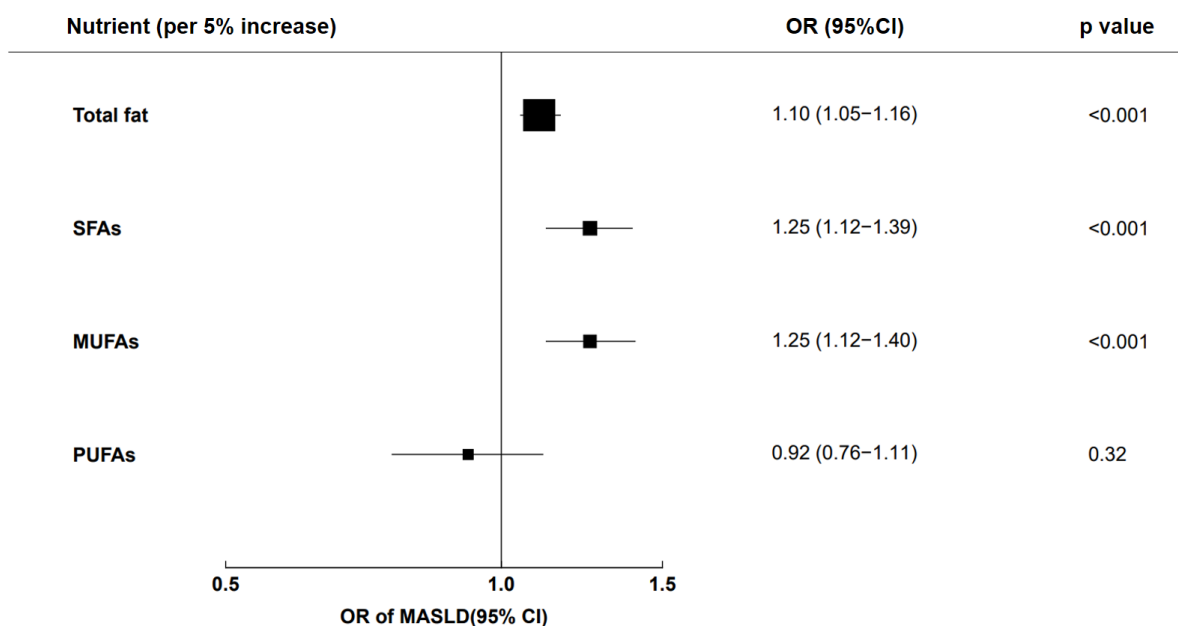


Figure 4.4 OR of MASLD (N=12,084), per 5% increase in dietary intake. 1,388 cases. Adjusted for sex, age, ethnicity, deprivation, education, region, smoking status, physical activity, total energy intake, fruits and vegetables intake.

#### 4.3.3 Individual dietary fatty acids and MASLD

Only linoleic, oleic, palmitic and stearic acid, as well as total omega 3 PUFAs and omega 6 PUFAs represented more than 1% of total energy intake, so only their results are considered relevant for this study. An increase in 1SD of stearic acid increased the odds of MASLD by 19% (OR: 1.19 [1.12, 1.26]) (Figure 4.5). Palmitic and oleic acid also presented positive associations with MASLD, with an increase in the odds of MASLD of 16% (OR: 1.16 [1.09-1.23]), and 13% (OR: 1.13 [1.06-1.19]), respectively).

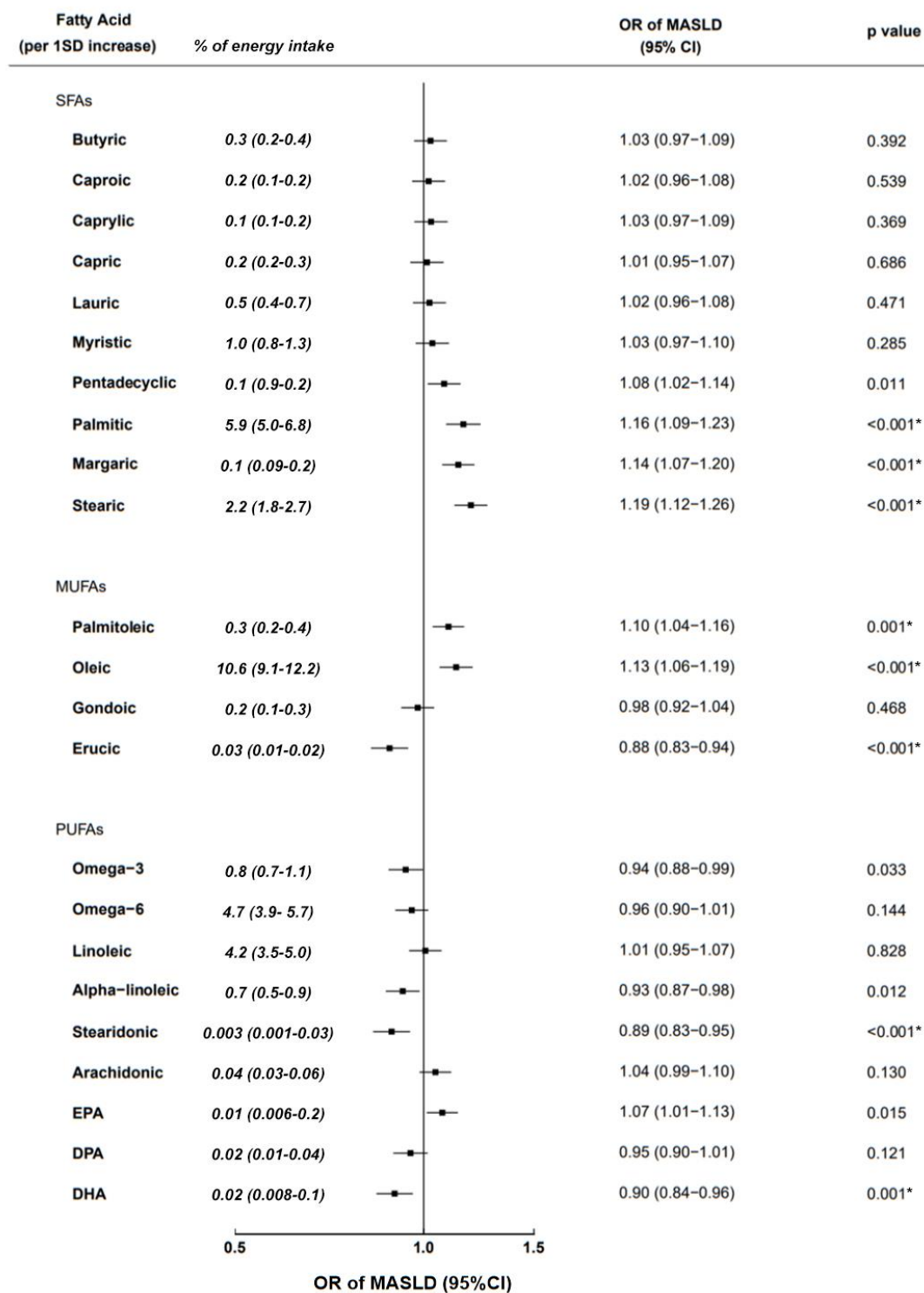


Figure 4.5 OR of MASLD (N=12,084), per 1SD increase in individual fatty acid intake. 1,388 cases of MASLD. Adjusted for sex, age, ethnicity, deprivation, education, region, smoking status, physical activity, total energy intake, fruits and vegetables intake.

EPA: eicosapentanoic acid DPA: docosapentanoic acid, DHA: docosaheaxaenoic acid.

#### 4.3.4 Subgroup analysis

No evidence of heterogeneity was observed across menopause status, visceral adiposity groups, and sex, for any of the fatty acid intake (Table 4.4).

Table 4.4 OR of MASLD per 5% of dietary fatty acids: subgroup analyses and tests for interaction

	<b>Total fat</b>	<b>SFAs</b>	<b>MUFAs</b>	<b>PUFAs</b>
<b>Overall</b>	1.10 (1.05 , 1.16)	1.25 (1.12 , 1.39)	1.25 (1.12 , 1.40)	0.92 (0.76 , 1.11)
<b>VAT&gt;2.8</b>	1.10 (1.04 , 1.16)	1.29 (1.15 , 1.45)	1.17 (1.03 , 1.32)	0.96 (0.77 , 1.20)
<b>VAT&lt;2.8</b>	1.04 (0.84 , 1.29)	1.27 (0.83 , 1.95)	1.08 (0.66 , 1.74)	0.59 (0.25 , 1.42)
<b><i>p heterogeneity</i></b>	<i>0.63</i>	<i>0.94</i>	<i>0.75</i>	<i>0.29</i>
<b>Men</b>	1.15 (1.07 , 1.23)	1.35 (1.18 , 1.56)	1.32 (1.13 , 1.55)	0.97 (0.73 , 1.28)
<b>Women</b>	1.06 (0.98 , 1.14)	1.14 (0.98 , 1.33)	1.18 (1.00 , 1.38)	0.87 (0.66 , 1.15)
<b><i>p heterogeneity</i></b>	<i>0.12</i>	<i>0.10</i>	<i>0.31</i>	<i>0.61</i>
<b>Premenopausal</b>	1.02 (0.94 , 1.11)	1.04 (0.87 , 1.24)	1.12 (0.92 , 1.35)	1.12 (0.92 , 1.35)
<b>Postmenopausal</b>	1.05 (0.89 , 1.23)	1.09 (0.78 , 1.53)	1.16 (0.82 , 1.64)	1.16 (0.82 , 1.64)
<b><i>p heterogeneity</i></b>	<i>0.78</i>	<i>0.84</i>	<i>0.85</i>	<i>0.72</i>

#### 4.3.5 Sensitivity analysis

After excluding participants with less than 4 dietary questionnaires, previous observation regarding the associations between total dietary fat intake and odds of MASLD were maintained (OR: 1.23 [1.11-1.36]). However, sensitivity analysis resulted in greater odds of MASLD per 5% increase in SFAs (OR: 1.37 [1.11-1.68]) or MUFAs intakes (OR: 1.64 [1.31- 2.05]) (Figure 4.6). The results for dietary PUFAs intake remained non-significant.

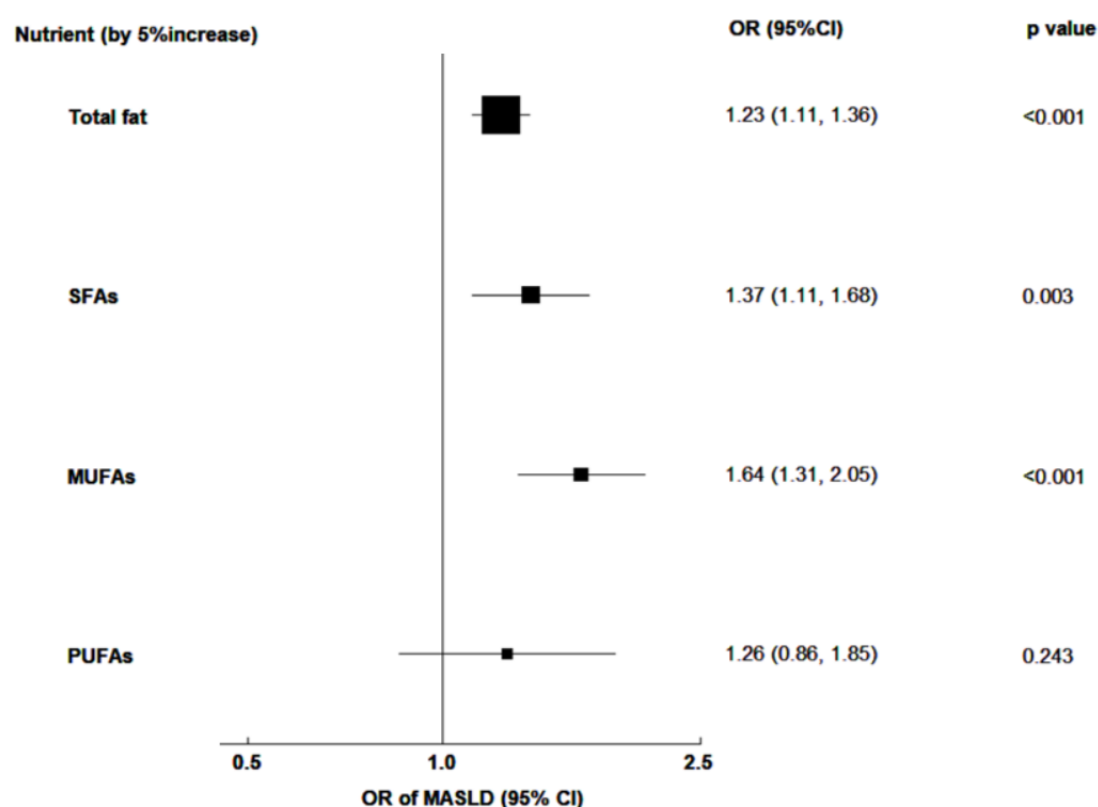


Figure 4.6 Sensitivity analysis. OR of MASLD (N=3,649), restricted to those who answered  $\geq 4$  WebQs. 424 cases. Adjusted for sex, age, ethnicity, deprivation, education, region, smoking status, physical activity, total energy intake, fruits and vegetables intake.

The results of the sensitivity analyses for individual fatty acids as exposures can be found in Figure 4.7. The results for palmitic, stearic and margaric acid became slightly stronger and remained significant.

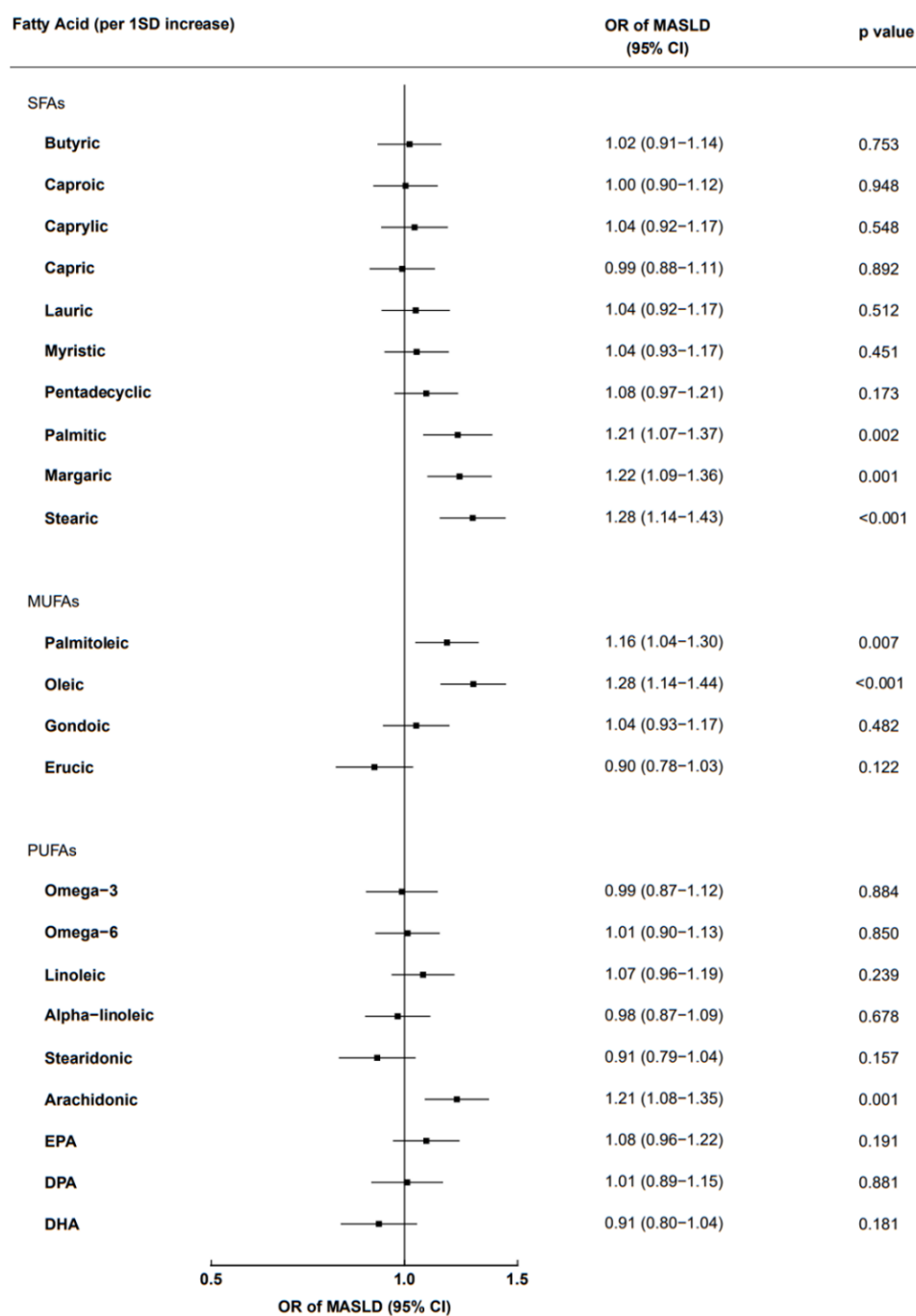


Figure 4.7 Sensitivity analysis, restricted to those who answered  $\geq 4$  WebQs. Odds ratio of MASLD (95%CI) per 1SD increase in dietary fatty acids (N=3,649).

EPA: eicosapentanoic acid DPA: docosapentanoic acid, DHA: docosahexaenoic acid.

## 4.4 Discussion

### 4.4.1 Main findings

This study analysed 12,031 participants from the UK Biobank and found that a 5% increase of dietary total fat, SFAs or MUFAs was associated with higher odds of MASLD, while there were no significant associations between dietary PUFAs and MASLD. When investigating intake of individual fatty acid it was observed that the association between SFAs and odds of MASLD were mainly driven by palmitic, margaric and stearic acid, whereas in regards MUFAs palmitoleic and oleic acid were major contributing factors. These are the individual fatty acids with the highest consumption within the fatty acid group in this population, which would explain the significant results. These results were consistent in sensitivity analyses, stronger in the case of MUFAs, and SFAs, and there was no observation of heterogeneity across sex, VAT groups or menopause status for any of the dietary exposures.

### 4.4.2 Comparison with previous studies

To my knowledge no previous studies have investigated the intake of dietary fatty acids and odd of MASLD, and as such these findings are novel. We can however compare the results to those that have previously studied NAFLD or liver fat as outcomes. In this sample, a 5% increase in SFAs was positively associated with MASLD, with an OR of 1.25 (1.12-1.39). This is consistent with many observational and interventional studies that assessed the role of SFAs, and found positive associations with liver fat (46, 47). The interventional studies carried out by Parry et al. Luukkonen et al. and Bjermo et al in controlled settings studied small samples that varied between 16 and 61 participants and observed higher liver fat in participants after they consumed high SFAs diets (31, 44, 45, 48). My study adds a population based perspective in which more participants can be studied, and their dietary reports reflect their diet in their daily lives, for a longer period of time. Chen et al. also studied SFAs and liver fat in the UK Biobank, and observed an OR of NAFLD of 1.27 (1.15-1.40) per 5% increase in dietary

SFAs (31). While my study presents very similar results, it provides a more precise and up-to-date outcome by using MASLD, while additionally analysing more reliable dietary intake estimates of at least 2 dietary questionnaires instead of 1. In addition, Friden et al found a positive association between SFAs and liver fat, but also included participants with one dietary questionnaire, which is not considered ideal in population based studies (130).

A study published by Tian et al. found that a dietary pattern high in PUFAs is associated with a lower risk of NAFLD in the UK Biobank (33). However, the results in this thesis do not show any associations between PUFAs intake and MASLD in either crude or fully adjusted models. This may be due to the fact that the MASLD definition allows for a better identification of cases, excluding other causes of steatosis more precisely than NAFLD, which is an exclusion diagnosis. Stearidonic acid presented an inverse association with MASLD in this study. However, due to its very low intake in this population (less than 0.002% of total energy intake), it is likely that these observations may be difficult to interpret and may not reflect a real association, but an incidental finding.

I found significant higher odds of MASLD per 5% increase of MUFAs in fully adjusted models and sensitivity analyses, which has not been observed in previous studies. Other observational studies did not find any significant associations (35-37, 41, 42) with any of the dietary fatty acids and liver fat. This may be due to their small samples (between 73 and 349 in the case of Allard et al, Zelber et al., Lopez Bautista et al) (35, 41, 42), less robust statistical analyses, or lack of a clear outcome definition in the case of Nouredin et.al. (36).

The lack of heterogeneity across VAT groups for all dietary exposures could be explained by the strong correlation between VAT and waist circumference (55, 131), one of the cardiometabolic factors to define MASLD.



#### 4.4.3 Strengths and limitations

This is the first study to explore the associations between dietary fatty acids and MASLD, and to further study these by looking at individual fatty acids and MASLD. The use of MASLD as a main outcome provides a positive identification of disease, moving away from the exclusionary criteria presented when studying NAFLD (which only considered steatosis excluding excessive alcohol intake) (7). This provides results that look at an outcome that is up to date with recent EASL guidelines, and considers participants from an overall metabolic perspective, rather than isolating liver fat content (101). The SLD classification clearly differentiates participants with excessive alcohol consumption into ALD and MetALD categories, which prevents the effect of alcohol to spear the associations studied. Additionally, the detailed data available from the UK Biobank recruitment assessment, along with linkage with hospital admissions data, allowed to account for relevant exclusions and confounders, and to consider multiple cardiometabolic factors, key for MASLD diagnosis. The use of MRIs provided a non-invasive and accurate measurement of steatosis (80). However, it is important to note that this represents a simplified definition of MASLD, as the same BMI categories were applied across all ethnicities, and diabetes was defined without incorporating HbA1c data or primary care diagnoses due to data availability constraints.

The data obtained from the Oxford WebQ provided detailed information about dietary fatty acids and individual fatty acids. Performing sensitivity analyses restricted to  $\geq 4$  dietary assessments was particularly relevant in the case of PUFAs, as their main sources are certain types of food products which in the UK may be consumed approximately one portion per week, or even less (107, 122). The multiple WebQ completed to measure diet attempted to account for variation in nutrient intake across time by administering it throughout different times of the week and year, in a way that prevents response fatigue (70). In addition, the use of the 21 individual fatty acid dataset allowed to observe if there were specific fatty acids among the

respective structural groups that were driving the associations more strongly. It is important to consider that self-reported dietary intake is likely to introduce measurement error and information bias, because food products considered unhealthy and rich in saturated fatty acids may have been underreported (130). Therefore, the associations observed may have been biased towards the null, reducing their strength.

One of the main limitations to consider is that dietary assessment and liver fat assessment were not done simultaneously, as mentioned in Chapter 3. Liver fat was measured multiple years after the WebQ, allowing time for diet to change. This was addressed by the exclusion of cardiovascular disease hospital episodes, and liver disease hospital episodes that could have affected diet significantly during those years, due to medical advice. In addition, the mean change in BMI between baseline and liver fat was calculated, and it was a mean 4.3% change.

Cholesterol lowering medication and blood pressure medication were self-reported at baseline, while diabetes was self-reported and included as any incident cases in hospital episodes. Therefore, there may have been changes that were not captured such as the prescription of cholesterol lowering medication or blood pressure, or a new diabetes diagnoses which does not require hospital admission.

It is important to consider that this population is mainly of White ethnicity (>90%), and UK based which does not allow for generalisation of these results in other populations. As previously mentioned, the low response rate of 5 % and the characteristics of the participants make it a sample susceptible to selection bias. In this study, the willingness to respond multiple dietary questionnaires, and even more than 4 could also be related to a group that is overall healthier (70). However, it has been shown that results from studies using the UK Biobank could be generalizable, as they have been consistent with studies that were adequately representative of the population of interest (124, 125).

## 4.5 Conclusions

Dietary intake of SFAs and MUFAs had a positive association with MASLD in the UK Biobank, while PUFAs intake was not significantly associated with MASLD. A better understanding of the role of dietary fatty acids and liver fat content will be key for MASLD prevention and improvement. Future research assessing dietary intake and liver fat content can be advanced: ideally, diet and imaging assessments can be measured with a short period of time between them. It will also be necessary to conduct long-term large prospective studies in more diverse populations, and to perform interventional studies that look at the effect of dietary fatty acids on liver fat content in participants with cardiometabolic factors. This could inform dietary guidelines and allow a better approach to MASLD prevention and treatment.

## 5 Plasma fatty acids and liver fat content

### 5.1 Introduction

I have described the associations between dietary fatty acids and different measures of liver fat and steatosis: liver fat content and MASLD in Chapters 3 and 4. As described in Chapter 1, Section 1.2, both diet and adipose tissue contribute to plasma FAs. Without data from dietary intake, plasma FAs can be considered an objective marker of diet, as they could reflect dietary intake indirectly (132). Therefore, levels of plasma fatty acids can be used as potential biomarkers of diet (132).

#### 5.1.1 Previous studies

Previous studies have observed that total plasma MUFAs% and SFAs% were associated with higher liver fat content in the UK Biobank, and that PUFAs% are associated with lower liver fat content (32, 133). Another study found that plasma PUFAs% was inversely associated with liver fat (134), but liver fat was measured indirectly by the fatty liver index (Table 5.1). The hepatic fatty acid metabolic pathways are different between fasted and postprandial state, which could influence the associations investigated (24, 96, 135, 136).

Table 5.1 Studies investigating plasma fatty acids and liver fat.

Study	Population	Plasma fatty acid assessment	Liver fat outcome	Covariates	Findings
Gnautic, 2023(133)	UK, N=1,088	NMR	Liver fat by MRI (continuous)	Age, sex, ethnicity, deprivation, physical activity, alcohol intake, smoking status, total energy intake. stratified by diabetes and BMI	Plasma SFAs and MUFAs presented positive associations, PUFAs presented inverse associations.
Friden, 2023(32)	UK, N=9,119	NMR	Liver fat by MRI (continuous)	Smoking, physical activity, age, stratified by sex.	Plasma SFAs and MUFAs presented positive associations, PUFAs presented inverse associations.
Makela 2022(134)	N= 1,533	Gas chromatography	Fatty liver index	Age, sex, physical activity, smoking, alcohol intake, total energy and carbohydrate intake	PUFAs was inversely associated with liver fat
Lee et al. ,2015 (137)	N=24	Gas chromatography	Liver fat by MR spectroscopy	Spearman correlation	VLDL-TG palmitoleic acid was associated with liver fat

### 5.1.2 VLDL-TG fatty acids

This analysis will study the proportion of individual fatty acids in VLDL-TG, described in Figure 5.1. Dietary fatty acids are absorbed in the duodenum, and are transported as chylomicrons, which are large triglyceride-rich lipoproteins. Chylomicron remnants enter the liver, and they mix with fatty acids coming from multiple sources: those from lipolysis of adipose tissue (breakdown of tissue into fatty acids), and fatty acids newly synthesized in the liver, by de novo lipogenesis (26). In addition FA from chylomicrons can enter the liver via *spillover* (138).

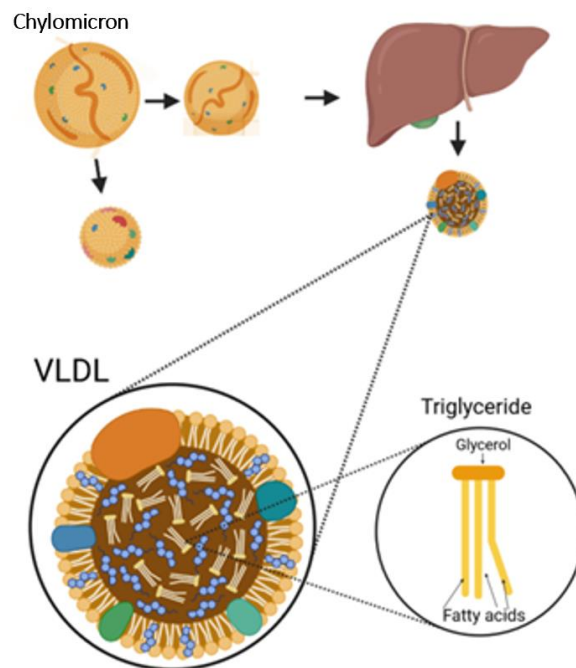


Figure 5.1 Very-low density lipoprotein (VLDL) triglyceride fatty acid composition. Created with biorender.com. FAs from this pool within the liver can then be partitioned into either an oxidation pathways in which they produce energy, or can enter an esterification pathway. FAs can be esterified to form TG which can become part of the liver TG pool in the cytosol (and remain in the liver) or can be incorporated into VLDL-TG, which are secreted from the liver into the bloodstream (24). Therefore, studying VLDL-TG fatty acid composition can provide additional information about the FAs that enter the liver FA pool.

There is evidence that VLDL-TG export is higher in livers with high levels of liver fat, due to the altered lipid metabolism (139). In addition, an observational study of 24 adults living with obesity overweight found that high liver steatosis, measured with MRI, was associated with high VLDL-TG palmitoleic acid and VLDL-TG linoleic acid in comparison to lower liver fat (137). Therefore, a steatotic liver may export more VLDL-TG than a liver without steatosis, and at the same time this may provide information about the types of fatty acids that are in the liver, showing if any is particularly more abundant. However, despite this observation, data is

limited regarding potential differences in the composition of fatty acids in the VLDL-TG in steatotic vs non-steatotic livers. This analysis will study the associations between steatosis and VLDL-TG fatty acid composition in a sample in the UK.

### 5.1.3 Aims

This chapter aims to assess the associations between plasma fatty acids (SFAs%, PUFAs%, MUFAs% and individual fatty acids %) and liver fat content in fasted and non-fasted samples, which will be done in the UK Biobank and in a sample from OCDEM. In addition, it aims to explore the associations between liver steatosis (liver fat>5.6%) and VLDL-TG fatty acid composition (VLDL-TG SFAs%, VLDL-TG MUFAs%, VLDL-TG PUFAs% and VLDL-TG individual fatty acids) at OCDEM. Finally, it aims to assess the associations between steatosis and blood biomarkers of cardiovascular risk (total cholesterol, LDL cholesterol, HDL cholesterol, VLDL cholesterol, non-HDL cholesterol and total triglycerides).

## 5.2 Methods

### 5.2.1 Study population

Two different datasets of participants were included in this analysis. The first group includes participants from the UK Biobank cohort, and the second group consisted of a dataset of participants that had previously been part of interventional studies performed by Professor Leanne Hodson and her team at the Oxford Centre of Diabetes, Endocrinology and Metabolism (OCDEM). Data obtained during assessment visits was stored and this study analysed a dataset that consists of participants from eight different interventional studies. It will be referred to as the OCDEM group, a sample of 125 participants, from different trials described in Chapter 2, Section 2.2.

Participants from the UK Biobank cohort were included in this analysis if they had data available for liver MRI and plasma fatty acids. Plasma was obtained from blood samples

collected at baseline. Individuals with prevalent chronic liver disease, diabetes, pregnancy, cardiovascular disease, endocrine diseases, high-risk drinkers, and cholesterol-lowering medication were excluded. Figure 5.2 describes the inclusion and exclusion criteria. Additional information about the UK Biobank dataset and variables code used for identification of exclusions can be found in Chapter 2, Section 2.1.9.

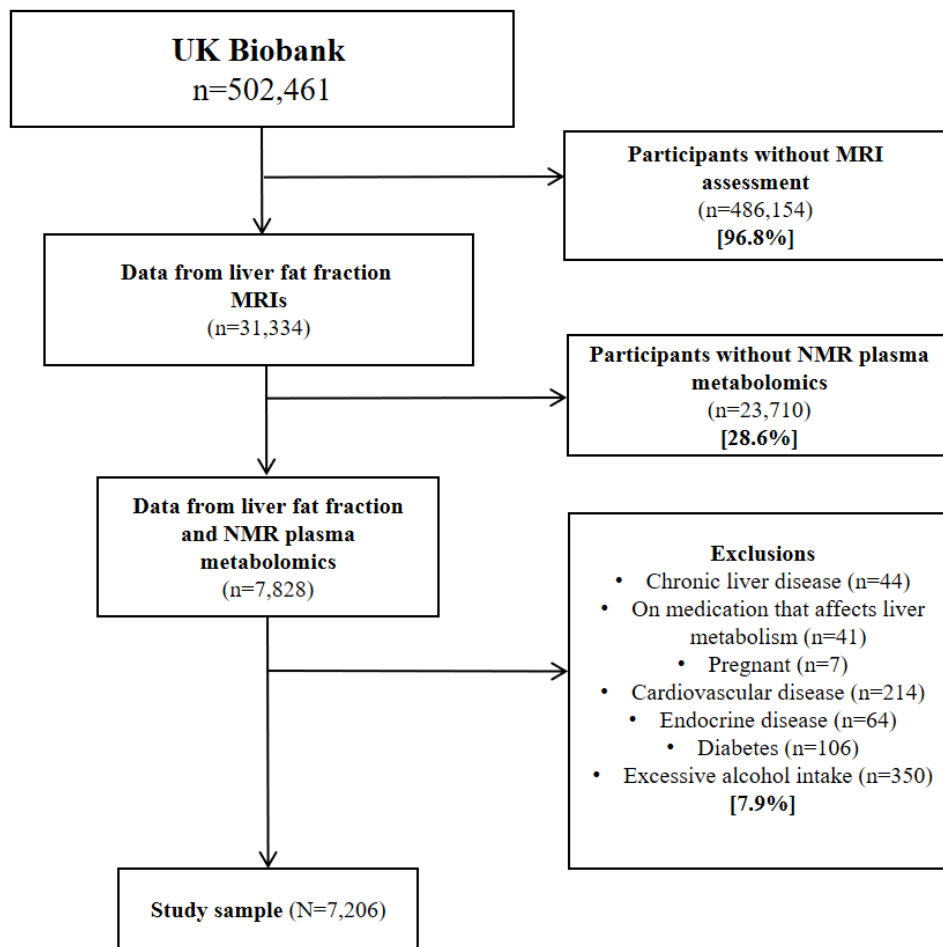


Figure 5.2 Flowchart of UK Biobank showing exclusion of participants included

### 5.2.2 Liver fat assessment

To measure liver fat content (%) in the UK Biobank, the liver fat content measurements were obtained from the abdominal MRI carried out as part of the first UK Biobank imaging visit. The measurements were made by AMRA, in ten different centres across the UK, and more details about how liver fat was estimated can be found in Chapter 2, Section 2.1.8 (79).



To measure liver fat content in OCDEM, a 3 Tesla MRI scanner, using proton spectroscopy, was used at the Oxford Centre of Magnetic Resonance at the John Radcliffe Hospital in Oxford. A single voxel (20x 20 x 20 mm) was positioned in the posterior part of the left liver lobe, and both water-suppressed and non-water-suppressed stimulated acquisition mode (STEAM) measurements were performed. The proportion of fat in the liver tissue was determined using the OXSA toolbox (94). Participants with liver fat >5.6% were considered to have steatosis (140).

### 5.2.3 NMR plasma fatty acids

Plasma fatty acids represent all fatty acid pools found in plasma, which consist of free fatty acids and any fatty acids contained in lipoproteins (132). Data obtained was collated to calculate total SFAs%, PUFAs% and MUFAs%, of plasma total fatty acids. For the UK Biobank sample, plasma was obtained at baseline during the first assessment visit, with a mean difference of 9 years before the liver MRI assessment. As described in Chapter 2, fasting was not required, and the time elapsed since last meal was documented. For the OCDEM samples, plasma fatty acid analysis was also determined by Nuclear Magnetic Resonance based metabolomics by Nightingale Health. These measurements, along with anthropometric measurements were taken within a week from the liver MRI assessment. More details on the OCDEM study visits are in Chapter 2.

### 5.2.4 VLDL-TG fatty acid composition measurement at OCDEM

To measure VLDL-TG fatty acid composition, plasma from blood samples was collected, and VLDL were isolated through ultracentrifugation. A lipid extraction was carried out, and fatty acid methyl esters were formed in order to use a gas chromatography analysis machine, which produced chromatogram peaks. These peaks were used to identify the fatty acids present in the samples, by comparing them to standards, and fatty acids were presented as mol%. More

information about the VLDL isolation and lipid extraction can be found in Chapter 2, Section 2.2.5.

### 5.2.5 Markers of cardiovascular risk at OCDEM

Participants' blood samples were also assessed by Nightingale to obtain the NMR concentrations of triglycerides and lipoproteins that are associated with a higher cardiovascular risk. The outcomes included were total cholesterol, LDL cholesterol, HDL cholesterol (which is suggested to be a marker of lower risk), VLDL cholesterol, non-HDL cholesterol and total triglycerides. A summary of all exposures and outcomes can be observed in Figure 5.3.

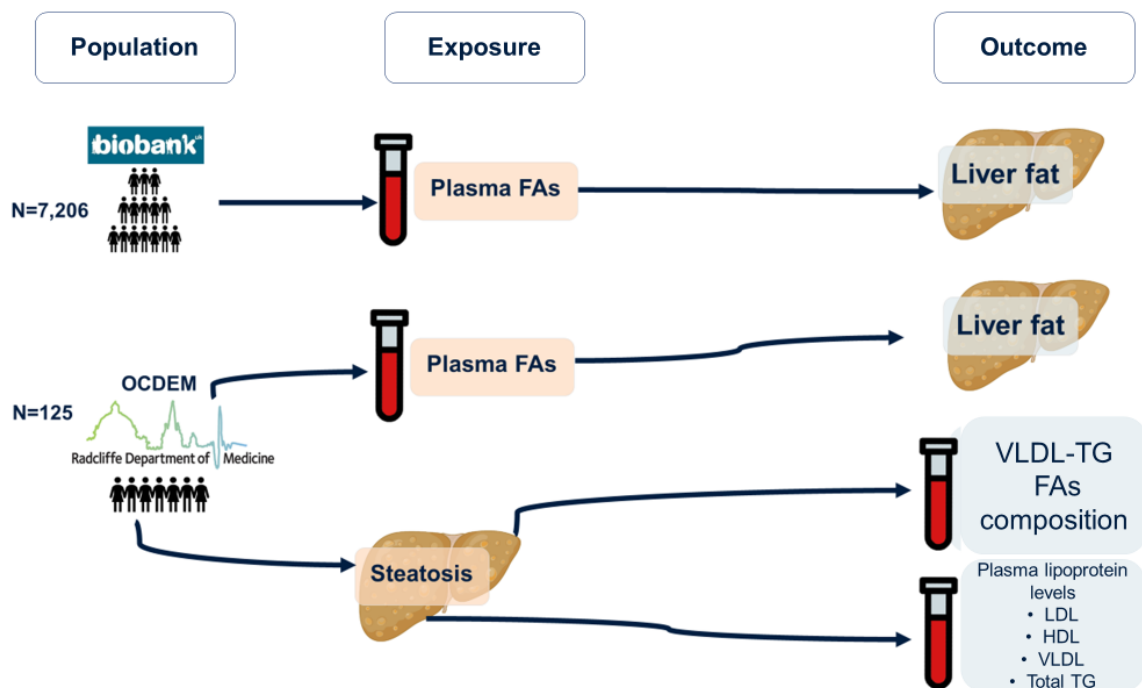


Figure 5.3 Summary of exposures and outcomes analysed in this chapter. FAs: fatty acids.

### 5.2.6 Covariates

In the UK Biobank, participants had data available about education, ethnicity, region, medication, as well as lifestyle characteristics about alcohol consumption and physical activity, that were taken collected during the assessment visit, at the same time as their blood samples.

In OCDEM, anthropometric measurements including BMI were taken during the assessment visit, and data was obtained about sex and age. There was no data about more detailed lifestyle or demographic characteristics, as the exclusion criteria for the individual studies was very specific and any excessive alcohol consumption or chronic conditions were excluded.

## 5.2.7 Statistical analyses

### 5.2.7.1 NMR plasma fatty acids and liver fat

Due to its skewed distribution, liver fat was transformed by calculating its natural logarithm in both sets of participants. For both UK Biobank and OCDEM samples linear regression models were built to estimate the geometric mean percentage difference in liver fat per 1% increase in plasma fatty acids. Beta coefficients obtained were transformed to obtain geometric mean % differences.

Models were adjusted for sex, age, ethnicity, deprivation, education, smoking, physical activity, alcohol risk and BMI in the UK Biobank sample, while age, sex and BMI were the only confounders included in the models in the analysis of the OCDEM sample. Results were presented in minimally adjusted models and for sequential adjustment of confounders, to assess the impact of the addition of covariates. Heterogeneity by sex and BMI (included as a continuous variable) was tested. This was carried out by including the variable as a confounder in a model, and then comparing it with a model in which the variable is introduced as an interaction term with LRT. The choice of linear exposures was informed by carrying out tests for departure of linearity. This was done by comparing models with the exposure as a continuous variable, with a model with the exposure as quintiles, using a LRT.

Assumptions of linearity were checked and none of them were violated. This was carried out by confirming the normality of residuals using Q-Q plots, with a mean of zero, plotting fitted versus residual values and checking for homoscedasticity, which are presented

in Ancillary Figures A1-A3. In addition, correlation matrices were calculated to assess multicollinearity, and no explanatory variables were highly correlated to each other.

#### 5.2.7.2 Sensitivity analysis

From the UK Biobank sample, a sensitivity analysis was carried out studying participants who reported having fasted for more than 8 hours at the time of the blood sample collection. These participants were studied to understand the effect of fasting in their blood measurements, and to compare them to the OCDEM sample, all of whom were sampled in the fasting state.

#### 5.2.7.3 Liver steatosis and VLDL-TG fatty acid composition

Wilcoxon rank tests were carried out to compare the VLDL-TG fatty acid % of SFAs, MUFAs, and PUFAs across participants with and without steatosis. The violation of assumptions of linearity can be observed in the Appendix, Ancillary Figures A4-A16, which confirmed the need of categorising the outcome. Participants were categorized into “High” and “Low” VLDL-TG fatty acids according to being higher or lower than the median of that fatty acid. Logistic regression models were built to estimate the odds of high levels of different VLDL-TG fatty acids% for participants with steatosis in comparison to controls. In addition, the roles of specific fatty acids within VLDL-TG were assessed independently of the structural group they belong to. This included myristic acid, palmitic acid, palmitoleic acid, stearic acid, oleic acid, linoleic acid, and alpha-linolenic acid.

#### 5.2.7.4 Liver steatosis and markers of cardiovascular risk

Due to violation of assumptions of linear associations, observed in Ancillary Figures A17-21, logistic regression models were built to estimate the odds of high TG, high VLDL, and high LDL, low HDL, high non-HDL and high Total cholesterol for participants with steatosis in comparison to controls. Participants were classified according to whether they were higher or lower than the median for each variable.

The significance level was set at 0.05 for all analyses. All statistical analyses were performed using Stata version 18.5 StataNow (Stata Corp, TX, United States)(141). Figures were created using RStudio Version 4 (R Core Team, Vienna, Austria), with package Jasper, and with BioRender.com (117).

## 5.3 Results

### 5.3.1 Participants characteristics

The final sample from the UK Biobank consisted of 7,206 participants, and their characteristics are presented in Table 5.2. The characteristics of the 125 participants of the OCDEM sample are presented in Table 5.3.

In the UK Biobank, the mean fasting time was 3 hours, whilst in the OCDEM studies participants were required to fast overnight before coming into the CRU. The mean percentages of plasma fatty acids were similar in both samples. On average, for both the OCDEM and UK Biobank samples, participants with steatosis had slightly higher plasma MUFAs% and SFAs% and slightly lower plasma PUFAs% than those without steatosis.

Table 5.2. Characteristics of UK Biobank participants by steatosis.

	<b>No steatosis</b> N=5,439	<b>Steatosis</b> N=1,767	<b>Overall</b> N=7,206	<b>p-value *</b>
Liver fat content <sup>a</sup>	2.5 (2.0,3.4)	9.7 (7.1,14.4)	3.0 (2.1,5.5)	<0.001 <sup>b</sup>
Age (years) <sup>c</sup>	63.2(7)	65.2 (7)	64.2 (7)	0.11 <sup>d</sup>
Men	2,376 (44%)	1,101 (62%)	3,477 (48%)	<0.001 <sup>b</sup>
Body mass index (BMI kg/m <sup>2</sup> ) <sup>a</sup>	25.3 (23.2,27.7)	28.9 (26.3,31.8)	26.0 (23.7,28.9)	<0.001
White ethnicity	5,280 (97.1%)	1,714 (97.0%)	6,994 (97.1%)	0.61
Least deprived	1,087 (20%)	354 (20.1%)	1,441 (20.0%)	0.19
Current smoker	295 (5%)	136 (7.7%)	431 (6.0%)	<0.001
University degree	4,180 (77%)	1,264 (71.5%)	5,444 (75.5%)	<0.001
Harmful alcohol intake <sup>^</sup>	203 (3.7%)	63 (3.6%)	266 (3.7%)	
Low physical activity	1,056 (19.4%)	477 (27.0%)	1,533 (21.3%)	<0.001
Total plasma fatty acids (mmol/l) <sup>c</sup>	11 (2.1)	12 (2.5)	11 (2.)	<0.001 <sup>d</sup>
SFAs% <sup>a</sup>	33 (32 ,34)	34 (33,35)	33 (32.7,35.0)	<0.001
PUFAs% <sup>a</sup>	44 (42,45)	41 (38,43)	43 (41.0,45.5)	<0.001
MUFAs % <sup>a</sup>	22 (21, 24)	24 (22,26)	22 (21, 24)	<0.001

Values are presented as N (proportion), and \*P values represent chi-squared tests <sup>a</sup>Values are median (interquartile range)

<sup>b</sup> P values represent Wilcoxon rank-sum test <sup>c</sup> Values are mean (standard deviation) <sup>d</sup> P values represent analysis of variance

<sup>^</sup>Defined by NICE (>35 Units per week in women or >50 in men).

In both groups, participants with steatosis had a higher mean BMI than controls. In the UK Biobank sample, the correlation between NMR plasma fatty acids % and dietary fatty acids, measured as % of energy intake, using at least 2WebQ, presented a low correlation. For SFAs it was 0.1274, for MUFAs 0.0445, and for PUFAs 0.290. Omega 3 and omega 6 PUFAs presented slightly stronger correlations, of 0.1879 and 0.1468 respectively.

Table 5.3. OCDEM participant characteristics by steatosis.

	No steatosis (N=82)	Steatosis (N=43)	Overall (N=125)	p-value
Liver fat content <sup>a</sup>	1.1 (0.7,2.2)	12.7 (6.8,20.7)	2.3 (0.9,6.8)	<0.001 <sup>b</sup>
Age (years) <sup>c</sup>	47.0 (8.7)	49.3 (9.4)	47.8 (9.0)	0.17 <sup>d</sup>
Men	24 (29.3%)	24 (55.8%)	48 (38.4%)	0.004
BMI (kg/m <sup>2</sup> ) <sup>a</sup>	25.4 (23.3,27.8)	31.3 (28.4,35.6)	27.2 (24.6,30.2)	<0.001 <sup>b</sup>
Fat mass %	33.1 (7.5)	39.1 (7.6)	35.1 (8.0)	<0.001
<b>NMR plasma fatty acids</b>				
Total TG (mmol/l)	1.0 (0.7,1.3)	1.3 (1.0,2.0)	1.1 (0.8,1.6)	0.003
VLDL (mmol/l)	0.6 (0.3)	0.7 (0.3)	0.7 (0.3)	0.039
HDL (mmol/l)	1.4 (0.3)	1.2 (0.3)	1.3 (0.3)	0.002
LDL (mmol/l)	1.8 (0.4)	1.9 (0.4)	1.8 (0.4)	0.49
Non-HDL cholesterol (mmol/l)	3.3 (0.7)	3.4 (0.7)	3.3 (0.7)	0.37
LDL/HDL ratio	1.3 (1.1, 1.8)	1.7 (1.2, 2.0)	1.5 (1.11, 1.8)	0.01
HOMA-IR	2.1 (1.5,2.8)	4.9 (3.2,7.0)	2.5 (1.7,3.6)	<0.001
Insulin U/L	10.2 (6.8,12.1)	20.1 (12.6,27.7)	11.0 (8.0,15.1)	<0.001
Glucose mmol/L	5.0 (4.8,5.3)	5.5 (5.3,6.0)	5.2 (4.9,5.6)	<0.001
Plasma SFAs% <sup>a</sup>	34.1 (33.0,34.8)	34.8 (33.4,35.8)	34.3 (33.0,35.2)	0.038
Plasma PUFAs% <sup>a</sup>	42.5 (40.7,44.3)	39.4 (36.9,41.9)	41.8 (39.4,43.5)	<0.001
Plasma MUFAs% <sup>a</sup>	23.8 (22.3,25.3)	26.4 (24.4,28.8)	24.3 (22.6,26.4)	<0.001
<b>VLDL-TG FA composition</b>				
SFAs% <sup>a</sup>	30.7 (27.9,33.4)	35.1 (31.1,37.8)	31.4 (29.0,34.9)	<0.001
PUFAs% <sup>a</sup>	18.3 (16.3,21.9)	15.6 (12.7,18.4)	17.7 (14.5,20.3)	0.001
MUFAs% <sup>a</sup>	48.6 (46.4,52.2)	48.7 (45.9,50.6)	48.6 (45.9,51.2)	0.53
Myristic (mol%)	1.8 (0.8)	2.3 (1.1)	2.0 (0.9)	0.011
Palmitic (mol%) <sup>a</sup>	26.1 (23.6,28.0)	28.9 (26.2,31.2)	26.7 (24.7,29.1)	<0.001
Palmitoleic (mol%)	4.2 (1.4)	4.7 (1.7)	4.4 (1.5)	0.12
Stearic (mol%) <sup>a</sup>	2.9 (2.5,3.3)	3.3 (2.7,3.7)	3.0 (2.5,3.5)	0.002
Oleic (mol%)	43.8 (4.8)	41.1 (3.8)	42.9 (4.6)	0.002
Linoleic (mol%)	16.1 (4.3)	13.4 (3.1)	15.2 (4.1)	<0.001
Alpha linoleic (mol%)	1.1 (0.7)	0.9 (0.5)	1.0 (0.7)	0.244

Values are presented as N (proportion), and \*P values represent chi-squared tests. <sup>a</sup> Values are median (interquartile range)

<sup>b</sup> P values represent Wilcoxon rank-sum test <sup>c</sup>Values are mean (standard deviation) <sup>d</sup> P values represent analysis of variance. VLDL: very-low density lipoprotein, LDL: low density lipoprotein, HDL: high density lipoprotein, TG: triglycerides. HOMA-IR: Homeostatic Model Assessment of Insulin Resistance

### 5.3.2 NMR plasma fatty acids and liver fat in UK Biobank

The sequential adjustment for confounders are presented in Figure 5.4. In minimally adjusted models by sex and age, there were positive associations between plasma SFAs% and MUFAs% and liver fat (+9.4 % [8.5, 10.3] difference, and +9.3 [8.7, 10.0] % difference in geometric mean) per 1% increase in plasma FA percentage, respectively), while there was an inverse association between plasma PUFAs% and liver fat (-6.9 % [-7.3, -6.4%] per 1% increase in plasma FA percentage). These were attenuated by the addition of confounders, but remained significant after fully controlling for all covariates.



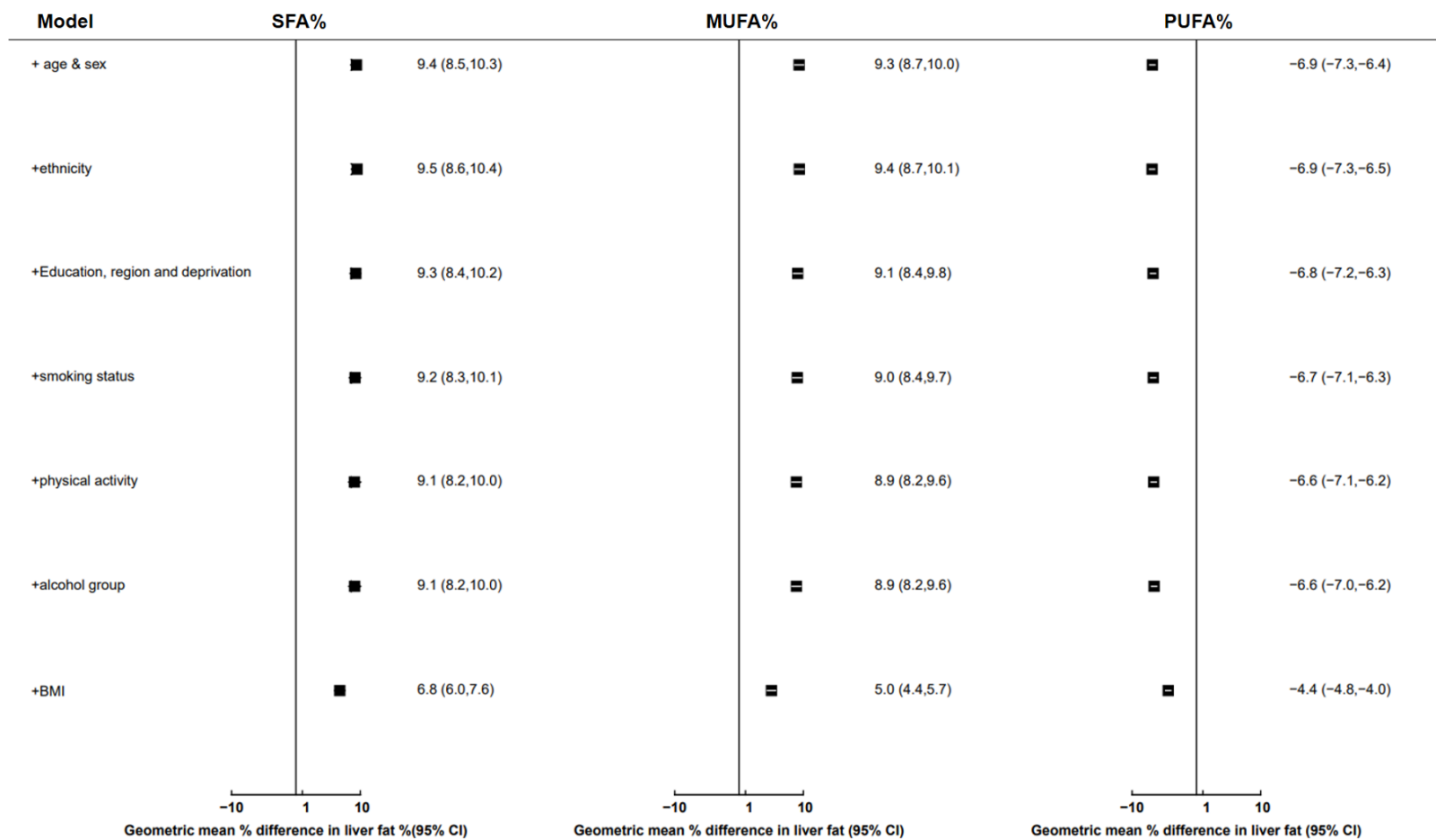


Figure 5.4. Geometric mean % difference in liver fat, per 1% increase in plasma FAs%: sequential adjustment for confounders (N=7,206), UK Biobank.

In fully adjusted models controlling for sex, age, ethnicity, deprivation, education, smoking, physical activity, alcohol risk and BMI, a positive association with liver fat was found with plasma SFAs% and MUFAs%. Per 1% increase in SFAs% there was a +6.8 geometric mean % difference (6.0, 7.6), and per 1% increase in MUFAs% there was a +5.0 geometric mean % difference (4.4, 5.7) in liver fat. In contrast, an inverse association was observed per 1% increase in PUFAs% (-4.4% [-4.8, -4.0]). BMI was the confounder with the highest % change in  $\chi^2$ , highlighting its relevance as a confounder in these associations.

There was no evidence of interaction between BMI with NMR plasma fatty acids in their association with liver fat (SFAs  $p$  heterogeneity [het]. = 0.71). PUFAs ( $p$  het. =0.36), MUFAs ( $p$  het. =0.22). There were no significant interactions by sex either with SFAs% ( $p$  het. =0.72), MUFAs% ( $p$  het. =0.57), or PUFAs% ( $p$  het. =0.94).

### 5.3.3 Sensitivity analysis in fasting group

Restricting the analyses to those who reported fasting for 8 hours or more limited the sample to 249 in total, which represented 3.5% of the main sample, and their mean plasma fatty acids proportions were similar. Fully adjusted models provided results very similar to those in the main sample, but with stronger positive associations with liver fat for SFAs% (9.6 % [4.9, 14.5%]) and an inverse for PUFAs% (-5.2% [-7.6, -2.8%] per 1% increase in plasma FAs). The positive association between plasma MUFAs% and liver fat was slightly attenuated (4.2%, [0.4, 8.1%]), (Figure 5.5).

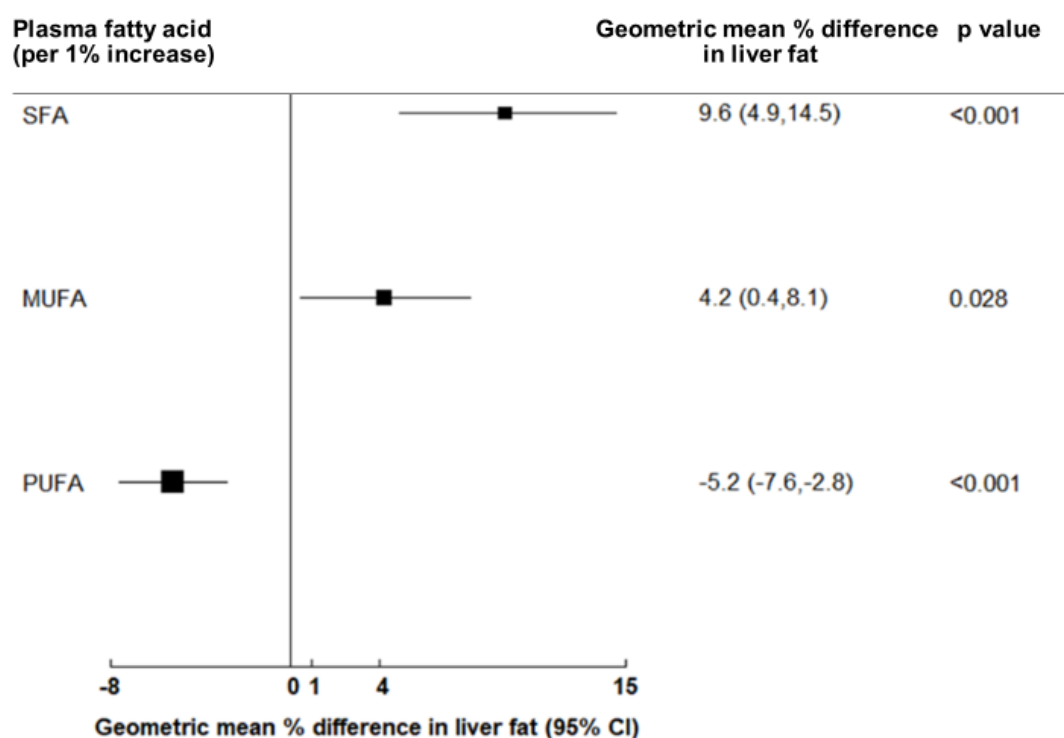


Figure 5.5 Geometric mean % difference in liver fat per 1% increase in plasma fatty acids in the UK Biobank. Sensitivity analysis restricted to participants fasting (N=249). Models were adjusted for sex, age, ethnicity, deprivation, education, smoking, physical activity, alcohol risk & BMI.

### 5.3.4 NMR plasma fatty acids and liver fat in OCDEM

The associations between NMR plasma fatty acids and liver fat in OCDEM are presented in Figure 5.6. In minimally adjusted models controlling for sex, there were significant positive associations between plasma SFAs and liver fat (19.5% [2.5, 39.3]), and between plasma MUFAs (18.3% [10.8, 26.3]) and liver fat. Conversely, there was an inverse association between plasma PUFAs and liver fat (-13.6%, [-18.4, -8.6]). Once age was added to the model, the associations in SFAs% became non-significant.

In fully adjusted models controlling for sex, age and BMI, the associations between plasma SFAs and liver fat, were non-significant. However, the associations between plasma MUFAs and PUFAs remained significant, with a difference of +7.5% (1.6, 13.7) and -4.9 (-9.1, -0.4), respectively. The addition of BMI substantially changed the  $\chi^2$  in all of the fatty acids (between 60-90%, data not shown).

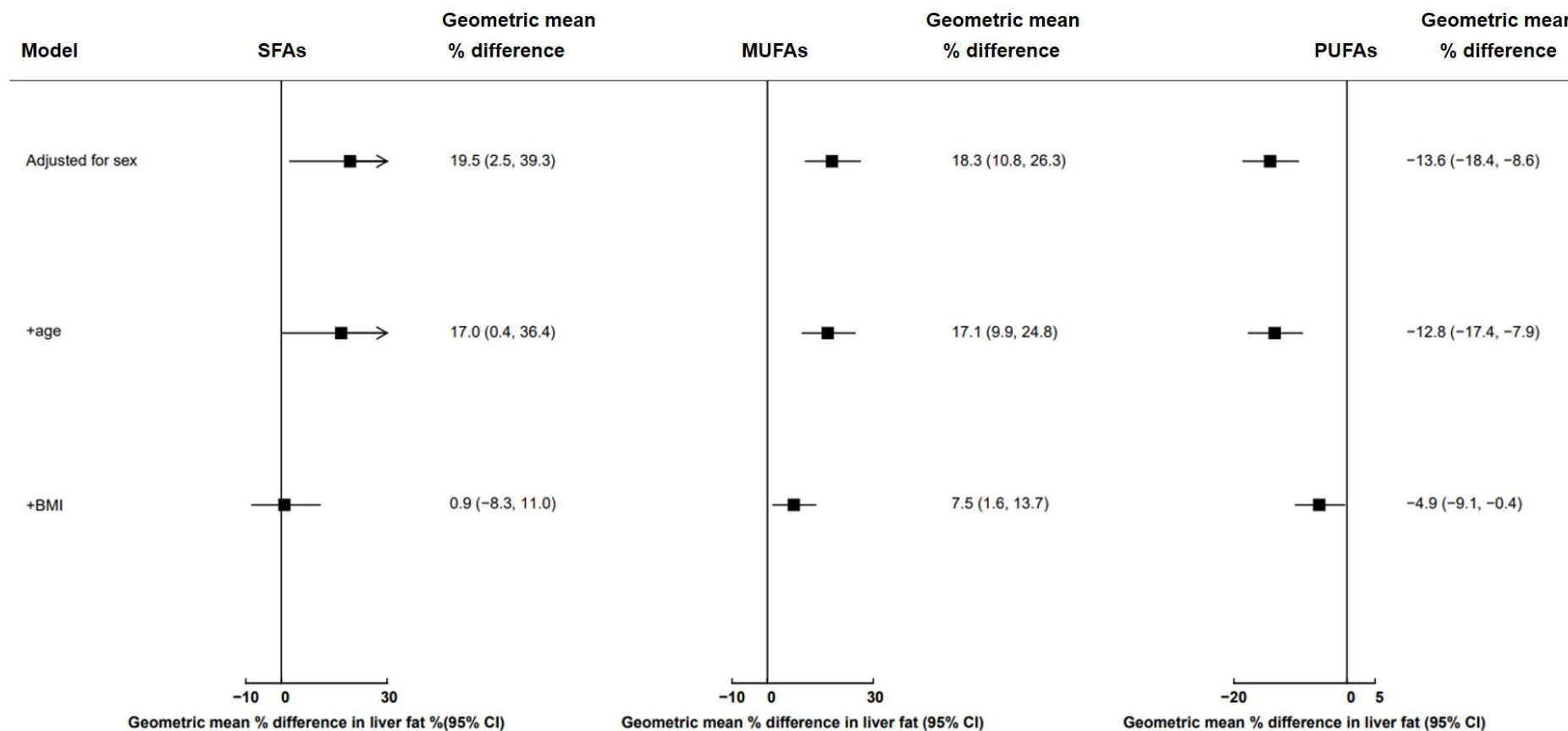


Figure 5.6 Geometric mean % difference in liver fat (95%CI) per 1% increase in NMR plasma fatty acids in OCDEM (N=125).

### 5.3.5 Comparison between OCDEM and UK Biobank

A comparison between UK Biobank and OCDEM analyses can be observed in Figure 5.7, presenting results from fully adjusted models. In fully adjusted models, the direction and significance of the associations was similar to the main sample per 1% increase in MUFAs% and PUFAs%, but there was no similarity in SFAs%.

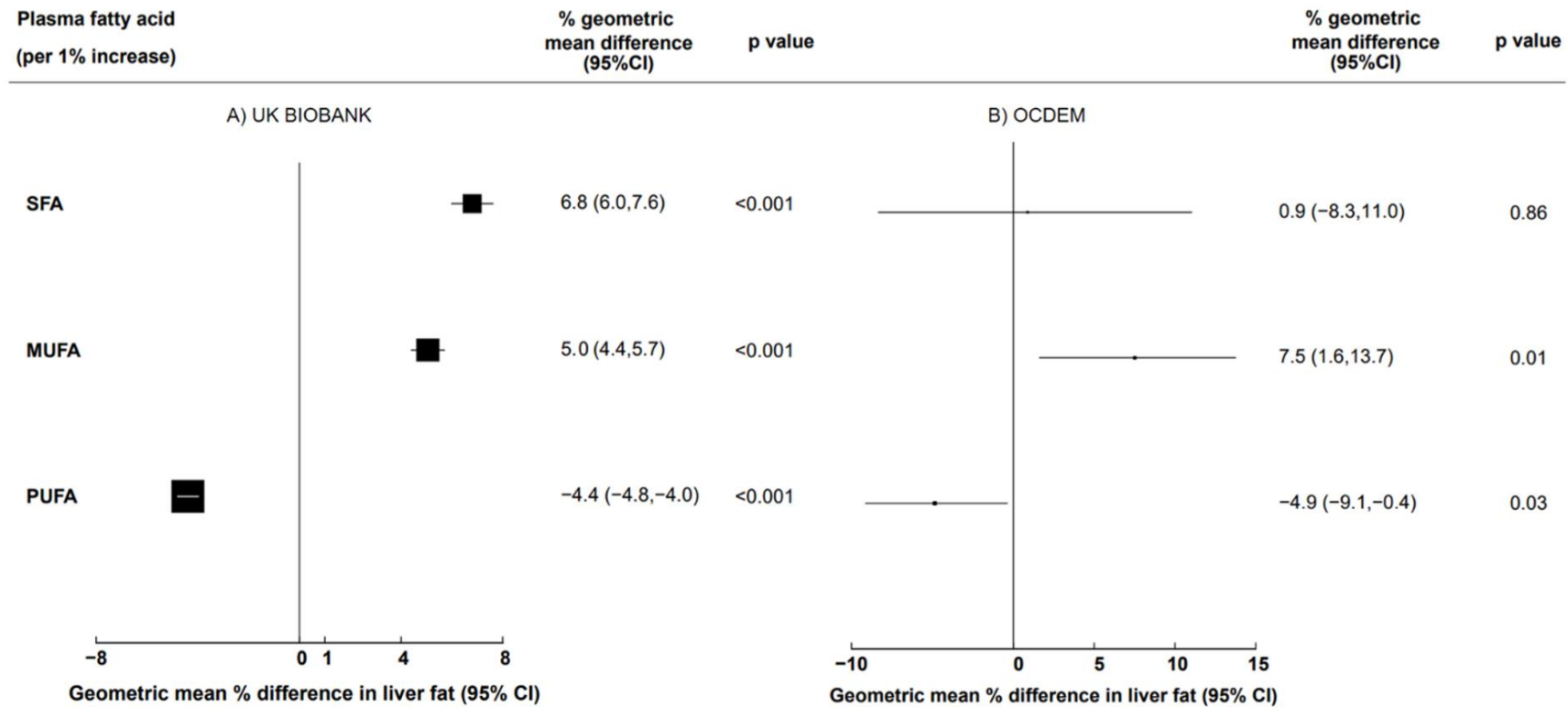


Figure 5.7. Geometric mean % difference in liver fat (95%CI) per 1% increase in NMR plasma fatty acids in UK Biobank (A) and OCDEM (B). Linear regression models were used to calculate the geometric mean % difference in liver fat, by 1% increase in plasma fatty acids. A) Models were adjusted for sex, age, ethnicity, deprivation, education, smoking, physical activity, alcohol risk & BMI (n=7,206) B) models were adjusted for age, sex and BMI (n=125)

### 5.3.6 Associations between liver steatosis and VLDL-TG fatty acid composition

In unadjusted comparisons of the VLDL-TG fatty acid composition of steatosis vs. controls using Wilcoxon rank sum tests, there was a significant median difference in SFAs% (+4.22% [95% CI 2.49-5.95],) and PUFAs% (-2.61% [-4.00, -1.22]), (Figure 5.8). There were no significant differences in MUFAs% (-0.01 [-1.51, 1.31]).

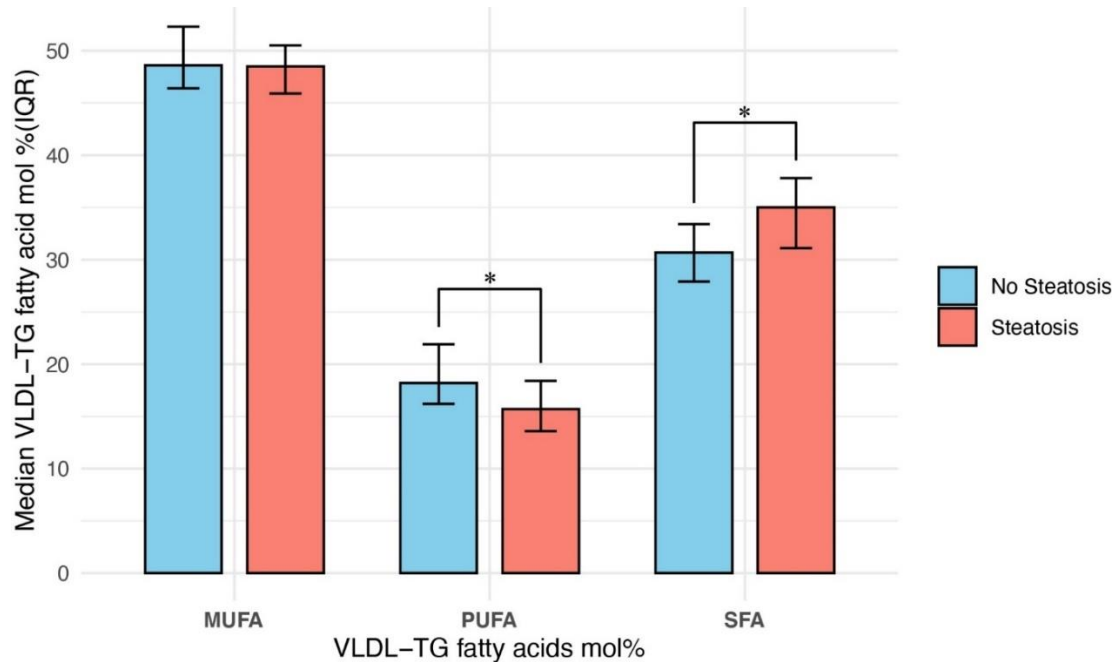


Figure 5.8 Comparison of mean VLDL-TG fatty acid composition in OCDEM participants with and without steatosis, unadjusted N=125. Medians of VLDL-TG fatty acids compared using Wilcoxon-rank tests \*p<0.05.

Participants with steatosis had higher odds of high VLDL-TG SFAs% than controls, when adjusting for sex and age (OR of high VLDL-TG SFAs: 3.85 [1.69, 8.75]) (Figure 5.9). After adjusting for BMI, the associations were not significant. For VLDL-PUFAs% and VLDL-MUFAs%, the associations with steatosis were non-significant in any of the models. Adjusting for BMI explained more than 90% of the associations between steatosis and high VLDL-SFAs and high VLDL-PUFAs, and none of the associations remained significant after its inclusion.

There was no evidence of an interaction between sex and steatosis in the association with high VLDL-TG MUFAs, PUFAs, or SFAs (p<sub>het</sub>=0.09, 0.94 and 0.42 respectively).

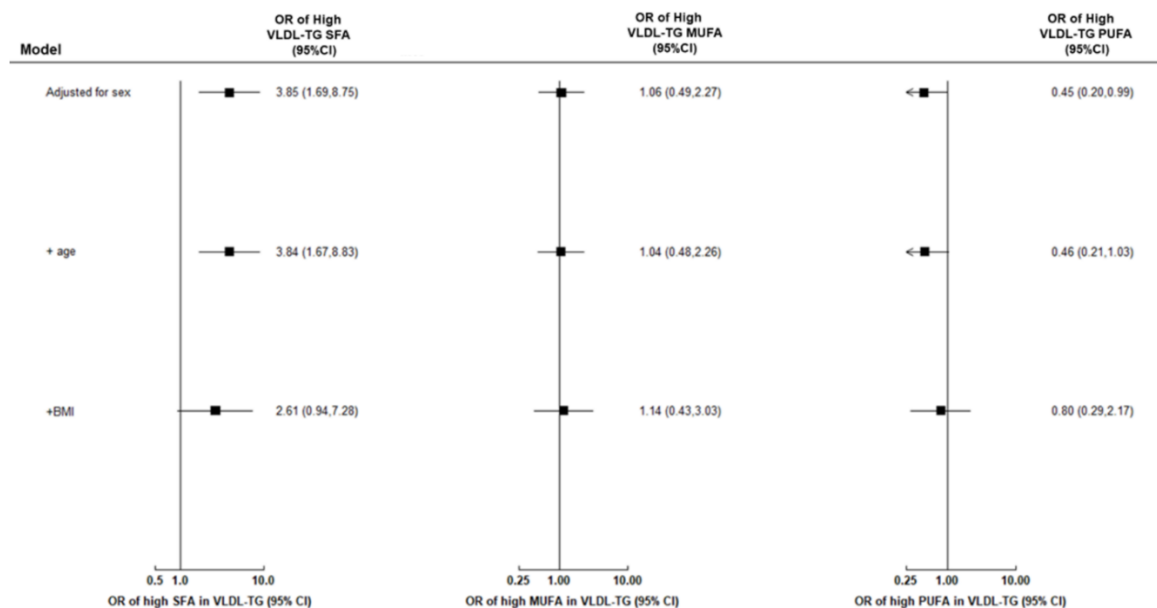


Figure 5.9 OR of high FAs VLDL-TG in participants with steatosis, with sequential adjustment for confounders. N=125. VLDL-TG: very-low density lipoprotein triglyceride

### 5.3.7 Associations between liver steatosis and VLDL-TG individual fatty acid composition in OCDEM

The associations observed in steatosis and VLDL-TG individual fatty acids were not substantially different from the results expected due to the structural group they belong to (SFAs, PUFAs or MUFAs). No significant associations were observed in minimally or fully adjusted models (Table 5.4).

Table 5.4 OR of high VLDL-TG individual fatty acids in participants with steatosis in OCDEM (N=125)

	Myristic	Palmitic	Palmitoleic	Stearic	Oleic	Linoleic	Alpha-Linoleic
Adjusted by sex	2.1 (1.0,4.7)	3.0 (1.3,6.6)	2.1 (0.9,4.8)	2.5 (1.1,5.6)	0.2 (0.0,0.5)	0.2 (0.1,0.5)	0.4 (0.2,1.0)
+age	2.1 (0.9,4.7)	2.9 (1.3,6.5)	2.0 (0.9,4.7)	2.6 (1.1,5.9)	0.2 (0.0,0.5)	0.2 (0.1,0.5)	0.4 (0.2,1.0)
+BMI	1.7 (0.6,4.8)	1.9 (0.7,5.4)	1.3 (0.5,3.8)	2.6 (0.9,7.4)	0.4 (0.0,1.2)	0.3 (0.1,1.0)	0.4 (0.2,1.2)



### 5.3.8 Associations between steatosis and markers of cardiovascular risk in OCDEM

In fully adjusted models controlling for age, sex and BMI, no associations between steatosis and markers of cardiovascular risk were significant (Figure 5.10). There was a significant interaction between total triglycerides and sex, therefore results are presented for men and women separately.

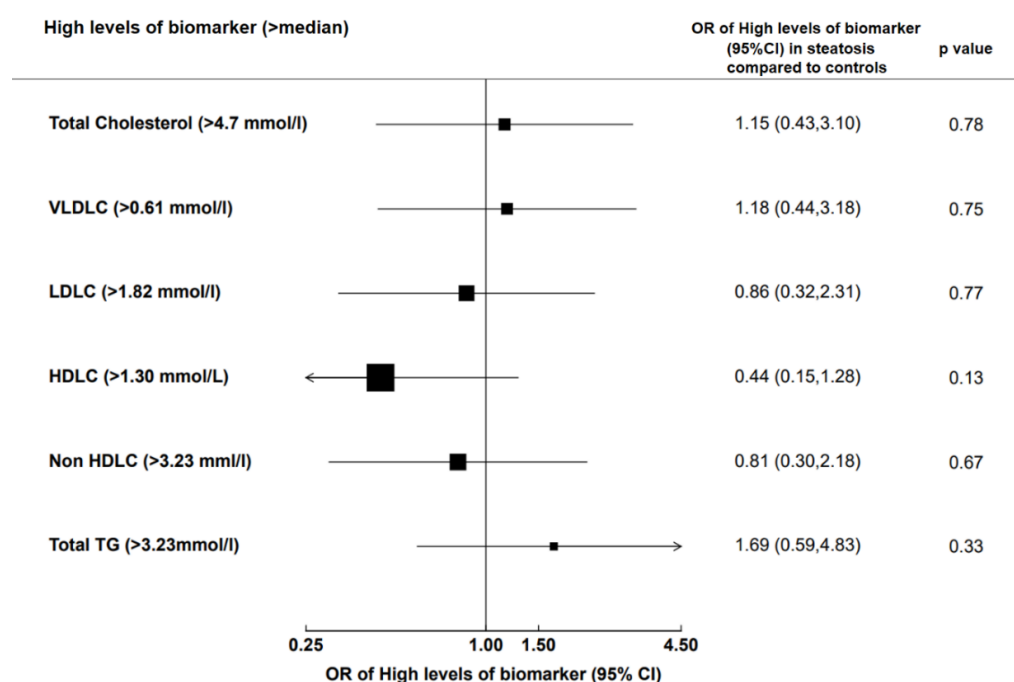


Figure 5.10 Odds ratio of high levels of biomarkers in participants with steatosis in comparison to controls (N=125). Results obtained from multivariate logistic regression models, adjusted by age, sex, and body mass index. VLDLC: very-low density lipoprotein cholesterol, LDLC: low density lipoprotein cholesterol, HDLC: high density lipoprotein cholesterol, TG: triglycerides.

Analysis showed that the positive association was observed mostly in men, with higher odds of high TG in comparison to controls, with no association in women (Table 5.5).

Table 5.5 Odds ratio (OR) of high TG in participants with steatosis, by sex, adjusted for age, and BMI (N=125)

	OR of Triglycerides>3.23mmol/l (95%CI)	
Overall	1.69 (0.59,4.83)	
Men	5.14 (1.45,18.15)	
Women	0.66 (0.18,2.51)	
		p heterogeneity= 0.023

Odds ratios of high levels of blood biomarkers, in minimally adjusted models and with sequential adjustment for confounders can be observed in Table 5.6. Only high levels of HDL (which is a protective biomarker) were inversely associated with steatosis in models adjusted by sex (OR: 0.32 [0.14-0.73]). However, after the addition of confounders, these associations became non-significant. High changes in  $\chi^2$  were observed after comparing models with BMI, which suggests that BMI may explain a large proportion of the associations between steatosis and levels of biomarkers.

Table 5.6 OR of high levels of markers of cardiovascular risk in participants with steatosis vs. controls with sequential adjustment for confounders (N=125).

	<b>Adjusted for sex</b>	<b>+ age</b>	<b>+BMI</b>
<b>Total cholesterol</b>	0.84 (0.39,1.82)	0.89 (0.41,1.94)	1.15 (0.43,3.10)
<b>VLDL</b>	1.64 (0.76,3.57)	1.66 (0.75,3.65)	1.18 (0.44,3.18)
<b>LDL</b>	0.81 (0.38,1.76)	0.81 (0.37,1.76)	0.86 (0.32,2.31)
<b>HDL</b>	0.32 (0.14,0.73)	0.33 (0.14,0.75)	0.44 (0.15,1.28)
<b>Non-HDL</b>	0.88 (0.41,1.91)	0.88 (0.40,1.94)	0.81 (0.30,2.18)
<b>Total triglycerides</b>	1.96 (0.88,4.40)	1.84 (0.81,4.17)	1.69 (0.59,4.83)

## 5.4 Discussion

### 5.4.1 Main findings

This observational study aimed to explore the associations between plasma fatty acids and liver fat, as well as the associations between steatosis and VLDL-TG fatty acids composition. In addition, it explored the associations between steatosis and plasma markers of CVD. In this study of two groups of adult participants from the UK, it was observed that after adjusting for confounders, a positive association was observed between plasma MUFAs% and liver fat, in both the UK Biobank and OCDEM samples. In addition, an inverse association was observed between plasma PUFAs% and liver fat content in both groups. In contrast, higher plasma SFAs% was positively associated with liver fat in the UK Biobank group but not in the OCDEM group.

In the OCDEM study group, after adjusting for BMI, no significant associations were observed between steatosis and VLDL-TG fatty acid composition. There were no significant associations between steatosis and individual fatty acids within VLDL-TG. When looking at the high changes in  $\chi^2\%$ , it can be observed that a high proportion of the association between steatosis and VLDL-TG is actually driven by BMI. Finally, steatosis was associated with higher odds of high levels of plasma TG than those without steatosis, only in men. There were no significant associations with other cardiovascular risk factors in fully adjusted models.

#### 5.4.2 Comparison with previous studies

The results presented in this chapter support previous evidence demonstrating the positive associations between plasma SFAs% and MUFAs% with liver fat, as well as the negative association between PUFAs% and liver fat (133). In a study by Gnautic et al, using linear regressions, it was observed that higher levels of liver fat were associated with lower circulating levels of PUFAs, and higher levels of SFAs and MUFAs (133). This thesis adjusted for similar confounders than that study, but analysed a larger sample of participants due to MRI availability, and provided further adjustment for physical activity, and for BMI, which was included as a confounder because it did not meet the statistical significance as an interaction. Another study published by Friden et al. in the UK Biobank also found associations in the same directions for plasma SFAs%, PUFAs% and MUFAs% (32). However, this thesis provided a more comprehensive of exclusion of participants with excessive alcohol intake and further adjustment in the model, while Friden et al. did not perform any exclusions based on alcohol intake (32). This is particularly relevant because, as previously mentioned, excessive alcohol intake can result in participants being classified under alcoholic steatotic liver disease, a distinct condition with a different pathophysiology to MASLD. Failing to exclude such individuals could lead to analysing a study sample that includes cases of alcoholic steatotic liver disease, despite my focus being solely on understanding the condition in the context of moderate

alcohol consumption. Therefore, it was essential in my study to exclude individuals in this category to ensure that the observed associations were not skewed by their inclusion.

The lack of significant associations between VLDL-TG SFAs and VLDL-TG PUFAs and steatosis after adjustment for BMI highlights the importance of BMI as a confounder in these associations, since BMI is strongly associated with both circulating lipid concentrations and liver fat content. This suggests that the associations between steatosis and VLDL-TG fatty acid composition are mostly confounded by BMI.

The positive associations with high plasma TG in participants with steatosis was only observed in men while there was no significant association for women. Higher levels of plasma TG measured via NMR were also associated with higher levels of liver fat in the study by Gnaniuc et al in the UK, which collected the data in a 5 year separation (133). This supports the evidence by including data in which exposure and outcome were measured at the same time. The different associations between steatosis and high plasma TG by sex could be explained by difference in triglycerides metabolism between men and women (142).

When comparing the two groups, the UK Biobank provides insights from a large population cohort, while the OCDEM sample provides more detailed measurement of outcomes, and more precise identification of exclusions. Furthermore, at OCDEM, blood and imaging measurements were taken in close temporal proximity, which was not possible in the UK Biobank sample. The mean plasma fatty acids values in both populations were similar, and they were both processed by the same provider, which allowed comparing results. OCDEM and UK Biobank protocols ensured a correct collection, and storage of blood samples, following protocols before they were delivered to the NMR assessment provider (74).

This is one of the few studies to look at the relationship between steatosis and VLDL-TG composition, and to explore associations within individual fatty acids. While a study in the

United States found that VLDL-TG palmitoleic acid is strongly associated with high liver fat, this was only observed in univariate associations in this sample, as after addition of confounders associations became non-significant because BMI was confounding the associations (137).

### 5.4.3 Strengths

The UK Biobank sample provides the largest dataset with both plasma fatty acids measurements and liver MRIs available worldwide (78). The use of the liver MRI to assess liver fat is convenient as it is a non-invasive method, with the most accurate quantification of liver fat, and with greater sensitivity than ultrasound (80, 140).

The detailed information collected from UK Biobank participants, allowed careful adjustment for confounders and exclusions. In a sensitivity analysis that studied participants who were fasting, their results were similar than those from the main sample. This could suggest that fasting does not significantly influence the associations in this population, reinforcing the robustness of the results obtained from the main sample, that is then not heavily confounded or dependent on fasting status.

### 5.4.4 Limitations

#### 5.4.4.1 Samples and design

Firstly, this study has the limitations of its cross-sectional design, which cannot establish causality. In addition, in the UK Biobank sample, the time between MRIs and collection of blood samples was 9 years on average, which would have been enough time for plasma fatty acid composition to change if a dietary change had been made. While participants from the UK Biobank with incident chronic disease were excluded during the period between plasma fatty acids and MRIS were taken, as an attempt to prevent conditions influencing measurements, changes in conditions such as hypercholesterolemia may have been missed.

This could be because they may not have needed a hospital admission, and would have been diagnosed in primary healthcare setting. This could have resulted in missing the identification of a case of MASLD.

Misclassification was not an issue for participants at OCDEM, as they were screened and any chronic condition made them ineligible at the time of the study, which was much more precise than the UK Biobank. The sample size available from the studies performed in OCDEM was relatively small (n=125 participants) which may have been a factor that did not enable enough power to assess some of the associations. This is reflected in the wide confidence intervals of many of the estimates, and in theory a larger sample size would present with less variance. More than 95% of the population in the UK Biobank was of White ethnicity, and there was no data about ethnicity available in OCDEM. While it could be representative of a subset UK population, results may only be generalisable to other similar European populations.

#### 5.4.4.2 Plasma fatty acid measurements

As described previously, some authors observed that plasma fatty acids are not good biomarkers of dietary intake, and their use has been criticised due to their low correlation with diet in many studies (143). However this depends mainly on the fatty acid measured: it is also proposed that as long they are correctly measured with the appropriate collection, storage, and quality control procedures, an objective measurement (132, 144). In addition, it is important to highlight that in this sample, there were low correlations between dietary fatty acids measured by WebQ and plasma fatty acids.

### 5.5 Conclusions

This study suggests that different plasma fatty acids relate with liver fat in different directions, and that differences in VLDL-TG fatty acid composition in steatotic livers compared with non-steatotic may be confounded by sex and BMI. In particular, it would be helpful to explore

these associations in larger samples, while still acknowledging the complexity of obtaining VLDL-TG measurements, which are more difficult in larger populations. Finally, it would be informative to assess these associations in different populations that may have been underrepresented in these samples. Still, the findings in this chapter suggest that the associations between plasma fatty acids and liver fat do not vary according to fasting status, and that they remain significant after adjusting for confounders. This would be particularly relevant for future studies, as reducing the need of fasting may increase feasibility of studies and volunteer participation.

## 6 Discussion

### 6.1 Summary of findings

#### *Chapter 3 Dietary fatty acids and liver fat*

In an observational analysis of 9,268 participants from the UK Biobank, higher levels of self-reported intakes of dietary SFAs, and lower intakes of PUFAs were associated with higher liver fat measured by MRI, with consistent findings in sensitivity analyses. In subgroup analyses by sex, it was observed that the associations between SFAs intake and liver fat were stronger in men. For SFAs, stronger associations with liver fat content were observed in those participants with higher levels of visceral adipose tissue.

#### *Chapter 4 Dietary fatty acids and MASLD*

In observational analyses of 12,301 participants from the UK Biobank, higher intakes of both SFAs and MUFAs were associated with higher odds of MASLD, while a higher intake of PUFAs was associated with lower odds of MASLD. When investigating the relationship between individual fatty acids and risk of MASLD, individual fatty acids demonstrated associations which were in alignment with the classification group in which they belonged.

#### *Chapter 5 Plasma fatty acids and liver fat. Steatosis and VLDL-TG fatty acid composition.*

In observational analyses of 7,201 participants from the UK Biobank, an increase in plasma SFAs measured by NMR was associated with higher levels in liver fat. No significant results were observed for higher plasma PUFAs or MUFAs after fully adjusting for confounders. Associations observed remained the same when analyses were restricted to fasting participants.



## 6.2 Findings contribution to current literature

As described in Chapter 1, the majority of the observational studies published so far on dietary fats and liver fat content or MASLD, have reported no associations between different types of fatty acids and liver fat, or had insufficient dietary or outcome measurements. This thesis adds novel knowledge to the field as described below.

### 6.2.1 Positive associations between dietary SFAs and plasma SFAs with liver outcomes across the thesis.

This thesis reinforces the findings of positive associations between dietary SFAs and liver fat content that were recently described in two observational studies carried out in the UK Biobank by Chen et al. and Friden et al., who reported positive associations between SFAs and liver fat using at least one WebQ (31, 32). However, the analyses in this thesis utilised at least 2 WebQs. Using only one 24-h dietary assessment is considered a poor estimate of true long term dietary intake, while using at least two reduces random measurement error due to within-person variability, and overall, provides a better estimate of usual dietary intake (31, 32). In addition, the analyses performed in this thesis outlined a set of exclusions based on excessive alcohol drinking, a variable that was adjusted for by Friden et al. but not used for exclusions (32). In this thesis, the influence of alcohol intake is carefully controlled for, by both exclusion of excessive alcohol drinkers and addition of alcohol intake as a confounder, reducing the likelihood of misclassifying participants as non-MASLD, MetALD, or ALD cases. In addition, this study differs in its adjustment approach, as results were presented both before and after including BMI as a confounder (32, 33, 133). This was key as BMI had a substantial impact in the associations. There was also a stronger association between SFAs and liver fat in participants with high VAT compared to those with low VAT, which had not been previously studied by other researchers as an interaction. This interaction was only present for SFAs and no other fatty acids, and we could hypothesise that high levels of VAT would contribute to a

pro-inflammatory state (through cytokines), adding to the inflammation and insulin-resistance associated with ceramides from SFAs, promoting more liver fat accumulation. Overall, these findings align with the recommendations made at the last SACN report, reinforcing the existing recommendations that dietary SFAs should be consumed in moderation <10% of intake (145). Following the findings of VAT, it may be particularly relevant to advise patients with high levels of VAT to reduce their SFAs intake.

### 6.2.2 Dietary and plasma PUFAs associations with liver fat outcomes across the analyses

This thesis found inverse associations between dietary PUFAs and liver fat. This could be explained by the use of at least two WebQs by this thesis, which are better at capturing nutrients that may not be eaten very frequently; particularly relevant in the case of PUFAs, and the sensitivity analyses. Food products rich in PUFAs are consumed approximately only once a week in the UK, so it would be less likely to capture this intake using only one measurement (122). Looking at omega 3 and omega 6 fatty acids did not yield any significant results; these did not show associations different from overall dietary PUFAs in this thesis analyses. These findings are similar to research carried out in the UK Biobank by Tian et al (33), in which a dietary pattern high in PUFAs was negatively associated with liver fat content, without significant associations for other patterns. In addition, trials have observed a negative association between dietary PUFAs and liver fat content in controlled environments in which excessive alcohol consumption would not be present (44, 47). However, dietary PUFAs showed no significant associations with MASLD, and in Chapter 3, the negative associations observed with liver fat became non-significant when moderate alcohol drinkers were included in the sample. More research in moderate alcohol drinkers may be needed to explore this question in depth.

### 6.2.3 Individual dietary fatty acids across analyses

The current analyses contributed to the literature as the first carried out on individual dietary fatty acids and liver fat outcomes using the UK Biobank individual fatty acids dataset, the largest dataset with individual fatty acids dietary intake. There were many advantages of this approach, as it provides data for 21 individual fatty acids, and allowed to observe whether any individual fatty acids had differing associations from the overall group they belonged to. In addition, I also studied omega 3 and omega6 PUFAs groups. While the individual fatty acids resource provided very valuable and detailed data about the individual fatty acids intake by using the data from the Oxford WebQ, only six individual fatty acids represented more than 1% of total energy intake, therefore only those results were considered relevant in this population. Nevertheless, findings for other individual fatty acids were also reported, as the dietary composition of individual fatty acids in foods could shift in the near future, and these results could serve as a baseline for future research.

From the SFAs group, palmitic, stearic and myristic all presented positive significant associations with liver fat, while from the PUFAs group, none of the FAs, including omega-3 and omega-6 groups, demonstrated significant associations after adjusting for multiple testing. For the MUFA group, only oleic acid represented more than 1% of energy intake and they were associated with higher odds of MASLD. All FAs demonstrated associations which were in alignment with the classification group in which they belonged.

### 6.2.4 Steatosis, VLDL-TG fatty acid composition

As described in Chapter 2, the methods required to isolate VLDL-TG are complex, and only carried out by a small number of research groups around the world due to its time and resources demand. As such there is limited available data looking at steatosis and different VLDL-TG fatty acid composition patterns. It has previously been reported that steatosis is associated with increased VLDL-TG secretion(146). It was therefore hypothesised that steatotic livers may

export VLDL-TG with higher composition of SFAs and lower composition of PUFAs, providing information of the liver fatty acid content (26, 96, 146-148). Associations in those directions were observed, but these crude associations did not remain significant after the addition of BMI, suggesting that BMI may be the main driver of the associations between steatosis and VLDL-TG fatty acid composition.

#### 6.2.5 The role of menopause status

One of the aims of this thesis was to explore the role of menopause status in the associations between dietary fatty acids and liver fat outcomes, given that postmenopausal women would represent a group metabolically different from premenopausal women, due to changes in oestrogen, and previous studies have observed higher prevalence of MASLD in postmenopausal women compared to premenopausal group (149). The lack of any significantly different results across menopausal status groups could be explained by both the sample size and the way that menopause was measured. While the UK Biobank age range at recruitment ranged between 40 and 70 years, considering MRIs were taken >10 years later makes the mean age of women who attended the imaging visit 63 years old. It is important to note that the menopause categories used in this thesis were not based on hormones levels or diagnoses, but on self-report and time at MRI, which introduces measurement error. To define menopause in this thesis, I considered women who reported going through menopause at the baseline visit, women over 55 years old at the time of MRI, and women who underwent gynaecological procedures that result in the onset of menopause.

#### 6.2.6 Sexual dimorphism

This thesis aimed to analyse the role of sex in the associations between dietary fatty acids and liver fat outcomes, and showed results from a sex specific subgroup analysis. It was observed that the positive associations between dietary SFAs and liver fat content were stronger in men than in women, which was not observed in any of the other analyses carried out with other fatty

acids in this thesis. It has been described in a previous study that men present lower dietary fatty acids oxidation, and increased de novo lipogenesis compared to women, which may lead to fatty acids esterification and storage in the liver, and more liver fat accumulation (150). On one hand, these findings were expected, as they align with established sex based metabolic differences. Oestrogen may play a significant role in lipid metabolism, and men show lower rates of fatty acid oxidation (151). This may predispose men to metabolise SFAs towards storage pathways that accumulate fat in the liver more easily (150, 151). In addition, after fully adjusting for confounders, no associations between steatosis and markers of cardiovascular risk were significant in Chapter 5, but an interaction by sex was observed in the associations between steatosis and plasma triglycerides, with a positive association only for men, highlighting a potential difference in triglyceride metabolism. In the context of this thesis, this could be related to an extent to the results from Chapter 3, in which men have stronger associations between SFAs and liver fat, than women. As liver fat is mainly TG, this could be driven by the same difference in metabolism, potentially due to the lack of the protective effect of oestrogen.

However, considering that most women in this sample belong to the postmenopausal group, the impact of oestrogen may not be present, and this could be an incidental finding, especially since these associations were not observed in other analyses. Still, it has already been pointed out in this chapter that the measurement and categorisation of menopause in this study are imperfect. These limitations highlight the importance of conducting future studies in populations with a wider age range to assess these associations in big datasets with sufficient statistical power for both pre-and post-menopausal women.

#### 6.2.7 The role of BMI and adiposity across analyses

BMI played an important role on most of the associations studied in this thesis, confounding the associations in which diet and plasma fatty acids were the exposure. There is value in

including BMI in the statistical models built in this thesis, and presenting results before and after adjustment, which has not been done by many of the articles published. BMI is related to liver fat outcomes, and to both dietary and plasma fatty acids exposures. In some patients living with obesity, higher amounts of free fatty acid are released into the bloodstream, secreted by dysfunctional adipose tissue. Obesity is also associated with excessive liver fat due to the increase in fat tissue and lipid accumulation within the liver cells (14, 152, 153). BMI could have also acted as a mediator in the associations studied in this thesis. However, no mediation analyses were conducted. Incorporating mediation analyses in future research could provide a more nuanced understanding of the indirect and direct effects of BMI, enhancing the depth and robustness of this findings. BMI and visceral adiposity were moderately correlated in this sample (0.68 spearman correlation), which is similar with a meta-analysis carried out about correlations between BMI and VAT (55). However, they are not the same entity as VAT which specifically refers to the metabolically dysfunctional adipose tissue in the abdominal cavity. One of the findings of this thesis, in which the associations between dietary SFAs and liver fat were stronger in those with high VAT, highlights the importance of assessing for effect modification by adiposity: in this case, it may be particularly advisable a greater reduction in dietary SFA intake for patients with high levels of VAT. However, in order to apply this advice further research would be needed to confirm these findings, and to determine specific cut-off values for high VAT.

#### 6.2.8 The role of alcohol intake across analyses

Alcohol intake, a variable challenging to measure as it was based on self-report, had a central role in this thesis. Very detailed data on alcohol consumption allowed the exclusion of participants with excessive drinking, as well as adjustment as a confounder. This variable is key as it is strongly associated with liver fat accumulation, MetALD and ALD (154). It is important however, that while the data on alcohol included a wide range of products, it was

only taken once at baseline, and could have changed since. Therefore, this could have led to misclassification of alcohol intake, introducing information bias. To exclude new cases of liver steatosis related to alcohol intake, any alcohol related fatty liver disease that was incident in hospitals also made the participants excluded, but it is also clear that hospitalisation likely occurs in very advanced and severe cases. As previously mentioned, GP data was not available for this thesis, but could have also been a helpful source to identify participants who had been referred to alcohol cessation or rehabilitation programs, which could have been noted by their GP.

Considering alcohol intake in this thesis exclusion criteria and as an adjustment in the models has allowed to observe unexpected findings in the case of PUFAs. In Chapter 3, after limiting the sample to people with very moderate alcohol intake, the associations between PUFAs and liver fat became non-significant, while in Chapter 4, there were no significant associations between PUFAs and MASLD, which by definition already excludes moderate alcohol drinking. While the alcohol measurement was imperfect, this may suggest that the inverse associations between liver fat and PUFAs are not observed when participants are very light drinkers. However, these findings are unexpected, considering the evidence on the associations between alcohol intake and an increase in DNL, which promotes liver fat accumulation. PUFAs may reduce DNL, and there may be a threshold of DNL at which the inverse associations between PUFAs and liver fat are observed. These findings could have been considered incidental, but as they were observed in two separate settings of moderate consumption, they require attention, and further investigations on associations between PUFAs and liver fat in low alcohol intake contexts are needed.

### 6.3 Methodological considerations

This thesis presents results from observational designs, which cannot establish causality, and presented associations observed at a specific time. There are many aspects of the design of the analyses carried out that are listed below that must be considered when interpreting the results.

#### 6.3.1 Sample characteristics (comparison across samples, size, power, chance findings)

The sample sizes in the UK Biobank varied from 7,201 to 12,013 participants, and the sample in OCDEM was 125. The analyses carried out in the UK Biobank have been able to detect modest associations between dietary exposures and outcomes, and used the largest dataset of liver MRIs available at the moment, which is a strength (78). Due to the age at recruitment and the time at liver MRI, most of the women included in the study were considered part of the Postmenopausal group, and subgroup analysis by menopausal status were not significant, which could be explained by lack of power, or a lack of association at all. Chance findings due to multiple testing could have arisen due to exploring many dietary exposures, producing type 1 errors. However, as an attempt to control for this, there was a correction for multiple testing by lowering the significance threshold to  $<0.0018$  in the case of multiple individual fatty acids.

#### 6.3.2 Ethnicity studied

While this thesis considered the role of Ethnicity in the statistical models, it has been pointed out in previous chapters that the UK Biobank participants were more than 90% White ethnicity, and OCDEM participants were 100% White, which limits the representativeness of this sample, and it is as a main limitation (155). While the UK Biobank represents the largest global repository of MRI liver fat with dietary data, there is still urgency to study these research questions across diverse populations, as steatosis profiles can vary significantly. For example, in Caucasian populations obesity-related steatosis is more prevalent than lean steatosis, whereas Asian populations tend to have a greater proportion of cases of lean steatosis (156).



Additionally, Hispanic adults may have a higher prevalence of MASLD than non-Hispanic adults, and this group was not included in this thesis (157). Moreover, eating behaviours, such as snacking and food preparation, vary significantly between Western and Eastern populations (156). Urbanization and lifestyle changes, including shifts in dietary patterns, may contribute differently to associations between dietary fatty acids and liver steatosis (156). While some studies attribute ethnic disparities to socioeconomic status and social determinants of health, which is partially accounted for by including Townsend index in this thesis, future research must recruit participants with data on both diverse ethnicity and socioeconomic status to better understand these differences. Including a broader range of ethnicity groups is necessary to ensure diverse representation, and to allow power for subgroup analyses, to observe if the associations between dietary fatty acids and liver fat are different across ethnicities. This approach would prevent potential recommendations and evidence-based advice from being limited to a single ethnicity, thereby promoting inclusivity and broader applicability.

### 6.3.3 Selection bias

While the UK biobank invitations to participate attempted to recruit a group as representative as possible from the UK population, only 5% of people contacted accepted to be involved, which could have introduced selection bias, and in particular healthy volunteer bias. This can be observed because compared to the general population, overall UK Biobank participants present some evidence of healthy volunteers selection bias: they report fewer self-reported health conditions and have lower all-cause mortality rates than the general population in both ages 70 to 74 for men and women (155). In addition, participants who completed a dietary assessment presented characteristics of health-consciousness, and they tended to be White, female, older, less deprived and more educated, which increases the selection bias (70).

While still acknowledging these sources of selection bias, it is important to observe that results from analyses using UK Biobank data may still be generalizable to the general population, considering previous studies (124, 125).

The data obtained from participants at the OCDEM trials was obtained from variety of volunteers. While the Oxford Biobank claims to be representative of Oxfordshire, there could be some healthy volunteer bias still, especially considering that there were different participants who volunteered who were not recruited via the Oxford Biobank (93). The selection criteria for these participants also considers that the participants are free from known metabolic disease, and providing a healthy sample- but also includes participants from steatohepatitis trials.

#### 6.3.4 Dietary intake assessment

As nutritional epidemiology studies, the discussion sections of Chapters 3 and 4 present the multiple challenges of self-reported dietary data (158). There are several limitations of self-reported dietary data that have been mentioned in the respective discussion sections that have to be considered, such as understanding that results may be influenced by social desirability bias, as participants may tend to report healthier choices than their usual diet, which would weaken the results observed here. Within person day-to-day variability in dietary intake may also introduce random error. There was an attempt to reduce this by only including participants with at least 2 24-hour dietary assessments, but the use of repeated 24 hour assessments does not completely account for long term variations in diets: even by repeating them in different days and months, it may still present some measurement error, and dietary intake may still change over time (130). The sensitivity analyses were restricted to participants who had completed  $\geq 4$  WebQs, which may have allowed for a better capture of dietary intake, providing more data per participant and reducing random measurement error and within-person variability. However, restricting the analysis to participants who completed  $\geq 4$  WebQs may

have introduced additional selection bias, as this subset is likely to represent a highly motivated and even more health-conscious group compared to the general population of Oxford WebQ responders (70). Still, repeated measurements are necessary in order to obtain more reliable usual dietary intakes. The 24-hour assessment is not as precise as a food diary or as a controlled intervention (3) but is much more suitable for cohort studies (71). This is particularly relevant for liver fat, which is an outcome that is strongly associated with excess of energy intake as evidenced by multiple intervention studies and the strong association with obesity. As strengths, it is a well-accepted method that can be administered online, reducing participants fatigue and the need of interviewers, carried out on different months to capture intakes that may vary through the seasons, calculating nutrients automatically, and at a low cost (67). There is data on eating patterns in the UK Biobank, but this is limited to eating disorders or binge eating, and it would have been useful to have data on eating speed, and pace, as it has been observed that high eating speed is associated with liver steatosis (159). In addition, while it is a very valuable resource, as described in this thesis, the Oxford WebQ is limited by the predefined food items in its database, which may lead to underrepresentation of certain products and the introduction of measurement error.

#### 6.3.5 Liver fat outcomes

For this thesis, the liver fat outcomes ascertainment used MRI, which is a non-invasive, accurate method to measure liver fat, and all MRI scanners were the same brand and type, using the same software across all imaging centres in the UK Biobank (79, 80). As previously mentioned, the availability of MRI measurements of liver fat allows for studying a large number of participants with a very detailed measurement of liver fat (120). However, genetic variants that make participants more susceptible to MASLD were not excluded from the analysis or adjusted for, which is a limitation of this project (160). In addition, MASLD is a multifactorial and fluid condition. Unlike acute outcomes such as stroke or a cancer diagnosis,

its nature allows for changes over time, and it can either improve, allowing a return to a healthy liver when steatosis is still reversible, or deteriorate to progress to more severe stages. It is also silent, undiagnosed in many cases and only screened for in patients with certain conditions. In addition, those who are screened by their GP are usually studied with ultrasound, which detects steatosis once it is very advanced, and it is then difficult to capture steatosis at early stages. Finally, participants who are hospitalised are the one in very advanced stages as well. Therefore, this population based cohort with imaging provides an effective approach to studying MASLD and diet, as it allows for the detection of MASLD cases that would otherwise go unnoticed, during asymptomatic stages.

#### 6.3.6 Reverse causality

Temporality is one of the considerations to look at when interpreting the results of this thesis. While there has been an effort to minimise the potential of reverse causality, by excluding participants with concomitant chronic diseases and conditions that may affect their diet or weight significantly, the excess (or lack of) liver fat accumulation may have preceded the dietary exposure, as data on liver MRIs was obtained once, many years after diet assessment, and changes over time were not calculated. As there are no routine liver ultrasounds or imaging prescribed for the general population, and due to the dynamic nature of liver fat accumulation which can change in only 2 weeks, shown in trials, it is not known if participants' levels of liver fat would have been similar at the time that diet was measured. This represents a major limitation of this thesis, as dietary changes may have been introduced and they may have represented significant changes in their liver (161). This is a limitation of this study that could have been remediated if there were data on repeated liver fat measurements, including one close to the time in which diet was measured. The repeated imaging studies of the UK Biobank were carried out on a subsample that did not complete any further dietary assessments, so there is no exposure measurement that could be included. In the case of the OCDEM dataset analysed

in Chapter 5, blood measurements and liver imaging were carried out on the same period, with only a few days of difference, but it is still a cross-sectional study that cannot establish temporality. Further research is needed to understand the progression of liver fat steatosis, as it can either worsen or improve over time. Since a single MRI does not provide information about the individual's past or future trajectory, there is a possibility of reverse causality, though unlikely, that should be considered.

There may have been undiagnosed chronic disease that were not correctly excluded in many cases, which could have led to changes in diet, and therefore including participants in this study who were no longer consuming similar nutrients as reported before. While data from hospital episodes was used to exclude any new cases, there could have been diagnoses outside this setting, such as primary healthcare that were not detected. While primary health data could have been helpful to detect additional medication or metabolic diseases, as mentioned before, it may be relevant only for incident dyslipidaemia or hypertension, or increased liver enzymes, which are more likely to be diagnosed than liver steatosis.

In the cases of Chapter 4, 5, and 6, the outcome of interest is liver fat or high liver fat as part of MASLD, which is different from other classic outcomes such as cardiovascular events or cancer diagnosis: as this is not an acute event with a clear set of signs and symptoms, and is usually not diagnosed. The valuable aspect of this sample is that it included diabetic participants, which were not studied in Chapter 4, and may benefit from specific advice, different from the overall population.

Residual confounding that was not measured should also be considered in these associations observed. While there was an attempt to understand this by assessing the change in chi-square, they were high in many cases, which estimates that there was high influence of residual confounding. It is important to be aware of potential residual confounding, that may be present

in this observational study, even after the attempt to measure variables as precisely as possible and performing the relevant exclusions.

Self-reported alcohol intake, physical activity and smoking may also have changed over time. These potential changes in behaviour, in particular in alcohol consumption, may have introduced misclassification of the covariate, which could have led to residual confounding as the influence of the true measure of these confounders would not have been controlled, making the associations observed less accurate. While excluding incident cases of “any alcoholic liver disease” hospital episodes, there may still have been participants who have changed their alcohol intake, and not reached the severity needed to go to the hospital, which is a limitation of this thesis.

#### 6.4 New evidence for current dietary guidelines

The positive associations between SFAs and MUFAs and liver fat outcomes, and the inverse associations between PUFAs and liver fat outcomes may be an opportunity to encourage more research about this in dietary trials, and inform guidelines on the dietary management of MASLD. In September 2024, a Delphi consensus published by 55 experts from 27 countries concluded that a reduction in SFAs consumption is advised for the approach to MASLD, due to the emerging accumulating evidence on its harmful associations with liver fat (21). However, reduction in SFAs is still not part of all official guidelines for MASLD approach, and there is still no clear advice on dietary fatty acid consumption, apart from reducing total fat and calorie intake. Currently, the latest guidelines recommend consumption of a Mediterranean diet, rather than specifying a group of fatty acids (101, 103, 104, 162). The proceedings from the American Association of the study of Liver Disease, states that there is not enough evidence on dietary approaches for MASLD from long-term trials that use biopsy to measure outcomes (162). While this would be a very precise measurement of liver histology and provide a very accurate measure of liver fat, it would mean the need of biopsies, likely hundreds, that are invasive and

carry more risks (162). MRI scans still provide good measurements of fat and are able to detect modest changes in liver fat over time, and a recent meta-analysis has shown that their implementation has similar performance as liver biopsies (120).

Unlike findings from other studies, no individual fatty acid presented associations with liver fat that were significantly different from the ones anticipated based on their classification by level of saturation (SFAs, MUFAs, and PUFAs). This suggests that dietary guidance could, for now, focus on the major structural fatty acid groups, as there is no consistent evidence supporting a particularly different role of any single fatty acid. Emphasizing individual fatty acids would add unnecessary complexity, while we still do not have enough evidence, and making guidelines more challenging to implement or less acceptable to the population.

In this thesis and other studies, the UK Biobank has provided valuable insights into the associations between different macronutrients and liver fat accumulation, including carbohydrates and protein. However, further research is needed to explore the complex interplay between macronutrients and other dietary components (20, 21, 32).

## 6.5 Recommendations for future research

The definition of MASLD is clear, but it has still not been fully implemented in new studies. The new set of definitions introduced in 2023 brought new challenges, and correct measurement of MASLD may not always be simple, as it requires access to imaging studies and adequate classification of alcohol intake. While the SLD definitions are still recent, correct implementation in further studies, along with adequate measurement tools, will be key to continue the research to understand this disease. The definition and the way steatosis and cardiometabolic factors are measured, and the correct measurement of alcohol to exclude cases of alcohol related steatosis, would be key for correct identification of MASLD cases in new studies. In addition, further research is needed to clarify whether plasma fatty acids are linked

to MASLD, diet, or both. Fatty acids in plasma are still challenged by some as valid markers of diet (143). Advancing our understanding of the role of plasma fatty acids in liver fat accumulation is still necessary, as their use as biomarkers is still not fully understood and prone to error. More research on diverse populations, wider age ranges (especially pre and post-menopausal women) and backgrounds, will be more helpful to understand how dietary fatty acids relate to liver outcomes in non-White populations. Large samples will be necessary to ensure adequate power, with adequate dietary assessment, and always considering the role of BMI in analyses (or other measurement of adiposity). The UK Biobank aims to have data from 100,000 MRIs by 2025, which after carrying out pertinent exclusions, would enable the study of this research question in an even bigger sample. The use of repeated imaging studies to assess changes in liver fat over time, rather than relying on a single measurement would be key. In further studies, it would be helpful to recruit participants with different genetic risk scores. Interventional studies that look at the effect of dietary fatty acids in liver fat content may also include data on individual fatty acids, recruit participants from different ethnicities, and attempt to collect more data on VLDL-TG fatty acid composition in larger samples.

## 6.6 Conclusion

Given the recent introduction of the new definition of MASLD, and the growing interest in questions about diet and liver fat, this thesis provides valuable insights into this rapidly evolving area of study, to understand better how fatty acids relate to liver fat. While acknowledging that both dietary habits and MASLD are complex and multifactorial, and challenging to measure in observational studies, these findings still offer valuable insights despite their limitations. This is the first study in the UK Biobank that studies diet and MASLD, and that uses the individual dietary fatty acids dataset. The novel analyses carried out in this thesis support previously published research on the positive associations between dietary SFAs and liver fat outcomes, reinforcing current recommendations. Still the evidence for dietary



PUFAs and MUFAs is not as clear. While there is a strong focus and interest in MASLD research in the field, there is an opportunity to understand and do high quality research on this topic, to improve the significant global burden of this disease.

MASLD (and overall excessive liver fat accumulation) is a complex, multifactorial condition that affects millions of people globally. Dietary habits are also complex behaviours, challenging to measure, that comprehend nutrients but also environment, setting and culture.

While pharmacological trials are in place and mainly focus on advanced stages of MASLD, there is still significant opportunity to focus on preventive approaches through lifestyle changes. With more robust, high-quality evidence on the role of diet in MASLD, dietary advice could become increasingly precise and impactful. Diet is already a core element of multidisciplinary approaches to MASLD, and this would further enhance efforts to reduce the global burden of the disease.

## References

1. Corless JK, Middleton HM, 3rd. Normal liver function. A basis for understanding hepatic disease. *Arch Intern Med*. 1983;143(12):2291-4.
2. van Herpen NA, Schrauwen-Hinderling VB. Lipid accumulation in non-adipose tissue and lipotoxicity. *Physiology & Behavior*. 2008;94(2):231-41.
3. Theise ND. *Robbins & Cotran Pathologic Basis of Disease*. Ninth Edition ed: Elsevier Inc; 2015. p. 821-81.
4. Lonardo A, Leoni S, Alswat KA, Fouad Y. History of Nonalcoholic Fatty Liver Disease. *International Journal of Molecular Sciences*. 2020;21(16):5888.
5. Ayonrinde OT. Historical narrative from fatty liver in the nineteenth century to contemporary NAFLD - Reconciling the present with the past. *JHEP Rep*. 2021;3(3):100261.
6. Glen J, Floros L, Day C, Pryke R. Non-alcoholic fatty liver disease (NAFLD): summary of NICE guidance. *Bmj*. 2016;354.
7. Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, et al. A multi-society Delphi consensus statement on new fatty liver disease nomenclature. *J Hepatol*. 2023.
8. Paik JM, Golabi P, Younossi Y, Mishra A, Younossi ZM. Changes in the global burden of chronic liver diseases from 2012 to 2017: the growing impact of NAFLD. *Hepatology*. 2020;72(5):1605-16.
9. Younossi Z, Tacke F, Arrese M, Chander Sharma B, Mostafa I, Bugianesi E, et al. Global Perspectives on Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. *Hepatology*. 2019;69(6):2672-82.
10. Younossi ZM, Blissett D, Blissett R, Henry L, Stepanova M, Younossi Y, et al. The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. *Hepatology*. 2016;64(5):1577-86.
11. Younossi Z, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. *Nature reviews Gastroenterology & hepatology*. 2018;15(1):11-20.
12. Wang J, He W, Tsai P-J, Chen P-H, Ye M, Guo J, et al. Mutual interaction between endoplasmic reticulum and mitochondria in nonalcoholic fatty liver disease. *Lipids in health and disease*. 2020;19(1):1-19.
13. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nature genetics*. 2008;40(12):1461-5.
14. Greenberg AS, Obin MS. Obesity and the role of adipose tissue in inflammation and metabolism<sup>23</sup>. *The American Journal of Clinical Nutrition*. 2006;83(2):461S-5S.
15. Albhaisi S, Chowdhury A, Sanyal AJ. Non-alcoholic fatty liver disease in lean individuals. *JHEP Reports*. 2019;1(4):329-41.
16. Parry SA, Hodson L. Influence of dietary macronutrients on liver fat accumulation and metabolism. *J Investig Med*. 2017;65(8):1102-15.
17. DeLany JP, Windhauser MM, Champagne CM, Bray GA. Differential oxidation of individual dietary fatty acids in humans. *The American journal of clinical nutrition*. 2000;72(4):905-11.
18. Lombardi R, Iuculano F, Pallini G, Fargion S, Fracanzani AL. Nutrients, Genetic Factors, and Their Interaction in Non-Alcoholic Fatty Liver Disease and Cardiovascular Disease. *International Journal of Molecular Sciences*. 2020;21(22):8761.
19. Hydes T, Alam U, Cuthbertson DJ. The Impact of Macronutrient Intake on Non-alcoholic Fatty Liver Disease (NAFLD): Too Much Fat, Too Much Carbohydrate, or Just Too Many Calories? *Frontiers in nutrition*. 2021;8:42.
20. Orliacq J, Pérez-Cornago A, Parry SA, Kelly RK, Koutoukidis DA, Carter JL. Associations between types and sources of dietary carbohydrates and liver fat: a UK Biobank study. *BMC Medicine*. 2023;21(1):444.

21. Zeng X-F, Varady KA, Wang X-D, Targher G, Byrne CD, Tayyem R, et al. The role of dietary modification in the prevention and management of metabolic dysfunction-associated fatty liver disease: An international multidisciplinary expert consensus. *Metabolism*. 2024;161:156028.
22. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, Abdelmalek MF, Caldwell S, Barb D, et al. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatology*. 2023;77(5):1797-835.
23. Yki-Järvinen H, Luukkonen PK, Hodson L, Moore JB. Dietary carbohydrates and fats in nonalcoholic fatty liver disease. *Nature Reviews Gastroenterology & Hepatology*. 2021;1-17.
24. Hodson L, Frayn KN. Hepatic fatty acid partitioning. *Current opinion in lipidology*. 2011;22(3):216-24.
25. Green CJ, Pramfalk C, Charlton CA, Gunn PJ, Cornfield T, Pavlides M, et al. Hepatic de novo lipogenesis is suppressed and fat oxidation is increased by omega-3 fatty acids at the expense of glucose metabolism. *BMJ Open Diabetes Res Care*. 2020;8(1).
26. Hodson L. Hepatic fatty acid synthesis and partitioning: the effect of metabolic and nutritional state. *Proc Nutr Soc*. 2019;78(1):126-34.
27. Kremmyda L-S, Tvřzicka E, Stankova B, Zak A. Fatty acids as biocompounds: their role in human metabolism, health and disease- A review. Part 2: Fatty acid physiological roles and applications in human health and disease. *Biomedical Papers of the Medical Faculty of Palacky University in Olomouc*. 2011;155(3).
28. Tvřzicka E, Kremmyda L-S, Stankova B, Zak A. Fatty acids as biocompounds: their role in human metabolism, health and disease-a review. part 1: classification, dietary sources and biological functions. *Biomedical Papers of the Medical Faculty of Palacky University in Olomouc*. 2011;155(2).
29. Tsompanaki E, Thanapirom K, Papatheodoridi M, Parikh P, de Lima YC, Tsochatzis EA. Systematic review and meta-analysis: the role of diet in the development of nonalcoholic fatty liver disease. *Clinical Gastroenterology and Hepatology*. 2021.
30. Nemer M, Osman F, Said A. Dietary macro and micronutrients associated with MASLD: Analysis of a national US cohort database. *Annals of Hepatology*. 2024;29(3):101491.
31. Chen S, Pang J, Huang R, You Y, Zhang H, Xue H, et al. Associations of Macronutrients Intake With MRI-determined Hepatic Fat Content, Hepatic Fibroinflammation, and NAFLD. *J Clin Endocrinol Metab*. 2023;108(12):e1660-e9.
32. Fridén M, Mora AM, Lind L, Risérus U, Kullberg J, Rosqvist F. Diet composition, nutrient substitutions and circulating fatty acids in relation to ectopic and visceral fat depots. *Clinical Nutrition*. 2023;42(10):1922-31.
33. Tian A, Sun Z, Zhang M, Li J, Pan X, Chen P. Associations between dietary fatty acid patterns and non-alcoholic fatty liver disease in typical dietary population: A UK biobank study. *Frontiers in Nutrition*. 2023;10:1117626.
34. Cui J, Li L, Ren L, Sun J, Zhao H, Sun Y. Dietary n-3 and n-6 fatty acid intakes and NAFLD: A cross-sectional study in the United States. *Asia Pacific Journal of Clinical Nutrition*. 2021;30(1).
35. López-Bautista F, Barbero-Becerra VJ, Ríos MY, Ramírez-Cisneros MÁ, Sánchez-Pérez CA, Ramos-Ostos MH, et al. Dietary consumption and serum pattern of bioactive fatty acids in NAFLD patients. *Annals of Hepatology*. 2020;19(5):482-8.
36. Nouredin M, Zelber-Sagi S, Wilkens LR, Porcel J, Boushey CJ, Le Marchand L, et al. Diet Associations With Nonalcoholic Fatty Liver Disease in an Ethnically Diverse Population: The Multiethnic Cohort. *Hepatology*. 2020;71(6):1940-52.
37. Alferink LJM, Kieft-De Jong JC, Erler NS, Veldt BJ, Schoufour JD, De Knecht RJ, et al. Association of dietary macronutrient composition and non-alcoholic fatty liver disease in an ageing population: The Rotterdam Study. *Gut*. 2019;68(6):1088-98.
38. Vernekar M, Singhal R, Joshi K, Amarapurkar D. Variation in the plasma levels of polyunsaturated fatty acids in control Vis-à-Vis nonalcoholic fatty liver disease subjects and its possible association with gut microbiome. *Metabolic syndrome and related disorders*. 2018;16(7):329-35.
39. Cheng Y, Zhang K, Chen Y, Li Y, Li Y, Fu K, et al. Associations between Dietary Nutrient Intakes and Hepatic Lipid Contents in NAFLD Patients Quantified by <sup>1</sup>H-MRS and Dual-Echo MRI. *Nutrients*. 2016;8(9).

40. Oya J, Nakagami T, Sasaki S, Jimba S, Murakami K, Kasahara T, et al. Intake of n-3 polyunsaturated fatty acids and non-alcoholic fatty liver disease: a cross-sectional study in Japanese men and women. *European journal of clinical nutrition*. 2010;64(10):1179-85.
41. Allard JP, Aghdassi E, Mohammed S, Raman M, Avand G, Arendt BM, et al. Nutritional assessment and hepatic fatty acid composition in non-alcoholic fatty liver disease (NAFLD): a cross-sectional study. *Journal of hepatology*. 2008;48(2):300-7.
42. Zelber-Sagi S, Nitzan-Kaluski D, Goldsmith R, Webb M, Blendis L, Halpern Z, et al. Long term nutritional intake and the risk for non-alcoholic fatty liver disease (NAFLD): a population based study. *Journal of hepatology*. 2007;47(5):711-7.
43. Meisinger C, Rospleszcz S, Wintermeyer E, Lorbeer R, Thorand B, Bamberg F, et al. Isocaloric Substitution of Dietary Carbohydrate Intake with Fat Intake and MRI-Determined Total Volumes of Visceral, Subcutaneous and Hepatic Fat Content in Middle-Aged Adults. *Nutrients*. 2019;11(5).
44. Bjermo H, Iggman D, Kullberg J, Dahlman I, Johansson L, Persson L, et al. Effects of n-6 PUFAs compared with SFAs on liver fat, lipoproteins, and inflammation in abdominal obesity: a randomized controlled trial. *The American journal of clinical nutrition*. 2012;95(5):1003-12.
45. Luukkonen PK, Sädevirta S, Zhou Y, Kayser B, Ali A, Ahonen L, et al. Saturated fat is more metabolically harmful for the human liver than unsaturated fat or simple sugars. *Diabetes care*. 2018;41(8):1732-9.
46. Rosqvist F, Iggman D, Kullberg J, Cedernaes J, Johansson H-E, Larsson A, et al. Overfeeding polyunsaturated and saturated fat causes distinct effects on liver and visceral fat accumulation in humans. *Diabetes*. 2014;63(7):2356-68.
47. Rosqvist F, Kullberg J, Ståhlman M, Cedernaes J, Heurling K, Johansson H-E, et al. Overeating saturated fat promotes fatty liver and ceramides compared with polyunsaturated fat: a randomized trial. *The Journal of Clinical Endocrinology & Metabolism*. 2019;104(12):6207-19.
48. Parry SA, Rosqvist F, Mozes FE, Cornfield T, Hutchinson M, Piche ME, et al. Intrahepatic Fat and Postprandial Glycemia Increase After Consumption of a Diet Enriched in Saturated Fat Compared With Free Sugars. *Diabetes Care*. 2020;43(5):1134-41.
49. Utzschneider KM, Bayer-Carter JL, Arbuckle MD, Tidwell JM, Richards TL, Craft S. Beneficial effect of a weight-stable, low-fat/low-saturated fat/low-glycaemic index diet to reduce liver fat in older subjects. *British journal of nutrition*. 2013;109(6):1096-104.
50. Bozzetto L, Prinster A, Annuzzi G, Costagliola L, Mangione A, Vitelli A, et al. Liver fat is reduced by an isoenergetic MUFA diet in a controlled randomized study in type 2 diabetic patients. *Diabetes care*. 2012;35(7):1429-35.
51. Wan Y, Wang D, Li H, Xu Y. The imaging techniques and diagnostic performance of ultrasound, CT, and MRI in detecting liver steatosis and fat quantification: A systematic review. *Journal of Radiation Research and Applied Sciences*. 2023;16(4):100658.
52. Khov N, Sharma A, Riley TR. Bedside ultrasound in the diagnosis of nonalcoholic fatty liver disease. *World J Gastroenterol*. 2014;20(22):6821-5.
53. Rhee JJ, Cho E, Willett WC. Energy adjustment of nutrient intakes is preferable to adjustment using body weight and physical activity in epidemiological analyses. *Public Health Nutr*. 2014;17(5):1054-60.
54. Preiss D, Sattar N. Non-alcoholic fatty liver disease: an overview of prevalence, diagnosis, pathogenesis and treatment considerations. *Clin Sci (Lond)*. 2008;115(5):141-50.
55. Mouchti S, Orliacq J, Reeves G, Chen Z. Assessment of correlation between conventional anthropometric and imaging-derived measures of body fat composition: a systematic literature review and meta-analysis of observational studies. *BMC Med Imaging*. 2023;23(1):127.
56. Milić S, Lulić D, Štimac D. Non-alcoholic fatty liver disease and obesity: biochemical, metabolic and clinical presentations. *World journal of gastroenterology: WJG*. 2014;20(28):9330.
57. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*. 2015;12(3):e1001779.
58. UK Biobank website [Available from: <https://www.ukbiobank.ac.uk/>].
59. UK Biobank Study Protocol. Protocol for a large-scale prospective epidemiological resource Protocol No: UKBB-PROT-09-06 2007. Available from: [https://www.ukbiobank.ac.uk/media/3sbeknnz/ukbiobank\\_protocol.pdf](https://www.ukbiobank.ac.uk/media/3sbeknnz/ukbiobank_protocol.pdf).

60. UK Biobank Research Ethics Approval

[Available from: <https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/ethics>.

61. UK Biobank Ethical Approval Renewal 2021 [Available from: <https://www.ukbiobank.ac.uk/media/p4yjfqcp/2021-nwrec-rtb-application-and-approval.pdf>.

62. UK Biobank Resource 100230: Procedures for participant consent. Available from: <https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=100230>.

63. UK Biobank Resource 100247: Touch Screen Questionnaire2011. Available from: <https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=100247>.

64. UK Biobank Resource 100238 Order of Data Collection 2011. Available from: <https://biobank.cts.ox.ac.uk/crystal/refer.cgi?id=100238>.

65. UK Biobank Resource 100235: Verbal Interview stage2012. Available from: <https://biobank.cts.ox.ac.uk/crystal/refer.cgi?id=100235>.

66. UK Biobank Resource 146620: Anthropometry2014.

67. Liu B, Young H, Crowe FL, Benson VS, Spencer EA, Key TJ, et al. Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. *Public health nutrition*. 2011;14(11):1998-2005.

68. UK Biobank Resource 118240: 24-hour dietary recall questionnaire Version 1.42024. Available from: <https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=118240>.

69. Perez-Cornago A, Pollard Z, Young H, van Uden M, Andrews C, Piernas C, et al. Description of the updated nutrition calculation of the Oxford WebQ questionnaire and comparison with the previous version among 207,144 participants in UK Biobank. *European Journal of Nutrition*. 2021;1-12.

70. Galante J, Adamska L, Young A, Young H, Littlejohns TJ, Gallacher J, et al. The acceptability of repeat Internet-based hybrid diet assessment of previous 24-h dietary intake: administration of the Oxford WebQ in UK Biobank. *British Journal of Nutrition*. 2016;115(4):681-6.

71. Greenwood DC, Hardie LJ, Frost GS, Alwan NA, Bradbury KE, Carter M, et al. Validation of the Oxford WebQ online 24-hour dietary questionnaire using biomarkers. *American journal of epidemiology*. 2019;188(10):1858-67.

72. Kelly RK, Pollard Z, Young H, Piernas C, Lentjes M, Mulligan A, et al. Evaluation of the New Individual Fatty Acid Dataset for UK Biobank: Analysis of Intakes and Sources in 207,997 Participants. *Nutrients*. 2022;14(17):3603.

73. Kawashima H. Intake of arachidonic acid-containing lipids in adult humans: dietary surveys and clinical trials. *Lipids in Health and Disease*. 2019;18(1):101.

74. UK Biobank Resource 3000: NMR metabolomics companion document. Available from: [biobank.ndph.ox.ac.uk/ukb/ukb/docs/nmrm\\_companion\\_doc.pdf](https://biobank.ndph.ox.ac.uk/ukb/ukb/docs/nmrm_companion_doc.pdf).

75. UK Biobank Resource 100227 : Blood sampling procedures using ACE2011. Available from: <https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=100227>.

76. Julkunen H, Cichońska A, Tiainen M, Koskela H, Nybo K, Mäkelä V, et al. Atlas of plasma NMR biomarkers for health and disease in 118,461 individuals from the UK Biobank. *Nat Commun*. 2023;14(1):604.

77. UK Biobank Resource 130 NMR metabolomics companion document (Phase 2 update). Available from: <https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=130>.

78. Littlejohns TJ, Holliday J, Gibson LM, Garratt S, Oesingmann N, Alfaro-Almagro F, et al. The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. *Nature communications*. 2020;11(1):1-12.

79. UK Biobank Resource 348 Body (Abdominal) MRI Scans2015.

80. Starekova J, Hernando D, Pickhardt PJ, Reeder SB. Quantification of liver fat content with CT and MRI: state of the art. *Radiology*. 2021;301(2):250-62.

81. Costa DN, Pedrosa I, McKenzie C, Reeder SB, Rofsky NM. Body MRI using IDEAL. *AJR Am J Roentgenol*. 2008;190(4):1076-84.

82. West J, Dahlqvist Leinhard O, Romu T, Collins R, Garratt S, Bell JD, et al. Feasibility of MR-based body composition analysis in large scale population studies. *PloS one*. 2016;11(9):e0163332.

83. Borga M, West J, Bell JD, Harvey NC, Romu T, Heymsfield SB, et al. Advanced body composition assessment: from body mass index to body composition profiling. *Journal of Investigative Medicine*. 2018;66(5):1-9.
84. UK Biobank Data Showcase: Data-Field 21000, Ethnic background [Available from: <https://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=21000>].
85. Yousaf SB, A. Townsend Deprivation Scores Report. UK Data Service; 2017.
86. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) Short and Long Forms.
87. A healthy lifestyle - WHO recommendations. Body Mass Index- BMI [Available from: <https://www.who.int/europe/news-room/fact-sheets/item/a-healthy-lifestyle---who-recommendations>].
88. NICE Alcohol-use disorders: prevention Public health guideline Available from: <https://www.nice.org.uk/guidance/ph24/resources/alcoholuse-disorders-prevention-pdf-1996237007557>.
89. Ryu S, Suh B-S, Chang Y, Kwon M-J, Yun KE, Jung H-S, et al. Menopausal stages and non-alcoholic fatty liver disease in middle-aged women. *European Journal of Obstetrics & Gynecology and Reproductive Biology*. 2015;190:65-70.
90. Singh A, Dhruve R, Kumar V, Sohal A. Prevalence of Steatotic Liver Disease (MASLD, MetALD, and ALD) in Post-Menopausal Women: NHANES 2017-2020. *Official journal of the American College of Gastroenterology | ACG*. 2024;119(10S):S1267-S8.
91. NICE guideline: Menopause identification and management 2015. Available from: <https://www.nice.org.uk/guidance/ng23/resources/menopause-identification-and-management-pdf-1837330217413>.
92. Gentile C, Weir T, Cox-York K, Wei Y, Wang D, Reese L, et al. The role of visceral and subcutaneous adipose tissue fatty acid composition in liver pathophysiology associated with NAFLD. *Adipocyte*. 2015;4(2):101-12.
93. Karpe F, Vasan SK, Humphreys SM, Miller J, Cheeseman J, Dennis AL, et al. Cohort Profile: The Oxford Biobank. *Int J Epidemiol*. 2018;47(1):21-g.
94. Purvis LAB, Clarke WT, Biasioli L, Valković L, Robson MD, Rodgers CT. OXSA: An open-source magnetic resonance spectroscopy analysis toolbox in MATLAB. *PLoS One*. 2017;12(9):e0185356.
95. Hodson L, McQuaid SE, Karpe F, Frayn KN, Fielding BA. Differences in partitioning of meal fatty acids into blood lipid fractions: a comparison of linoleate, oleate, and palmitate. *Am J Physiol Endocrinol Metab*. 2009;296(1):E64-71.
96. Heath RB, Karpe F, Milne RW, Burdge GC, Wootton SA, Frayn KN. Selective partitioning of dietary fatty acids into the VLDL TG pool in the early postprandial period. *Journal of lipid research*. 2003;44(11):2065-72.
97. Folch J, Lees M, Stanley GHS. A simple method for the isolation and purification of total lipides from animal tissues. *Journal of Biological Chemistry*. 1957;226(1):497-509.
98. Okuno T, Yokomizo T. Basic Techniques for Lipid Extraction from Tissues and Cells. In: Yokomizo T, Murakami M, editors. *Bioactive Lipid Mediators: Current Reviews and Protocols*. Tokyo: Springer Japan; 2015. p. 331-6.
99. Burdge GC, Wright P, Jones AE, Wootton SA. A method for separation of phosphatidylcholine, triacylglycerol, non-esterified fatty acids and cholesterol esters from plasma by solid-phase extraction. *Br J Nutr*. 2000;84(5):781-7.
100. Eder K. Gas chromatographic analysis of fatty acid methyl esters. *J Chromatogr B Biomed Appl*. 1995;671(1-2):113-31.
101. EASL-EASD-EASO Clinical Practice Guidelines on the management of metabolic dysfunction-associated steatotic liver disease (MASLD). *J Hepatol*. 2024;81(3):492-542.
102. Bischoff SC, Bernal W, Dasarthy S, Merli M, Plank LD, Schütz T, et al. ESPEN practical guideline: Clinical nutrition in liver disease. *Clin Nutr*. 2020;39(12):3533-62.
103. Eslam M, Sarin SK, Wong VW-S, Fan J-G, Kawaguchi T, Ahn SH, et al. The Asian Pacific Association for the Study of the Liver clinical practice guidelines for the diagnosis and management of metabolic associated fatty liver disease. *Hepatology International*. 2020;14(6):889-919.

104. Montemayor S, García S, Monserrat-Mesquida M, Tur JA, Bouzas C. Dietary Patterns, Foods, and Nutrients to Ameliorate Non-Alcoholic Fatty Liver Disease: A Scoping Review. *Nutrients*. 2023;15(18):3987.
105. Das A, Tang YLM, Althumiri NA, Garcia-Larsen V, Schattenberg JM, Alqahtani SA. Fatty acid composition but not quantity is an important indicator of non-alcoholic fatty liver disease: a systematic review. *Eur J Clin Nutr*. 2023.
106. Hodson L, Rosqvist F, Parry SA. The influence of dietary fatty acids on liver fat content and metabolism. *Proceedings of the Nutrition Society*. 2020;79(1):30-41.
107. Willett W. *Nutritional epidemiology*. New York: Oxford University Press; 2017.
108. Lonardo A, Nascimbeni F, Ballestri S, Fairweather D, Win S, Than TA, et al. Sex Differences in Nonalcoholic Fatty Liver Disease: State of the Art and Identification of Research Gaps. *Hepatology*. 2019;70(4):1457-69.
109. Sargeant JA, Gray LJ, Bodicoat DH, Willis SA, Stensel DJ, Nimmo MA, et al. The effect of exercise training on intrahepatic triglyceride and hepatic insulin sensitivity: a systematic review and meta-analysis. *Obes Rev*. 2018;19(10):1446-59.
110. Tomova GD, Arnold KF, Gilthorpe MS, Tennant PW. Adjustment for energy intake in nutritional research: a causal inference perspective. *medRxiv*. 2021.
111. Willett W, Willett W. *Implications of Total Energy Intake for Epidemiologic Analyses. Nutritional Epidemiology*: Oxford University Press; 2012. p. 260–86.
112. Robeva R, Mladenović D, Vesković M, Hrnčić D, Bjekić-Macut J, Stanojlović O, et al. The interplay between metabolic dysregulations and non-alcoholic fatty liver disease in women after menopause. *Maturitas*. 2021;151:22-30.
113. Lee DK. Data transformation: a focus on the interpretation. *Korean J Anesthesiol*. 2020;73(6):503-8.
114. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *Bmj*. 1995;310(6973):170.
115. Coulston AMD, Linda ; Ferruzzi, Mario ; Boushey, Carol J. *Nutrition in the Prevention and Treatment of Disease, Chapter 1: Research Methodology*. United States: Elsevier Science & Technology 2017.
116. Baranowski T, Willett W. 24 Hour Recall and Diet Record Methods. *Nutritional Epidemiology*: Oxford University Press; 2012. p. 49-69.
117. Arnold M. Jasper makes plots. R package version 2–266. <https://github.com/arnhew99/Jasper>. 2020.
118. Parry SA, Rosqvist F, Cornfield T, Barrett A, Hodson L. Oxidation of dietary linoleate occurs to a greater extent than dietary palmitate in vivo in humans. *Clin Nutr*. 2021;40(3):1108-14.
119. Madsen L, Petersen RK, Kristiansen K. Regulation of adipocyte differentiation and function by polyunsaturated fatty acids. *Biochim Biophys Acta*. 2005;1740(2):266-86.
120. Azizi N, Naghibi H, Shakiba M, Morsali M, Zarei D, Abbastabar H, et al. Evaluation of MRI proton density fat fraction in hepatic steatosis: a systematic review and meta-analysis. *Eur Radiol*. 2024.
121. Willett W, Willett W. *Nature of Variation in Diet. Nutritional Epidemiology*: Oxford University Press; 2012. p. 0.
122. Jackson A. KT, Williams C., Hughes I., Chipman K., Rowland I. . Advice on fish consumption: benefits & risks. *Scientific Advisory Committee on Nutrition*; 2004.
123. Hebert JR, Hurley TG, Peterson KE, Resnicow K, Thompson FE, Yaroch AL, et al. Social Desirability Trait Influences on Self-Reported Dietary Measures among Diverse Participants in a Multicenter Multiple Risk Factor Trial<sup>12</sup>. *The Journal of Nutrition*. 2008;138(1):226S-34S.
124. Allen NE, Lacey B, Lawlor DA, Pell JP, Gallacher J, Smeeth L, et al. Prospective study design and data analysis in UK Biobank. *Sci Transl Med*. 2024;16(729):eadf4428.
125. Batty GD, Gale CR, Kivimäki M, Deary IJ, Bell S. Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *bmj*. 2020;368.
126. Song SJ, Lai JC, Wong GL, Wong VW, Yip TC. Can we use old NAFLD data under the new MASLD definition? *J Hepatol*. 2023.

127. Park HJ, Lee S, Lee JS. Differences in the prevalence of NAFLD, MAFLD, and MASLD according to changes in the nomenclature in a health check-up using MRI-derived proton density fat fraction. *Abdom Radiol (NY)*. 2024.
128. Hagström H, Vessby J, Ekstedt M, Shang Y. 99% of patients with NAFLD meet MASLD criteria and natural history is therefore identical. *J Hepatol*. 2024;80(2):e76-e7.
129. Younossi ZM, Paik JM, Stepanova M, Ong J, Alqahtani S, Henry L. Clinical profiles and mortality rates are similar for metabolic dysfunction-associated steatotic liver disease and non-alcoholic fatty liver disease. *J Hepatol*. 2024;80(5):694-701.
130. Willett W, Willett W. Correction for the Effects of Measurement Error. *Nutritional Epidemiology*: Oxford University Press; 2012. p. 287–304.
131. Jabłonowska-Lietz B, Wrzosek M, Włodarczyk M, Nowicka G. New indexes of body fat distribution, visceral adiposity index, body adiposity index, waist-to-height ratio, and metabolic disturbances in the obese. *Kardiologia Pol*. 2017;75(11):1185-91.
132. Hodson L, Skeaff CM, Fielding BA. Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake. *Progress in Lipid Research*. 2008;47(5):348-80.
133. Gnatiuc Friedrichs L, Trichia E, Aguilar-Ramirez D, Preiss D. Metabolic profiling of MRI-measured liver fat in the UK Biobank. *Obesity (Silver Spring)*. 2023;31(4):1121-32.
134. Mäkelä TNK, Tuomainen TP, Hantunen S, Virtanen JK. Associations of serum n-3 and n-6 polyunsaturated fatty acids with prevalence and incidence of nonalcoholic fatty liver disease. *Am J Clin Nutr*. 2022;116(3):759-70.
135. Feingold KR. *Introduction to Lipids and Lipoproteins*. South Dartmouth (MA); 2000.
136. Parry SA, Rosqvist F, Peters S, Young RK, Cornfield T, Dyson P, et al. The influence of nutritional state on the fatty acid composition of circulating lipid fractions: implications for their use as biomarkers of dietary fat intake. *Ups J Med Sci*. 2021;126.
137. Lee JJ, Lambert JE, Hovhannisyan Y, Ramos-Roman MA, Trombold JR, Wagner DA, et al. Palmitoleic acid is elevated in fatty liver disease and reflects hepatic lipogenesis<sup>1234</sup>. *The American Journal of Clinical Nutrition*. 2015;101(1):34-43.
138. Piché ME, Parry SA, Karpe F, Hodson L. Chylomicron-Derived Fatty Acid Spillover in Adipose Tissue: A Signature of Metabolic Health? *J Clin Endocrinol Metab*. 2018;103(1):25-34.
139. Heeren J, Scheja L. Metabolic-associated fatty liver disease and lipoprotein metabolism. *Molecular Metabolism*. 2021;50:101238.
140. Green CJ, Parry SA, Gunn PJ, Ceresa CD, Rosqvist F, Piché M-E, et al. Studying non-alcoholic fatty liver disease: the ins and outs of in vivo, ex vivo and in vitro human models. *Hormone molecular biology and clinical investigation*. 2020;41(1).
141. StataCorp T, United States. *StataNow 18.5*. 2024.
142. Wang X, Magkos F, Mittendorfer B. Sex differences in lipid and lipoprotein metabolism: it's not just about sex hormones. *J Clin Endocrinol Metab*. 2011;96(4):885-93.
143. Marchioni DM, de Oliveira MF, Carioca AAF, Miranda AAM, Carvalho AM, Oki E, et al. Plasma fatty acids: Biomarkers of dietary intake? *Nutrition*. 2019;59:77-82.
144. Baylin A, Kim MK, Donovan-Palmer A, Siles X, Dougherty L, Tocco P, et al. Fasting whole blood as a biomarker of essential fatty acid intake in epidemiologic studies: comparison with adipose tissue and plasma. *Am J Epidemiol*. 2005;162(4):373-81.
145. Saturated Fats and health, Scientific Advisory Committee on Nutrition. 2019.
146. Fabbrini E, Mohammed BS, Magkos F, Korenblat KM, Patterson BW, Klein S. Alterations in Adipose Tissue and Hepatic Lipid Kinetics in Obese Men and Women With Nonalcoholic Fatty Liver Disease. *Gastroenterology*. 2008;134(2):424-31.
147. Fabbrini E, Sullivan S, Klein S. Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, and clinical implications. *Hepatology*. 2010;51(2):679-89.
148. Mittendorfer B, Yoshino M, Patterson BW, Klein S. VLDL Triglyceride Kinetics in Lean, Overweight, and Obese Men and Women. *J Clin Endocrinol Metab*. 2016;101(11):4151-60.
149. Kim EY, Lee YJ, Kwon YJ, Lee JW. Age at menopause and risk of metabolic dysfunction-associated fatty liver disease: A 14-year cohort study. *Dig Liver Dis*. 2024.
150. Pramfalk C, Pavlides M, Banerjee R, McNeil CA, Neubauer S, Karpe F, et al. Sex-specific differences in hepatic fat oxidation and synthesis may explain the higher propensity for NAFLD in men. *The Journal of Clinical Endocrinology & Metabolism*. 2015;100(12):4425-33.



151. Ko SH, Jung Y. Energy Metabolism Changes and Dysregulated Lipid Metabolism in Postmenopausal Women. *Nutrients*. 2021;13(12).
152. Sun B, Karin M. Obesity, inflammation, and liver cancer. *Journal of Hepatology*. 2012;56(3):704-13.
153. Paniagua JA. Nutrition, insulin resistance and dysfunctional adipose tissue determine the different components of metabolic syndrome. *World J Diabetes*. 2016;7(19):483-514.
154. Mackowiak B, Fu Y, Maccioni L, Gao B. Alcohol-associated liver disease. *J Clin Invest*. 2024;134(3).
155. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *American journal of epidemiology*. 2017;186(9):1026-34.
156. Han MAT, Yu Q, Tafesh Z, Pysopoulos N. Diversity in NAFLD: A Review of Manifestations of Nonalcoholic Fatty Liver Disease in Different Ethnicities Globally. *J Clin Transl Hepatol*. 2021;9(1):71-80.
157. Ochoa-Allemant P, Marrero JA, Serper M. Racial and ethnic differences and the role of unfavorable social determinants of health across steatotic liver disease subtypes in the United States. *Hepatol Commun*. 2023;7(12).
158. Willett W, Willett W. Overview of Nutritional Epidemiology. *Nutritional Epidemiology*: Oxford University Press; 2012. p. 1-16.
159. Zhang M, Sun X, Zhu X, Zheng L, Bi Y, Li Q, et al. Association between fast eating speed and metabolic dysfunction-associated steatotic liver disease: a multicenter cross-sectional study and meta-analysis. *Nutr Diabetes*. 2024;14(1):61.
160. Samarasinghe SM, Hewage AS, Siriwardana RC, Tennekoon KH, Niriella MA, De Silva S. Genetic and metabolic aspects of non-alcoholic fatty liver disease (NAFLD) pathogenicity. *Egyptian Journal of Medical Human Genetics*. 2023;24(1):53.
161. Winters-van Eekelen E, Verkouter I, Peters HP, Alsema M, de Roos BG, Schrauwen-Hinderling VB, et al. Effects of dietary macronutrients on liver fat content in adults: a systematic review and meta-analysis of randomized controlled trials. *European Journal of Clinical Nutrition*. 2021;75(4):588-601.
162. Chalasani N, Younossi Z, Lavine JE, Charlton M, Cusi K, Rinella M, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. *Hepatology*. 2018;67(1).

## 7. Appendix

### i. Ancillary Figures

#### Linearity assumption diagnostics

The plots displayed have checked for the assumptions of linearity by studying the distribution and mean of residuals and homoscedasticity.

UK Biobank

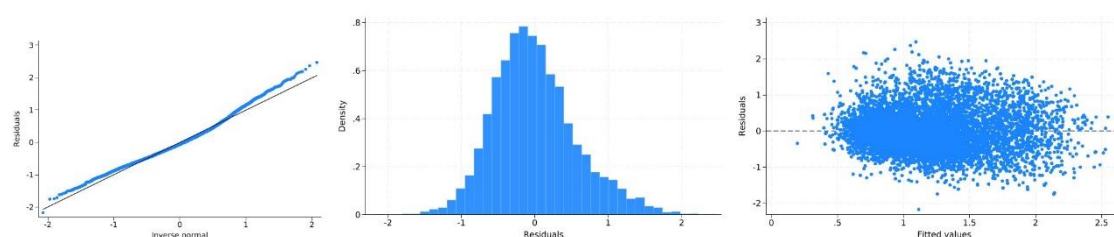


Fig A1 Linearity assumption diagnostics for Plasma SFAs%.

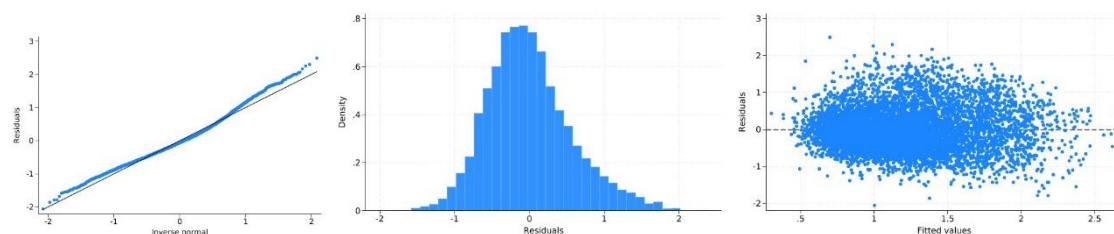


Fig A2 Linearity assumption diagnostics for Plasma MUFAs%.

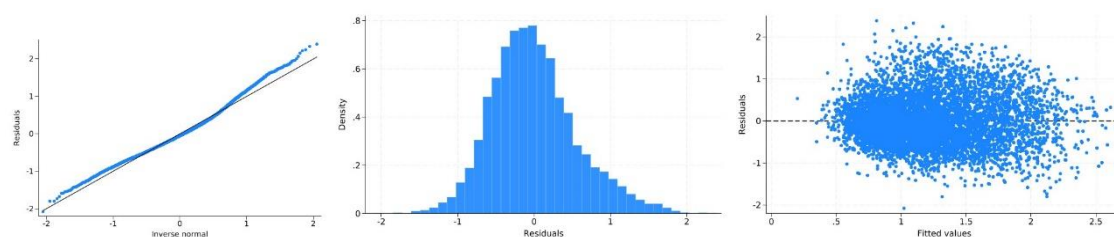


Fig A3 Linearity assumption diagnostics for PUFAs%.

## OCDEM

### *Plasma fatty acids*

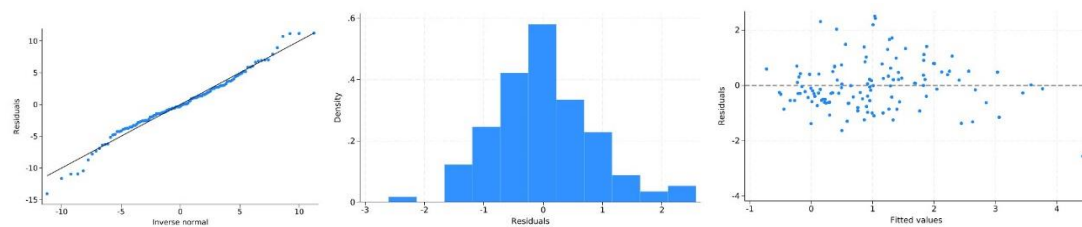


Fig A4 Linearity assumption diagnostics for Plasma SFAs%

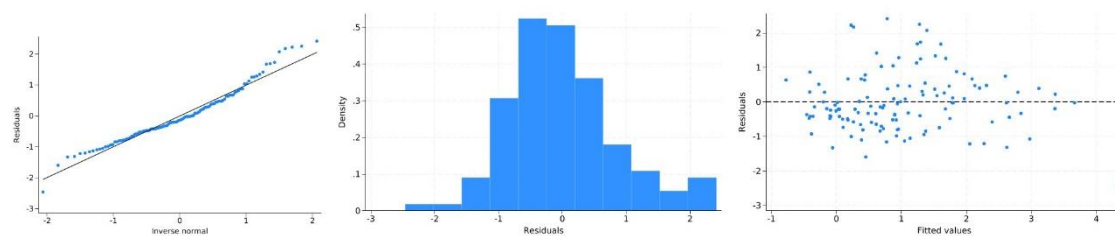


Fig A5 Linearity assumption diagnostics for Plasma PUFAs%.

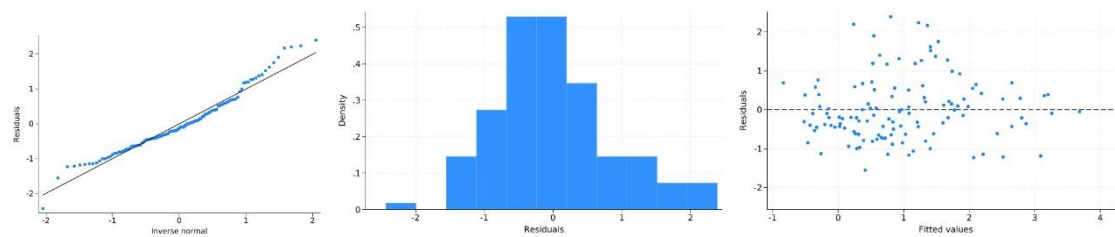


Fig A6 Linearity assumption diagnostics for Plasma MUFAs%.

## *VLDL TG Fatty acids composition*

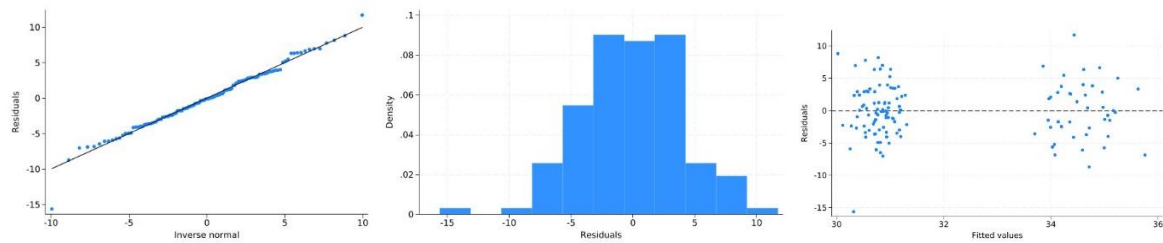


Fig A7 Linearity assumption diagnostics for VLDL-TG SFAs%.

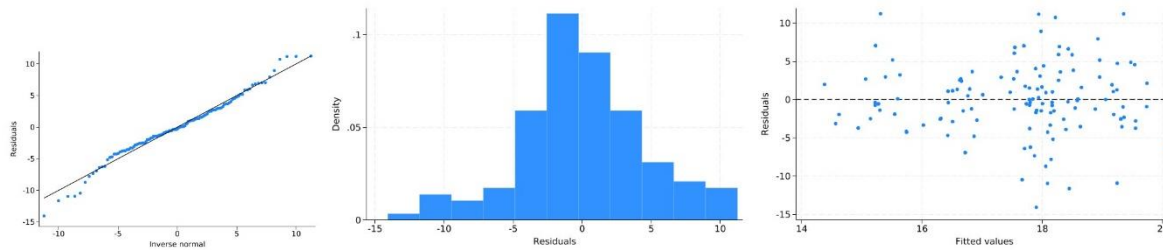


Fig. A8. Linearity assumption diagnostics for VLDL-TG PUFAs%

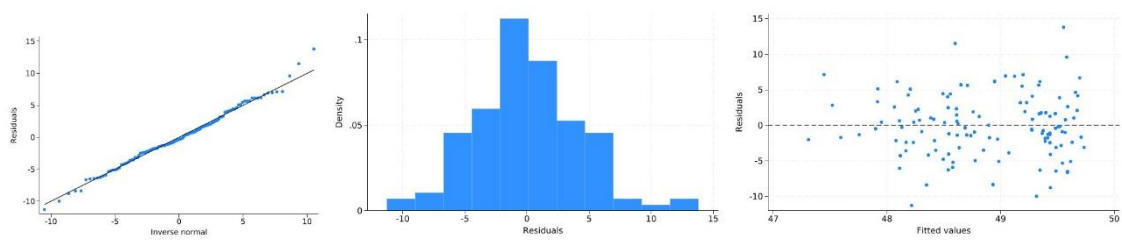


Fig. A9 Linearity assumption diagnostics for VLDL-TG MUFAs%

### *Individual fatty acids in VLDL-TG*

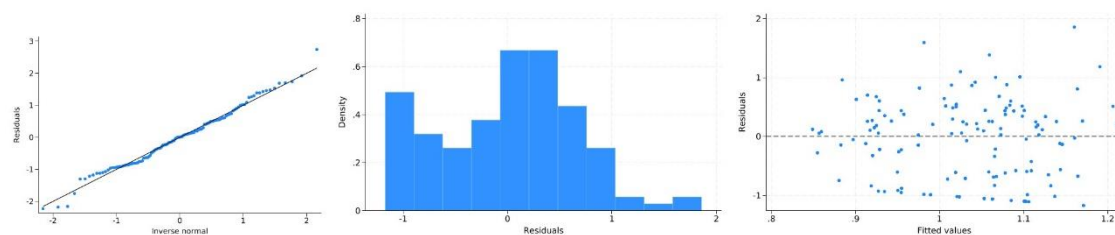


Fig. A10 Linearity assumption diagnostics for VLDL-TG Alpha linoleic %

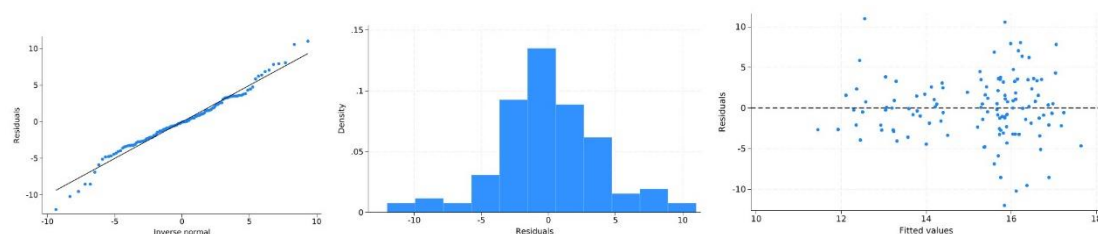


Fig. A11 Linearity assumption diagnostics for VLDL-TG Linoleic %

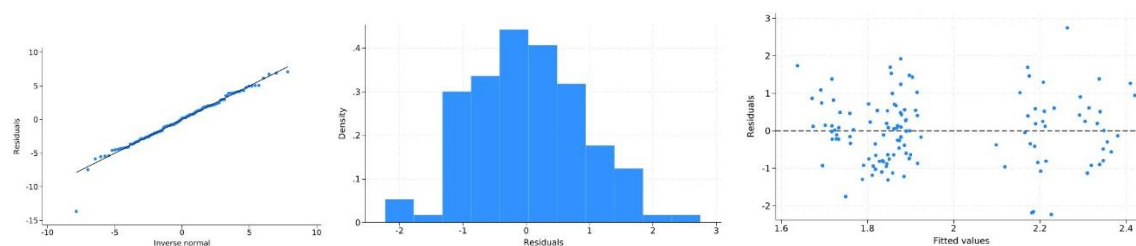


Fig. A12 Linearity assumption diagnostics for VLDL-TG myristic %

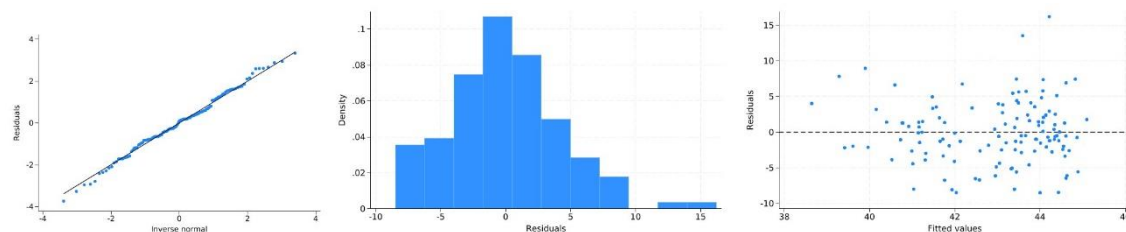


Fig. A13 Linearity assumption diagnostics for VLDL-TG oleic %

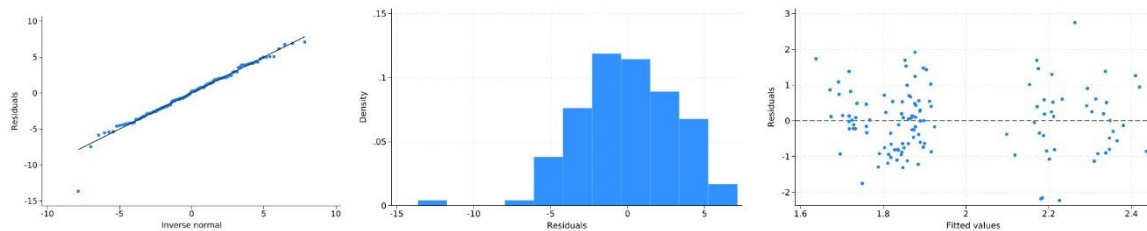


Fig. A14 Linearity assumption diagnostics for VLDL-TG palmitic%

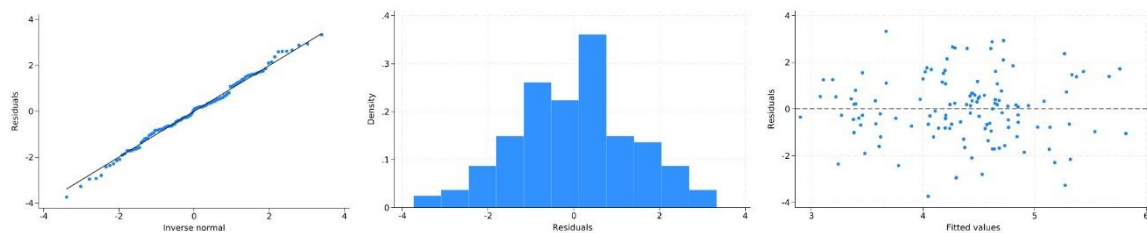


Fig. A15 Linearity assumption diagnostics for VLDL-TG palmitoleic%

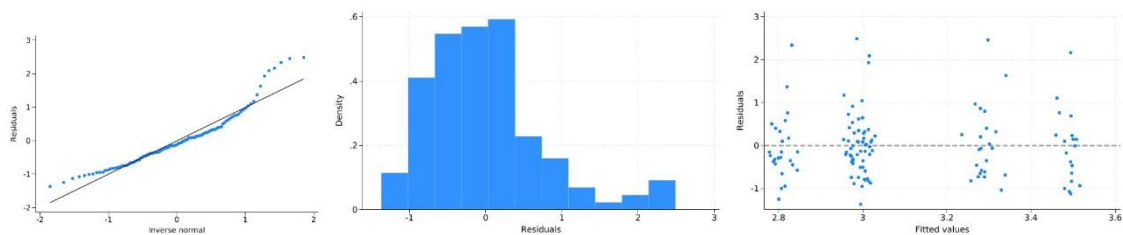


Fig. A16 Linearity assumption diagnostics for VLDL-TG stearic%

## Markers of cardiovascular risk

### Total cholesterol

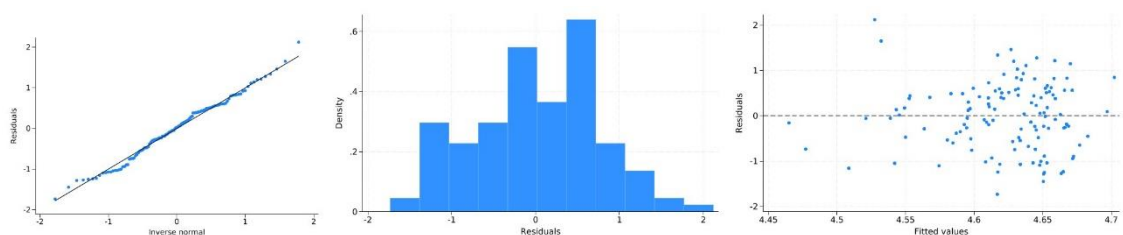


Fig. A17 Linearity assumption diagnostics for Total cholesterol

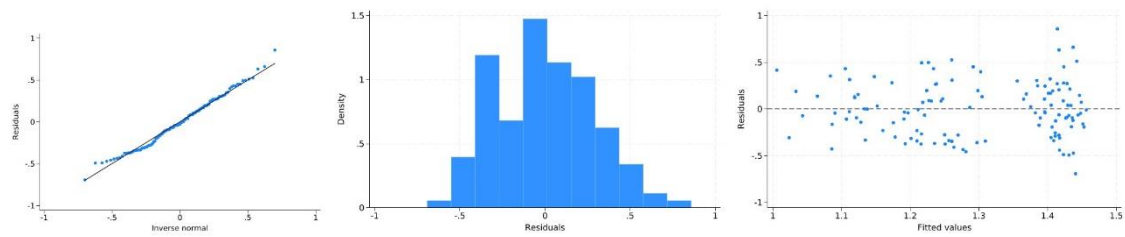


Fig. A18 Linearity assumption diagnostics for HDL Cholesterol

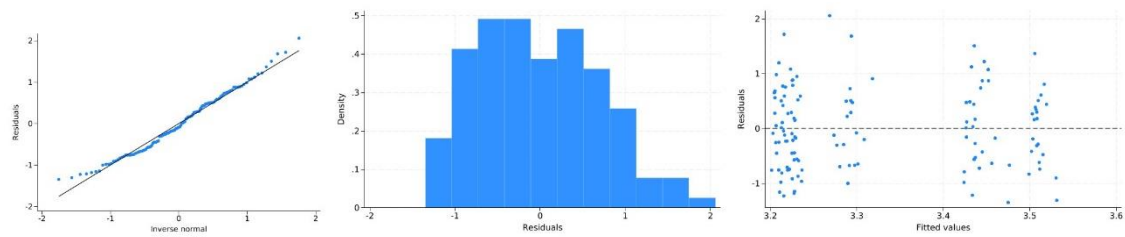


Fig. A19 Linearity assumption diagnostics for Non HDL Cholesterol

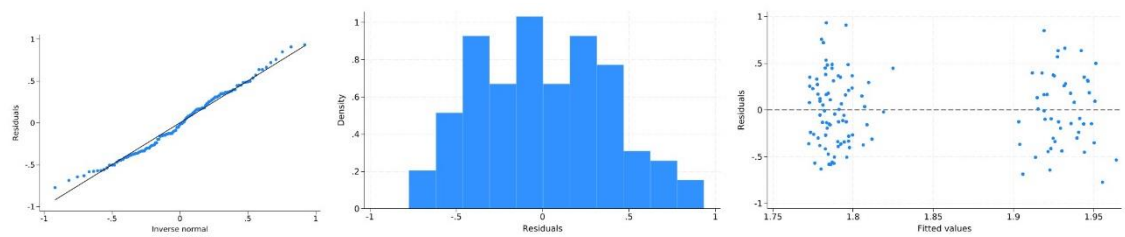


Fig. A20 Linearity assumption diagnostics for LDL Cholesterol

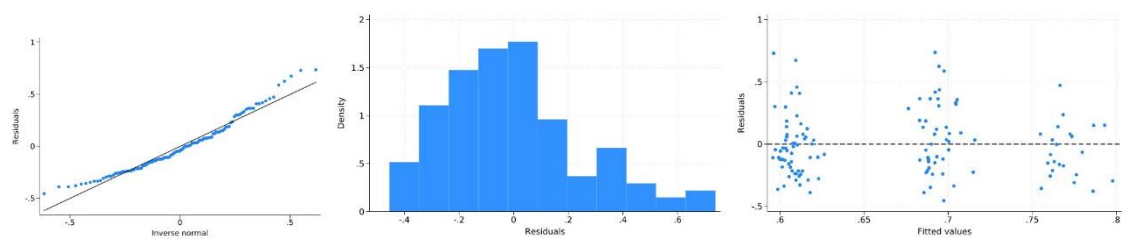


Fig. A21 Linearity assumption diagnostics for VLDL Cholesterol