Role of Fatty Acid Composition in Non-Alcoholic Fatty Liver Disease: a Dietary Perspective

Results from Interventional and Observational Studies

MICHAEL FRIDÉN
Dissertation presented at Uppsala University to be publicly examined in Sal IV, Universitetshuset, Biskopsgatan 3, Uppsala, Thursday, 2 May 2024 at 12:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in Swedish. Faculty examiner: Universitetslektor Helen Lindqvist (Department of Internal Medicine and Clinical Nutrition, University of Gothenburg).

Abstract

The overall aim of this doctoral thesis was to investigate the role of circulating, liver and dietary fatty acids in non-alcoholic fatty liver disease (NAFLD). For circulating and liver fatty acids, special emphasis was given to fatty acids reflecting diet.

In paper I, circulating cholesteryl ester (CE) linoleic acid (18:2n-6), which is considered a good biomarker of dietary intake of 18:2n-6, was cross-sectionally inversely associated with liver fat in n=308 50-year old men and women. Several fatty acids reflecting both exogenous intake and endogenous metabolism were associated with liver fat, basal fat oxidation and resting energy expenditure (REE). No association between fatty acids and liver fat, except for docosahexaenoic acid (22:6n-3) and liver fat, were attenuated after adjusting for REE.

In paper II, phospholipid (PL) 22:6n-3 in liver tissue, a potential biomarker of dietary intake of 22:6n-3, was cross-sectionally inversely associated with liver fibrosis in n=60 men and women with biopsy-verified NAFLD. This finding was not replicated in plasma. Several other fatty acids reflecting both exogenous intake and endogenous metabolism were associated with fibrosis. Pooled saturated fatty acids (SFA) were generally positively associated whereas monounsaturated (MUFA) and polyunsaturated fatty acids (PUFA) were inversely associated with fibrosis.

In paper III, no clear (i.e. imprecise) associations were observed for any of the nutrient or food substitutions with incidence NAFLD cirrhosis or hepatocellular carcinoma (HCC), over a median follow-up of 24 years in n>77 000 middle-aged to elderly men and women.

In paper IV, a 12-month randomized controlled trial (RCT) was conducted to investigate the effects of a low-carbohydrate high PUFA (LCPUFA) diet and a healthy Nordic diet (HND) on liver fat in men and women with type 2 diabetes (T2D) or prediabetes. The comparator diet (usual care (UC)) aligned with the Nordic Nutrition Recommendations. Liver fat decreased more in the LCPUFA diet and the HND versus UC. No difference in liver fat was observed between LCPUFA and HND. The LCPUFA diet and the HND improved several other cardiometabolic markers compared to UC, with more favorable improvements in the HND group.

In conclusion, findings from this thesis suggest that higher intakes of dietary unsaturated fatty acids (in particular PUFA) and lower intakes of SFA may be of importance for the prevention and treatment of NAFLD (at least for liver fat and fibrosis). Findings from this thesis also suggest that fatty acids reflecting both diet and endogenous metabolism may play a role in NAFLD.

Keywords: Fatty acids, Diet, Biomarkers, NAFLD, Liver fat, Fibrosis, Cirrhosis, HCC

Michael Fridén, Department of Public Health and Caring Sciences, Clinical Nutrition and Metabolism, Uppsala Science Park, Uppsala University, SE-75185 Uppsala, Sweden.

© Michael Fridén 2024

ISSN 1651-6206
URN urn:nbn:se:uu:diva-524500 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-524500)
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from the respective publishers.
The contribution of Michael Fridén to the papers included in this thesis was as follows:

I. Designed and planned the study in collaboration with supervisors, compiled all FFQs, analysed the data, wrote the first draft of the manuscript and revised the paper in collaboration with supervisors and co-authors.

II. Designed and planned the study in collaboration with supervisors and co-authors, collected a fraction of the liver biopsies for storage, compiled and analysed the data, wrote the first draft of the manuscript and revised the paper in collaboration with supervisors and co-authors.

III. Designed and planned the study in collaboration with supervisors and co-authors, wrote the plan for the Swedish Ethical Review Authority, analysed the data and wrote the first draft of the manuscript.

IV. Registered the study in Clinical Trials, wrote the statistical analysis plan in collaboration with supervisors and co-authors, wrote the first draft for the amendment to the Swedish Ethical Review Authority, coordinated all parts of the study, extracted a fraction of the plasma fatty acids, compiled and analysed the data and wrote the first draft of the manuscript.
# Contents

## Introduction

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
</tr>
</tbody>
</table>

## Background

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
</tr>
</tbody>
</table>

- MASLD: a change in nomenclature ......................................................... 8
- NAFLD ..................................................................................................... 8
- T2D and NAFLD .................................................................................... 9
- Assessment of NAFLD ........................................................................ 10
- Deposition of intrahepatic fat ............................................................. 11
- Fatty acids .......................................................................................... 13
- Fatty acids as biomarkers of dietary intake ............................................ 16
- Circulating and liver fatty acids and NAFLD ........................................ 17
- Dietary fatty acids and NAFLD ............................................................ 18
- Dietary patterns and NAFLD ................................................................. 20
- Research gaps in the field of fatty acids and NAFLD .......................... 21

## Aims

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
</tr>
</tbody>
</table>

- Specific aims .................................................................................... 22

## Methods

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
</tr>
</tbody>
</table>

- Study descriptions .............................................................................. 23
- Paper I (POEM) ................................................................................. 23
- Paper II (AM-02 NASH) .................................................................. 23
- Paper III (SMC and COSM) ............................................................... 24
- Paper IV (NAFLDiet) ........................................................................ 24
- Ethics ................................................................................................ 27
- Assessment of exposure ..................................................................... 27
- Assessment of outcomes ................................................................. 30
- Assessment of confounders .............................................................. 31
- Substitution modelling ..................................................................... 35
- Statistical analyses .......................................................................... 37

## Main results

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
</tr>
</tbody>
</table>

- Paper I ............................................................................................. 39
- Paper II ............................................................................................ 40
- Paper III ........................................................................................... 42
- Paper IV ........................................................................................... 44
Non-alcoholic fatty liver disease (NAFLD) is characterized by excess fat in the liver, affecting one third of the global population [1]. In 20-40 % of patients, the steatotic liver progresses to non-alcoholic steatohepatitis (NASH), further characterized by inflammation with or without accompanying fibrosis [2–5]. Cirrhosis (i.e. the advanced stage of fibrosis) due to NAFLD is now the fastest growing indication for liver transplantation globally [6]. Cirrhosis also increases the risk of hepatocellular carcinoma (HCC) [1]. Diet is a cornerstone in the prevention and treatment of NAFLD and short-term randomized controlled trials (RCT) have shown that saturated fatty acids (SFA) increase liver fat when compared to polyunsaturated fatty acids (PUFA) [7–9]. However, few long-term studies have been conducted to investigate the role of dietary fat quality on liver fat, cirrhosis and HCC. Furthermore, whether and how circulating and tissue fatty acids (derived both endogenously and from diet) may be implicated in the pathophysiology of liver fat accumulation and fibrosis development in NAFLD is still unclear. The latter approach to investigate fatty acid biomarkers as reflective of dietary intake may circumvent potential issues derived from measurement error of self-reported dietary data. This thesis aims to investigate the role of fatty acids, in particular dietary fatty acids, in NAFLD.
Background

MASLD: a change in nomenclature

In the summer of 2023, a multi-society Delphi consensus statement was released proposing a change in nomenclature from NAFLD to Metabolic-dysfunction Associated Steatotic Liver Disease (MASLD) [10]. The minor modifications to the disease criteria were the addition of any cardiometabolic risk factors in the presence of steatosis to be defined as MASLD. Nearly 100% of individuals with NAFLD also meet the criteria for MASLD in Sweden [11]. The change in nomenclature occurred after the planning phase for each paper included in this thesis. To stay true to the initial plan, we chose to keep the definition and nomenclature of NAFLD throughout the manuscript.

NAFLD

NAFLD is a spectrum of diseases ranging from simple steatosis (NAFL) to NASH which may progress to fibrosis, cirrhosis and HCC [1]. NAFLD is defined by the presence of excess liver fat in the absence of any competing risk factors for fatty liver (e.g. significant alcohol consumption ≥ 20 grams/day in women and ≥ 30 grams/day in men or viral hepatitis) [12]. Excess liver fat is in turn defined by >5.6% using the proton density fat fraction (PDFF) from magnetic resonance imaging (MRI) or >5% of hepatocytes containing macrovesicular lipid droplets using histological grading. The global prevalence of NAFLD is estimated to be 32-38% and may reach 70% in individuals with type 2 diabetes (T2D) [13,14]. The more inflammatory state of the disease spectrum, NASH, is grossly estimated to affect 1.5-6.45% of the general population and 37% of individuals with T2D [13,15,16]. NAFLD is strongly associated with incident cardiovascular disease (CVD) [17], T2D [18], severe liver disease [19], extrahepatic cancers [20] and all-cause mortality [19,21].

NAFLD is a highly heterogeneous disease with multiple underlying pathophysiological mechanisms involved. As such, certain genetic polymorphisms (e.g. patatin-like phospholipase domain-containing protein 3 (PNPLA3) (I148M) and transmembrane 6 superfamily 2 (TM6SF2) (E167K)) are now recognized to be strongly associated with both simple steatosis and more progressive forms of NAFLD, but not with insulin resistance, hyperglycaemia or hyperlipidaemia [22,23].
As simple steatosis progresses to NASH, through mechanisms that are not yet fully known, the liver is exposed to inflammatory mediators eventually causing liver injury. As a natural wound-healing response to this damage, liver cells called hepatic stellate cells (HSC) are activated, initiating the process of fibrogenesis, or scarring. If sustained, chronic fibrogenesis causes an excess of extracellular matrix, which is deposited in the liver, leading to fibrosis and cirrhosis [24]. Interestingly, Hagström et al. demonstrated that the degree of liver fibrosis in NAFLD was the strongest predictor of severe liver disease and all-cause mortality over a mean follow-up of 20 years [19]. Elucidating the mechanisms by which the liver in patients with NAFLD progresses to more severe stages of fibrosis is thus highly relevant.

Cirrhosis due to NAFLD is a rapidly growing aetiology. Around 20-40 % of people with NAFL will develop NASH, of which 20 % will develop cirrhosis [2–5]. However, the gradual development from simple steatosis to cirrhosis may take decades to develop and may go unnoticed in a large proportion of patients with the disease. Compensated cirrhosis is the state by which the liver is able to function properly despite severe scarring whereas decompensated cirrhosis is when the liver starts to deteriorate, causing complications such as portal hypertension, ascites, varices bleedings and hepatic encephalopathy [25].

Although 20 % of HCC cases from NAFLD do not present any clinical manifestation of cirrhosis, fibrotic scarring remains a strong risk factor for HCC [1]. HCC constitutes around 90 % of primary liver cancers and is mainly caused by viral hepatitis, alcohol overconsumption and NAFLD [1,26]. The age-standardized incidence rate of HCC in Sweden was estimated to be 7.77/100 000 individuals in 2019 [27], with NAFLD being one of the leading causes [28].

**T2D and NAFLD**

NAFLD is present in about 55-65% of patients with T2D, with prevalence being somewhat higher in European countries [13,16]. In a meta-analysis from 2021, Mantovani et al. demonstrated that pooling studies of biopsy-confirmed NAFLD and/or NAFLD diagnosed by imaging was associated with an approximate doubling in the risk of T2D over a median follow-up of 5 years, compared to those not having NAFLD [29]. Interestingly, the risk of T2D increased as the severity of hepatic steatosis (using ultrasonography) or fibrosis increased [29]. The former has been replicated in a Swedish cohort of 106 patients with NAFLD where steatosis (using stereological point counting) from liver biopsies predicted the risk of T2D incidence over 23 years of follow-up [30]. Using repeated measures of liver biopsies from the same study showed that a reduction in steatosis between the two time points was associated with a lower risk of T2D over follow-up [30]. As the severity of NAFLD...
parallels the risk of T2D. T2D may further exacerbate the progression from simple steatosis to fibrosis [31], thereby suggesting an intricate relationship between T2D and NAFLD severity [32].

Assessment of NAFLD

Hepatic fat deposition can be assessed using different fatty liver indices (e.g. the fatty liver index (FLI) or the hepatic steatosis index (HSI)), through liver biopsies or by non-invasive imaging techniques (e.g. ultrasound, computed tomography (CT), magnetic resonance spectroscopy (MRS) or MRI). MRI is a common imaging modality used in randomized trials for the quantification of steatosis and has been shown to correlate strongly with both MRS and liver biopsies [33–37]. MRI, as MRS, uses the PDFF method to estimate the amount of liver fat in a given cross-sectional area of the liver. In a recent meta-analysis of almost 600 individuals, the pooled correlation coefficient from nine different studies on PDFF-assessed liver fat (using MRI and MRS) versus histology was 0.85 [37]. However, as the amount of liver fat increased using histology-based quantification, the discrepancy between PDFF and histology increased as well. This finding may be explained by the fact that PDFF and liver histology quantifies liver fat based on different fundamental principles. A pathologist determines the degree of steatosis based on the number of macrovesicular lipid droplet-containing hepatocytes, whereas PDFF assesses proton densities in fat versus fat and water in a cross-sectional area of the entire liver. As the liver contains other cell types than hepatocytes (as well as extracellular matrix), the PDFF-method will inherently underestimate the amount of liver fat compared to liver biopsies [37]. This is of importance to remember when interpreting findings from studies using different methodologies to quantify liver fat. Although highly correlated with other methods, MRI is however expensive and time-consuming, limiting its use in clinical practice.

Despite limitations of sampling variability [38], liver biopsy remains the gold standard method for discriminating simple steatosis from NASH. Liver biopsy is further considered the most reliable method for staging the severity of fibrosis in NAFLD. Staging usually range from 0 (no fibrosis) to 4 (cirrhosis). Stage 2 and above is referred to clinically significant fibrosis and stage 3 and above is commonly referred to advanced fibrosis. However, liver biopsies are not without complications. Sensation of pain occurs in up to 84 % of patients whereas bleeding occurs in 1 in 500 biopsies with severe bleeding occurring in 1 in 2500 to 1 in 10 000 biopsies [39]. As such, other biomarkers of liver fibrosis have been extensively evaluated of which the FIB-4 index and transient elastography (TE) are commonly used in clinical practice and research. FIB-4 is calculated based on values on platelet counts, alanine ami-
notransferase (ALT) and aspartate aminotransferase (AST) from blood samples as well as age and is a relatively cheap first step in reliably excluding advanced fibrosis. If advanced fibrosis cannot be ruled out, TE usually works as an additional complement to measure liver stiffness.

Deposition of intrahepatic fat

In 2005, Donnelly et al. demonstrated that 59 % of the intrahepatic triglyceride (IHTG) content in patients with NAFLD was derived from non-esterified fatty acids (NEFA) from adipose tissue lipolysis, while 26.1 % and 14.9 % was derived from de novo lipogenesis (DNL) and diet, respectively [40]. This was eloquently shown using four days of stable isotope infusion coupled with liver biopsy. Although the relative proportion of DNL to total IHTG is higher in obese NAFLD subjects than in obese controls, plasma NEFA is still the major contributor to IHTG in both phenotypes [41]. This is also evident after a meal despite the postprandial inhibition of insulin-mediated suppression of adipose tissue lipolysis [42]. The three main pathways of which IHTG accumulation may occur is depicted in Figure 1 below.
Figure 1. Main pathways for the accumulation of fat in the liver. Intrahepatic triglyceride (IHTG) deposition is the sum of 1) fatty acids brought to the liver by lipolysis and/or lipoprotein lipase mediated spillover 2) de novo synthesized fatty acids from carbohydrates (and protein) and 3) fatty acids delivered from the food as well as the disposal of fatty acids from the liver via very low density lipoprotein (VLDL)-particle secretion and/or through complete oxidation (forming carbon dioxide) or ketogenesis (forming ketone bodies). The figure contains additional information on the metabolism of fatty acids and the complex interplay between fatty acids in both liver tissue and circulating lipoproteins. Some of the potential mechanisms by which fatty acids and their derivatives influence IHTG deposition, inflammation and fibrosis are also illustrated. ACC, Acetyl-CoA Carboxylase; CE, Cholesteryl Esters; ChREBP, Carbohydrate-Responsive Element-Binding Protein; CPT, Carnitine Palmitoyltransferase; DAG, Diacylglycerols; D5D, Delta-5 Desaturase; D6D, Delta-6 Desaturase; FAS, Fatty Acid Synthase; PL, Phospholipids; PPAR-α, Peroxisome Proliferator-Activated Receptor α; SCD-1, Stearoyl-CoA Desaturase 1; SREBP-1c, Sterol Regulatory Element-Binding Protein 1c; TAG, Triacylglycerols.

One should however bear in mind that the contribution of the different sources of fatty acids to the total IHTG pool is not only dependent on the metabolic phenotype of the individual. The composition of the diet is also an important contributor and has been thoroughly discussed in a recently published review article [43]. For example, individuals with NAFLD are characterized by two to three fold higher DNL rates compared to individuals without NAFLD [41]. This might be partially explained by the close association between NAFLD and whole-body insulin resistance whereby excess circulating insulin in combination with a selective hepatic insulin resistance exacerbates the DNL process, although the latter pathway-specific hypothesis of insulin resistance has been questioned lately [44]. However, it might also be due to the fact that individuals with NAFLD consume a diet richer in carbohydrates and more lipogenic simple sugars (e.g. fructose) [45], as has been reported in multiple observational studies [43]. Compared to fat, a diet with a higher proportion of
carbohydrates has been shown to increase DNL over days to weeks of follow-up [46–48]. In addition, within the carbohydrate category, fructose, as compared to glucose has been shown to stimulate DNL more in several controlled feeding trials [49,50]. Noteworthy, the impact on DNL from fructose versus glucose has not been shown to consistently translate to changes in liver fat between groups [43], except in a short-term study by Lecoultre et al., comparing 3g/kg of additional fructose vs glucose in a hypercaloric setting using sugar-sweetened soda as the vehicle [51]. Although interesting, the extent to which these findings may be translated to a real-world setting is questionable, as fructose is not found in isolation and the amount consumed in the trials is far above what is typically ingested for the average individual (3g/kg equals ~1000 kcal from fructose alone in an individual weighing 80 kg). The impact of fatty acids on DNL and liver fat will be discussed further below.

Fatty acids

Fatty acids are lipid molecules of varying lengths of carbon chains with a carboxyl-group (-COOH) at one end of the chain and a methyl-group (-CH3) at the other. Fatty acids may differ in both their length and saturation, i.e. the degree to which all carbon atoms are saturated with hydrogen atoms. The degree of saturation of a fatty acid is divided into three main categories: saturated (e.g. palmitic acid (16:0) where all carbon atoms are saturated and no double bonds exist), monounsaturated (e.g. oleic acid (18:1n-9) where all but two carbon atoms are saturated and one double bond between two carbons exists) and polyunsaturated (e.g. linoleic acid (18:2n-6) where many carbon atoms remain to be saturated and two or more double bonds exist). The unsaturated fatty acids are further classified into n-7 and n-9 for monounsaturated fatty acids (MUFA) and n-3 (omega-3) and n-6 (omega-6) for PUFA. Of these, two are essential and thus required from the diet: 18:2n-6 and α-linolenic acid (18:3n-3). In addition, fatty acids may be presented in either cis- or trans-configuration, meaning that the two hydrogen atoms of an unsaturated fatty acid attached to the two carbons with a double bond between them are either on the same side (cis) or on the opposite sides (trans) of each other. As trans-fatty acids present as straight chains of carbon atoms they resemble SFA more so than cis-unsaturated fatty acids that have a more asymmetrical shape.

The nomenclature of a fatty acid derives from the number of carbon atoms, the number of double bonds and at what carbon position from the methyl-group end the first double bond is located. For example, linoleic acid is a PUFA denoted 18:2n-6. The fatty acid contains 18 carbons with 2 double bonds of which the first double bond is positioned (n) at the 6th carbon from the methyl-group end, making it into an omega-6 fatty acid. Both trivial names and the nomenclature system will be used to describe fatty acids in this thesis, with an emphasis on the nomenclature system.
Fatty acids may undergo both elongation and desaturation, meaning that fatty acids can become longer and more unsaturated. Multiple enzymes are involved in this process. Please see Figure 1 for a summary of the main enzymes discussed in this thesis. Stearoyl-CoA desaturase 1 (SCD-1) is responsible for desaturating 16:0 to palmitoleic acid (16:1n-7) or stearic acid (18:0) to 18:1n-9. Delta-5 desaturase (D5D) is responsible for desaturating dihomo-\(\gamma\)-linolenic acid (20:3n-6) to arachidonic acid (20:4n-6) (or eicosatetraenoic acid (20:4n-3) to eicosapentaenoic acid (20:5n-3) for omega-3 fatty acids). Delta-6 desaturase (D6D) is responsible for desaturating 18:2n-6 to \(\gamma\)-linolenic acid (18:3n-6) (or 18:3n-3 to stearodonic acid (18:4n-3) for omega-3 fatty acids). In addition to these desaturase enzymes, several elongases are involved in adding carbons to fatty acid chains (e.g. elongase 6 is responsible for elongating 16:0 to 18:0). These product-to-precursor fatty acid ratios are commonly used to estimate the activity of the enzymes (e.g. 16:1n-7/16:0 is used to estimate SCD-1 activity).

Table 1. A summary of fatty acids discussed in this thesis.

<table>
<thead>
<tr>
<th>Trivial name</th>
<th>Nomenclature</th>
<th>Fatty acid class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myristic acid</td>
<td>14:0</td>
<td>SFA</td>
</tr>
<tr>
<td>Pentadecaenoic acid</td>
<td>15:0</td>
<td>SFA</td>
</tr>
<tr>
<td>Palmitic acid</td>
<td>16:0</td>
<td>SFA</td>
</tr>
<tr>
<td>Heptadecaenoic acid</td>
<td>17:0</td>
<td>SFA</td>
</tr>
<tr>
<td>Stearic acid</td>
<td>18:0</td>
<td>SFA</td>
</tr>
<tr>
<td>Arachidic acid</td>
<td>20:0</td>
<td>SFA</td>
</tr>
<tr>
<td><strong>Behenic acid</strong></td>
<td><strong>22:0</strong></td>
<td><strong>SFA</strong></td>
</tr>
<tr>
<td>Lignoceric acid</td>
<td>24:0</td>
<td>SFA</td>
</tr>
<tr>
<td>Palmitoleic acid</td>
<td>16:1n-7</td>
<td>MUFA</td>
</tr>
<tr>
<td>Vaccenic acid</td>
<td>18:1n-7</td>
<td>MUFA</td>
</tr>
<tr>
<td><strong>Oleic acid</strong></td>
<td><strong>18:1n-9</strong></td>
<td><strong>MUFA</strong></td>
</tr>
<tr>
<td>(\alpha)-linolenic acid</td>
<td>18:3n-3</td>
<td>Omega-3 PUFA</td>
</tr>
<tr>
<td>Eicosapentaenoic acid</td>
<td>20:5n-3</td>
<td>Omega-3 PUFA</td>
</tr>
<tr>
<td>Docosapentaenoic acid</td>
<td>22:5n-3</td>
<td>Omega-3 PUFA</td>
</tr>
<tr>
<td><strong>Docosahexaenoic acid</strong></td>
<td><strong>22:6n-3</strong></td>
<td><strong>Omega-3 PUFA</strong></td>
</tr>
<tr>
<td>Linoleic acid</td>
<td>18:2n-6</td>
<td>Omega-6 PUFA</td>
</tr>
</tbody>
</table>
MUFA, Monounsaturated Fatty Acids; PUFA, Polyunsaturated Fatty Acids; SFA, Saturated Fatty Acids. The four highlighted fatty acids will be emphasized in the thesis and readers may want to remember these four in particular.

Fatty acids are further incorporated into other lipid moieties such as phospholipids (PL), cholesteryl esters (CE) and triacylglycerols (TAG). PL are glycerophospholipids with a glycerol backbone attached to two fatty acids and a phosphorous group. PL are important cell membrane lipids by which phosphatidylcholine is the most abundant. TAG are glycerolipids with a glycerol backbone attached to three fatty acids. CE are sterol lipids with a cholesterol molecule attached to a single fatty acid. In the diet, 90-95% of all fatty acids are incorporated into TAG. The most abundant dietary fatty acids are 16:0, 18:1n-9 and 18:2n-6. The proportions of these fatty acids, along with all other fatty acids, differ depending on the food source. In addition, proportions of fatty acids also differ between compartments (e.g. liver tissue vs plasma/serum) and lipid fractions (e.g. PL, CE and TAG). Figure 2 below will help to disentangle this.

**Figure 2**. Proportions (% of total fatty acids) of palmitic acid (16:0), oleic acid (18:1n-9) and linoleic acid (18:2n-6) in different foods and over multiple compartments and lipid fractions. The proportions are approximate values gathered from the literature. CE, Cholesteryl Esters; PL, Phospholipids; TAG, Triacylglycerols.
Fatty acids as biomarkers of dietary intake

Self-reported dietary intake is prone to misreporting, caused by various factors such as social desirability, wrongly estimated portion sizes, limitations to the food composition database and participation burden. Underreporting of diet is a common feature in both observational and interventional studies and is generally more pronounced in certain groups of individuals (people with higher body mass index (BMI)) and for certain foods and nutrients (e.g. snacks and added sugar) [52,53]. This may introduce differential measurement error [52,53], leading to biased estimates of the causal relationship between diet and disease. Biomarkers of dietary intake have therefore been developed to potentially tackle some of these issues.

18:2n-6 is an essential fatty acid that is not synthesized in the human body and is therefore considered a good biomarker of dietary intake [7,8,54]. Using data from controlled dietary intervention trials, Zock, et al. calculated that for each 10 percent increase of dietary 18:2n-6 (in place of other fatty acids), 18:2n-6 in circulating CE increased by 9.3 g/100 g fatty acids [55]. However, predictive equations such as these have not been used to any large extent in the literature. 18:3n-3 is also considered essential, albeit not as strongly correlated with dietary intake as 18:2n-6 [56]. Although essential, enzymes responsible for the conversion of these fatty acids to non-essential fatty acids are regulated by various factors such as genetic polymorphisms and sex, thereby influencing the proportion of 18:2n-6 and 18:3n-3 in plasma [54,57]. Due to the limited conversion rate from 18:3n-3 to longer omega-3 fatty acids [58], 20:5n-3 and 22:6n-3, found in relatively high amounts in fatty fish, are also considered good biomarkers of these foods [54,59,60]. Multiple observational and well controlled dose-response feeding studies have been conducted to validate these fatty acids as reflective of dietary composition [54]. On the contrary, proportions of 18:2n-6 in circulation may be influenced by the consumption of long-chain omega-3 fatty acids [61], whereas proportions of long-chain omega-3 fatty acids in plasma may be influenced by the proportion of 18:2n-6 in the diet [7,62]. In addition, the proportion of 18:2n-6 in VLDL-TAG is lowered in response to increases in DNL-derived 16:0 from carbohydrate consumption, as demonstrated by Hudgins LC, et al. in the mid 90’s [48]. In response to large quantities of liquid sugar in free-living individuals, meaningful decreases in 18:2n-6 have also been observed in lipid fractions such as PL and CE when compared to water [63]. As such, careful consideration must be applied when interpreting findings of circulating and tissue fatty acids presented in proportions.

Lastly, pentadecaenoic acid (15:0) and heptadecaenoic acid (17:0), two odd-chain SFA may be considered to reflect dairy fat intake [64]. However, these fatty acids are also found in a variety of marine water fish and may even be synthesized in smaller quantities from gut-derived propionate through the consumption of certain fibers [65,66]. Intervention studies are also somewhat
conflicting [67,68]. Warensjö, et al. showed an increase in plasma 15:0 and 17:0 after consuming a diet rich in milk fat (butter, high-fat cheese and cream) compared to a diet rich in rapeseed oil over 3 weeks of follow-up [67]. In support of these findings, Bethancourt et al. showed an increase in 15:0 in PL when comparing a group consuming 3.3 servings of high-fat dairy/day with a group consuming <3 servings of non-fat milk/week, but no statistically significant differences in 15:0 were observed between the high-fat dairy group and the low-fat dairy group. No differences in 17:0 were observed [68,69]. Interestingly, in multivariable linear regression analyses, adjusting for liver fat content at follow-up attenuated the associations between dairy fat and 15:0 and 17:0, potentially suggesting that changes in liver fat may impact the validity of these biomarkers [68].

Not all fatty acids are good candidates for the use as biomarkers of dietary intake. Although present in large amounts in various foods (e.g. milk, butter, meat, olive oil and rapeseed oil), fatty acids such as 16:0, 18:0 and 18:1n-9 are also endogenously synthesised through DNL, elongation and desaturation processes (Figure 1). 16:1n-7, the desaturation product of 16:0 by SCD-1 is only present in small amounts in foods (e.g. macadamia oil and sea buckthorn oil), and may therefore be a suggestive candidate marker carbohydrate induced DNL [70], although evidence is not conclusive [71]. Increased plasma levels of 16:1n-7 may also reflect desaturation of 16:0 or de novo synthesis of ceramide species from both dietary and DNL-derived 16:0 [8]. Lastly, careful consideration should be given to the time period for which these fatty acids are used to reflect past diet exposure. Fatty acids derived from adipose tissue may reflect longer-term intake (several months) while fatty acids derived from circulating CE, PL and TAG may reflect diet in the past days to weeks [54].

Circulating and liver fatty acids and NAFLD

Circulating fatty acids reflecting dietary intake have been shown to associate with liver fat content in both cross-sectional [72–75] and prospective studies [73,74], using MRI, ultrasonography and different indices to assess liver fat and NAFLD. 18:2n-6 and 22:6n-3 have consistently been shown to be inversely associated with both liver fat and NAFLD [73,75,76]. Fewer studies have examined the relationships between circulating fatty acids and fibrosis in NAFLD.

Patients with NAFLD are characterized with lower concentrations of liver unsaturated fatty acids and higher concentrations of liver SFA [77–79]. Limited evidence exist for liver FA in relation to NAFLD fibrosis and cirrhosis. A lower proportion of 22:6n-3 in livers of patients with NAFLD has been repeatedly shown [77,78]. Whether or not exogenous and/or endogenous 22:6n-3 is implicated in the pathophysiology of NAFLD remain elusive, but several plausible mechanisms have been proposed that may link this omega-3 fatty
acid to liver fat accumulation and fibrosis. In 2020, Green et al. showed a reduction in DNL and an increase in net fat oxidation rate after eight weeks of omega-3 supplementation when compared to placebo [80]. In vitro studies have furthermore shown 22:6n-3 to reduce HSC proliferation in a dose-response manner [81], suggesting an important role of 22:6n-3 in the prevention of cirrhosis. Interestingly, 22:6n-3 of different lipid fractions in the liver correlate moderately to strongly with different lipid fractions in plasma, suggesting that plasma 22:6n-3 may be a good proxy for liver 22:6n-3 [82].

As for the fatty acids that may not be as good biomarkers of dietary intake as 18:2n-6 and 22:6n-3, 16:1n-7 and the fatty acid ratio 16:1n-7/16:0 (reflecting SCD-1 activity) have been consistently associated with liver fat and NAFLD using both circulating fatty acids and liver fatty acids [83,84]. However, whether these fatty acids reflect dietary intakes and/or an altered metabolism remain unclear. Interestingly, as of this moment, a phase III trial (the ARMOR trial) is undertaken to investigate the effects of a partial SCD-1 inhibitor (Aramchol) on NASH and fibrosis development, with promising results from a 52-week long phase IIb trial (the ARREST trial) on liver histology and liver enzymes [85].

Dietary fatty acids and NAFLD

Isocaloric randomized trials of dietary fat quality have consistently shown that replacing SFA with PUFA (or a mixture of PUFA and MUFA) lower the content of liver fat over 3-10 weeks [7–9,86]. In 2012, Bjermo et al. showed in a 10-week randomized eucaloric trial (HEPFAT) of individuals with abdominal obesity that replacing SFA from butter with n-6 PUFA from margarine, sunflower oil and sunflower seeds decreased liver fat by approximately 1.2 % in absolute numbers [9]. Results from the HEPFAT trial was later replicated in lean individuals using a hypercaloric design over 7 weeks as well as in individuals with overweight and obesity over 8 weeks (LIPOGAIN-1 and LIPOGAIN-2, respectively). The magnitude of effect was around 0.52-1.63 % in absolute numbers between groups and intervention diets were based on muffins baked with either palm oil or sunflower oil [7,8]. An independent group of researchers from Finland further replicated these findings in 2018, but this time using a hypercaloric design of an additional 1000 kcal over 3 weeks whereby the SFA-group received foods such as coconut oil, butter and blue cheese and the PUFA/MUFA-group received foods such as olive oil, pecan nuts, pesto and butter. The magnitude of effect was around 2 % in absolute numbers between groups [86]. Attempts to disentangle the mechanisms of such weight-independent effects have elucidated multiple potential key mediators. Luukkonen et al. showed in 2018 that replacing SFA with a mixture of MUFA and PUFA decreased the rate of adipose tissue lipolysis, as assessed
using stable isotope methodology under euglycemic hyperinsulinemic conditions, thereby limiting the supply of fatty acids to the liver for TAG synthesis [86]. In addition, changes in ceramide species were observed between groups, a finding that was later replicated by Rosqvist et al. in the LIPOGAIN-2 trial [8,86]. Ceramides are a class of sphingolipids that are partly synthesized de novo by the SFA 16:0 and are implicated in the process of hepatic insulin resistance, lipogenesis, fat oxidation, hepatic fatty acid uptake, production of pro-inflammatory cytokines and HSC activation [87], all of which are involved in the pathogenesis of NAFLD. Interestingly, adjusting for the change in C16-containing ceramides attenuated the effect on liver fat between SFA and PUFA in the LIPOGAIN-2 trial [8].

PUFA are oxidized to a greater extent than SFA in the postprandial state [88]. Using isotope-labelled fatty acids, Parry et al. showed in a randomized crossover feeding study of n=24 men and women that 18:2n-6 was more rapidly oxidized than 16:0, six hours after completion of a mixed meal [88]. Greater postprandial oxidation rates from 18:2n-6 compared to other SFA have been shown previously in smaller studies [89], although not conclusively [90]. The partitioning of fatty acids to oxidative pathways instead of storage has been suggested to be a plausible mechanism behind the effect on liver fat content when comparing PUFA versus SFA. However, whether these single meal differences may translate to differences in fat oxidation during fasting or over longer time periods of time is unclear. In addition, differences in energy expenditure, partly following from differences in substrate utilization has been proposed to play a role as well. Findings from well-controlled feeding studies are however mixed. Some have failed to see a difference in substrate oxidation rates and resting energy expenditure (REE) when assessed using indirect calorimetry [86,91] or when using a surrogate marker of hepatic fat oxidation (D-3 hydroxybutyrate) [7,8], whereas others have observed differences in these measures [92,93]. Heterogeneity between studies may be explained by differences in study population characteristics, methodologies used to assess substrate oxidation and energy expenditure as well as the choice of contrasting foods.

PUFA have also been shown to inhibit transcriptional factors involved in the regulation of DNL, such as sterol regulatory element-binding protein 1c (SREBP-1c) and carbohydrate-responsive element-binding protein (ChREBP) [94,95], which may offer yet another plausible mechanism for the effect on liver fat from replacing SFA with PUFA. However, although rodent studies have shown increases in the expression of lipogenic proteins after consuming SFA versus PUFA [96], human studies have failed to see a difference in rates of fasting DNL between groups [86,97]. Elucidating the mechanisms by which SFA increase liver fat compared to PUFA in isolated dietary substitution trials has been shown to be difficult and several other mechanisms have been proposed in the literature, such as fatty-acid specific hepatic uptake, changes in the microbiome and endotoxemia [86].
Dietary patterns and NAFLD

Although trials of isolated replacements of SFA for MUFA/PUFA are of great importance to understand the impact of dietary fatty acids on NAFLD per se, investigating dietary patterns may be useful to understand the interplay of fat quality and other nutrients in the diet. The Mediterranean diet has been shown in both short-term (6-12 weeks) and long-term (18 months) trials to improve liver fat and other cardiometabolic risk markers compared to other diets [98–101]. Interestingly, in the DIRECT PLUS trial, a hypocaloric Mediterranean diet supplemented with extra flavonoids decreased liver fat over 18 months by almost double compared to a traditional Mediterranean diet, despite similar weight loss, suggesting that other dietary compounds than dietary fat (and other macronutrients) may be of importance [99].

Low-carbohydrate (LCHF) diets have been under investigation with regards to liver fat, with mixed results. Some studies have shown a favour towards the LCHF-diets compared to higher carbohydrate diets [100,102,103], whereas others have shown the opposite [104,105]. A recent meta-analysis of 12 isocaloric study comparisons showed a null-effect on liver fat between LCHF diets and high-carbohydrate low-fat (HCLF) diets [106]. Potential explanations for study heterogeneity may be due to differences in weight between comparator groups in ad libitum trials as well as differences in the content of SFA among the high-fat diets. In a recent study by Dalby Hansen et al., a LCHF diet that emphasised SFA-rich food sources such as high-fat dairy, red meat and butter showed a decrease in weight and hemoglobin A1c (HbA1c) levels in individuals with T2D compared to a HCLF diet [102]. However, no change in the activity of NAFLD (assessed using the NAFLD activity score (NAS) from liver biopsies) was shown between groups. Interestingly, the LCHF group also increased low-density lipoprotein (LDL)-cholesterol compared to the HCLF diet. Holmer et al. showed that a LCHF diet decreased liver fat to a similar extent as an intermittent fasting (5:2) diet protocol compared to usual care [103]. However, the usual care group showed improvements in liver stiffness (a biomarker of liver fibrosis) compared to the LCHF group, suggesting that a LCHF diet may be beneficial for some aspects of the NAFLD spectra, but not others. Comparing the long-term effect on liver fat of a LCHF diet that emphasises the replacement of SFA with PUFA with usual care are limited.

A diet rich in whole-grains, fruits and vegetables, lower relative amounts of SFA and higher relative amounts of PUFA has been shown to improve cardiometabolic risk markers in individuals with the metabolic syndrome, central obesity or hyperlipidemia [107–109]. A short-term RCT over 12 weeks demonstrated improvements in liver fat assessed by ultrasonography after consuming a diet rich in whole-grains compared to a diet lower in whole-grains [110]. Whole-grain intake has also been associated with liver fat content cross-sectionally [111] and prospectively [112]. Contrasting a whole-
grain rich diet low in SFA and higher in PUFA (such as the healthy Nordic diet) with other diets has not been investigated with regards to liver fat in a long-term RCT.

Research gaps in the field of fatty acids and NAFLD

From the background literature it is evident that many research gaps remain. How well do beneficial effects of PUFA consumption from short-term controlled trials generalize to the wider population? What are the mechanisms explaining the effect on liver fat accumulation following the consumption of SFA versus PUFA? What is the role of non-dietary fatty acids in NAFLD? How does dietary fat quality impact more severe forms of NAFLD? Is a low-carbohydrate diet emphasising good fat quality beneficial for reducing liver fat and other cardiometabolic risk markers? Is a more local diet focusing on foods familiar to the population any better than similar, but broader, dietary guidelines for reducing liver fat? By addressing these remaining questions, missing pieces to the larger puzzle are collected, allowing us to not only understand underlying pathophysiological mechanisms, but also to provide a foundation for dietary guidelines for the prevention and treatment of NAFLD and its consequences.
Aims

The overall aim of this doctoral thesis was to investigate the role of circulating, liver and dietary fatty acids in NAFLD. For circulating and liver fatty acids, special emphasis was given to fatty acids reflecting diet.

Specific aims

To investigate associations between fatty acids in circulating cholesteryl esters and liver fat content and energy metabolism in a population-based sample of middle-aged Swedish men and women (paper I).

To investigate associations between liver and plasma fatty acids in multiple lipid fractions and significant liver fibrosis in adult Swedish men and women with biopsy-proven NAFLD (paper II).

To investigate associations between isocaloric substitutions of SFA and SFA-rich food sources with other macronutrients and food sources and NAFLD-related cirrhosis and all-cause HCC in a population-based sample of middle aged to elderly Swedish men and women (paper III).

To investigate the effects of comparing three different dietary strategies: a low-carbohydrate high PUFA diet (LCPUFA) versus usual care (UC), a healthy Nordic diet (HND) versus UC and a LCPUFA diet versus a HND on 12-months changes in liver fat content and other cardiometabolic risk markers in adult Swedish men and women with prediabetes or T2D (paper IV).
Methods

Study descriptions

Paper I (POEM)
In paper I, baseline data from the population-based Prospective investigation of Obesity, Energy and Metabolism (POEM) cohort were used. The main aim of the POEM cohort is to study the pathophysiological links between obesity and vascular dysfunction and future CVD. A total of n=2008 individuals were invited to participate one month after their 50th birthday. Of these, n=502 (25%) signed a written informed consent and agreed to participate [113]. The study population for this particular study consisted of those individuals who had data on serum CE fatty acids and liver fat (n=308). For secondary analyses, the study population consisted of those individuals contributing data on serum CE fatty acids and measures of energy metabolism: respiratory quotient (RQ) (n=481) and REE (n=478).

Paper II (AM-02 NASH)
In paper II, cross-sectional data from the AM-02 NASH study were used. The main aim of the AM-02 NASH study is to evaluate the ability of non-invasive imaging biomarkers to discriminate between NASH and NAFL for the purpose to use in future clinical trials of NASH therapeutics. Patients were recruited from the Departments of Gastroenterology and Hepatology Hospital and from the Swedish CArdioPulmonary BioImage Study “SCAPIS”. Inclusion criteria were: individuals aged 18–70 with clinically suspected NAFLD and at least one of the following: imaging indicative of NAFLD, ALT more than 1,5 × upper limit of normal, CK18 M30 more than 180 U/L and/or biopsy showing NAFLD within three months prior to screening visit. Exclusion criteria included other liver diseases and high intake of alcohol [82]. A total of n=134 were screened for eligibility of which n=68 had NAFLD verified by liver biopsy. Of these, n=60 had data on liver and plasma fatty acids and constituted the study sample of this analysis.
Paper III (SMC and COSM)

In paper III, associations between macronutrient and food substitutions and NAFLD cirrhosis and all-cause HCC were investigated using two prospective cohort studies: the Swedish Mammography Cohort (SMC) and the Cohort of Swedish Men (COSM). The SMC cohort was initiated in 1987 and invited women from central Sweden who were born between 1914-1948 whereas COSM was initiated in 1997 and invited men from central Sweden who were born between 1918-1952, both with the aim of investigating associations between diet and chronic disease outcomes [114]. Participants are followed up with questionnaires on diet and lifestyle factors every 10 years (1987-1990, 1997, 2008/2009 and 2019/2020). Baseline diet in 1997 was used to define start of follow-up in this study. A total of n=100 303 men and n=56 030 women (pooled n = 156 333) were invited to participate in 1997, whereby n=88 087 agreed (56 %). After the exclusion of participants with incorrect/missing personal identification number, history of cancer, implausible energy intake (defined by cut-offs of <500 or >3500 kcal for women and <800 and >4200 kcal for men) and those who died before start of follow-up (1 January, 1998), n=79 729 remained and constituted the study sample for the HCC outcome. After further excluding individuals with high consumption of alcohol (defined by >20 grams/day for women and >30 grams/day for men) and with prevalent liver disease at baseline, n=77 059 remained and constituted the study sample for the NAFLD cirrhosis outcome.

Paper IV (NAFLDiet)

In paper IV, a three-arm parallel designed RCT was conducted to investigate the effects of a LCPUFA diet versus UC, a HND versus UC and a LCPUFA diet versus a HND on 12-month changes in liver fat in individuals with T2D or prediabetes (Figure 3). Men and women aged 30-75 years with prediabetes (defined as having a fasting glucose level of ≥5.6 mmol/L or an HbA1c level of ≥39 mmol/mol) or T2D (duration ≤10 years) with no insulin treatment and a BMI between 25-40 kg/m² were recruited from population-based cohort studies (Alla Nya Diabetiker i Uppsala (ANDiU) and EpiHealth) as well as through web-based advertisements. A total of n=150 individuals were randomized into three different diet groups using computer-generated stratified randomization based on sex and T2D status. Out of n=150 individuals, n=148 were informed of their diet allocation and n=2 dropped out before being informed.
Figure 3. Protocol figure of paper IV. On each monthly or bimonthly meeting (M0-M12), participants received key food items in a bag. BIA, Bioelectrical Impedance Analysis; HND, Healthy Nordic Diet; LCPUFA, Low-Carbohydrate Polyunsaturated Fatty Acids; MRI, Magnetic Resonance Imaging; UC, Usual Care; WFD, Weighed Food Diaries.
Table 2. Description of the four papers included in this thesis.

<table>
<thead>
<tr>
<th></th>
<th>Paper I</th>
<th>Paper II</th>
<th>Paper III</th>
<th>Paper IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study design</strong></td>
<td>Cross-sectional study</td>
<td>Cross-sectional study</td>
<td>Prospective cohort study</td>
<td>Randomized controlled trial</td>
</tr>
<tr>
<td><strong>Cohort</strong></td>
<td>POEM</td>
<td>AM-02 NASH</td>
<td>COSM and SMC</td>
<td>NAFLDiet trial</td>
</tr>
<tr>
<td><strong>Study sample size</strong></td>
<td>n=308</td>
<td>n=60</td>
<td>n=77 059 - 79 729</td>
<td>n=148</td>
</tr>
<tr>
<td><strong>Study sample</strong></td>
<td>50-year old Swedish men and women</td>
<td>Swedish men and women (18-70 years of age) with biopsy-confirmed NAFLD</td>
<td>Middle-aged to elderly Swedish men and women without cancer (and liver diseases for NAFLD cirrhosis) at baseline</td>
<td>Swedish men and women (30-75 years of age) with prediabetes or T2D, with no insulin treatment and a BMI between 25-40 kg/m²</td>
</tr>
<tr>
<td><strong>Exposure</strong></td>
<td>Serum CE fatty acids</td>
<td>Liver- and plasma fatty acids in multiple lipid fractions</td>
<td>Substitution of SFA and SFA-rich foods with other macronutrients and foods</td>
<td>A LCPUFA diet, a HND and UC</td>
</tr>
<tr>
<td><strong>Main outcome</strong></td>
<td>Liver fat (MRI)</td>
<td>Significant liver fibrosis ≥ F2 (biopsy)</td>
<td>NAFLD cirrhosis and HCC (ICD codes via linkages to electronic health records)</td>
<td>Liver fat changes during 12 months (MRI)</td>
</tr>
<tr>
<td><strong>Adjustment</strong></td>
<td>Adjustment was made for the following confounders: BMI, VO₂max, carbohydrates, fat, alcohol, vitamin E and sex</td>
<td>Adjustment was made for the following confounders: BMI, age, PNPLA3 (I148M)</td>
<td>Adjustment was made for the following confounders: age, sex, total energy intake or total food mass, dietary variables specified in each substitution model, current smoking status, education, family history of CVD, physical activity, sleep and BMI</td>
<td>Adjustment was made for the following baseline characteristics: baseline value of liver fat, T2D diagnosis and sex</td>
</tr>
</tbody>
</table>

CE, Cholesteryl Esters; COSM, Cohort of Swedish Men; CVD, Cardiovascular Disease; HCC; Hepatocellular Carcinoma; HND, Healthy Nordic Diet; LCPUFA; Low-Carbohydrate Polyunsaturated Fatty Acids; MRI, Magnetic Resonance Imaging; NAFLD, Non-Alcoholic Fatty Liver Disease; NASH, Non-Alcoholic Steatohepatitis; PNPLA3, Patatin-Like Phospholipase Domain-Containing Protein 3; SFA, Saturated Fatty Acids; SMC, Swedish Mammography Cohort; T2D, Type 2 Diabetes; UC, Usual Care; VO₂max, Maximal Aerobic Capacity.
 Ethics

Ethical approvals were retrieved from the regional ethical review board in Uppsala (paper I), the regional ethical review board at Karolinska Institutet (paper III) and the Swedish Ethical Review Authority (paper II, III and IV). All studies were conducted in accordance with the declaration of Helsinki. The randomized trial by which paper IV is based on is registered at ClinicalTrials.gov (NCT04527965) with the a priori statistical analysis plan (SAP) attached.

Assessment of exposure

In paper I and II, lipids from fasting serum/plasma samples were extracted using the Folch extraction method with the addition of a 2:1 solution of chloroform:methanol. The different lipid fractions (CE, PL, TAG and NEFA) were then separated using thin-layer chromatography (TLC). After methylation, fatty acids were profiled using gas chromatography (GC) (Agilent Technologies system, 6890N or Agilent Technologies system, 7890B), as previously described in detail [76]. Peak area retention times were compared with reference standard fatty acid methyl esters (FAMEs). Fatty acids were measured in plasma in the AM-02 NASH cohort and in serum in the POEM cohort. The relative amounts of major fatty acids in serum and plasma are highly correlated [115]. Fatty acids were presented as proportions of total fatty acids by dividing the area under the peak from the retention time of the fatty acid of interest with the total area from all fatty acids.

A total of 13 CE fatty acids were measured in the POEM cohort whereas 14 CE fatty acids, 14 TAG fatty acids and 19 PL fatty acids were measured in the AM-02 NASH study. All fatty acids are listed in Table 1 below. In both the POEM and AM-02 NASH cohorts, SCD-1, D5D and D6D were estimated using product-to-precursor ratios of the following fatty acids respectively: 16:1n-7/16:0, 20:4n-6/20:3n-6 and 18:3n-6/18:2n-6. In addition, in the AM-02 NASH cohort, two ratios reflecting elongase activities were assessed: 18:0/16:0 and 20:4n-6/18:2n-6. Fatty acid classes were constructed by pooling all individual fatty acids within that class (i.e. SFA, MUFA, PUFA, omega-3 PUFA and omega-6 PUFA) in the AM-02 NASH cohort.
Table 3. Fatty acids measured in paper I and II.

<table>
<thead>
<tr>
<th></th>
<th>Paper I</th>
<th></th>
<th>Paper II</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CE</td>
<td>CE</td>
<td>TAG</td>
<td>PL</td>
</tr>
<tr>
<td><strong>SFA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14:0</td>
<td>Myristic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15:0</td>
<td>Pentadecanoic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16:0</td>
<td>Palmitic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17:0</td>
<td>Heptadecanoic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18:0</td>
<td>Stearic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20:0</td>
<td>Arachidic acid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22:0</td>
<td>Behenic acid</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>24:0</td>
<td>Lignoceric acid</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>MUFA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16:1n-7</td>
<td>Palmitoleic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18:1n-9</td>
<td>Oleic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18:1n-7</td>
<td>Vaccenic acid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Omega-3 PUFA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18:3n-3</td>
<td>α-linolenic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20:5n-3</td>
<td>Eicosapentaenoic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22:5n-3</td>
<td>Docosapentaenoic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22:6n-3</td>
<td>Docosahexaenoic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Omega-6 PUFA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18:2n-6</td>
<td>Linoleic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18:3n-6</td>
<td>γ-linolenic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20:3n-6</td>
<td>Dihomo-γ-linolenic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20:4n-6</td>
<td>Arachidonic acid</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CE, Cholesteryl Esters; MUFA, Monounsaturated Fatty Acids; PL, Phospholipids; POEM, Prospective investigation of Obesity, Energy and Metabolism; PUFA, Polyunsaturated Fatty Acids; SFA, Saturated Fatty Acids; TAG, Triacylglycerols.

In paper III, diet was assessed using a validated food-frequency questionnaire (FFQ), covering 96 food and beverage items. Response categories ranged from never to ≥ 3 times per day, with additional space for answers to open-ended questions regarding food that is commonly consumed in Sweden. Average intakes of foods were calculated by multiplying the frequency by which the participant consumed the food with age-specific portion sizes. Data was linked to the Swedish National Food Agency food composition database and nutrients were obtained. Total energy intake was calculated by adding carbohydrates, fiber, SFA, MUFA, PUFA, protein and alcohol whereas total food intake was calculated by adding all food groups in grams. For the food substitution models, food items were categorized into butter, margarine and vegeta-
ble oils, milk, fermented milk, fish, unprocessed red meat, fruits and vegetables, alcoholic beverages, sweets and other foods. A priori determined substitutions of interest were: SFA with PUFA, SFA with fiber, SFA with carbohydrates, unprocessed red meat with fish, butter with margarine and vegetable oil and milk with fermented milk. As the aim was to investigate isocaloric nutrient and food substitutions, other relevant food substitutions were not deemed feasible due to differences in caloric intake when foods are expressed in grams. The FFQ has been validated against multiple 24-hour recalls, dietary records as well as biomarkers from subcutaneous adipose tissue [116–120]. Correlation coefficients between the FFQ and 7-day dietary food records were 0.76 for carbohydrates, 0.7 for fermented milk and 0.3-0.7 for red and processed red meat. For SFA and PUFA, correlation coefficients between the FFQ and fourteen 24-hour recall interviews were 0.75 and 0.49, respectively [116].

In paper IV, participants received both orally presented as well as written information on their respective diet. The LCPUFA group was instructed to consume a diet lower in carbohydrates and higher in fat, with a focus on unsaturated fatty acids (in particular PUFA) over SFA. Focus points for the diet was to eat at least 40 grams of nuts per day, two table spoons of seeds per day, two table spoons of sunflower oil per day and to replace rice, pasta, bread, cereals and potatoes with other lower-carbohydrate sources. The LCPUFA group received key food items such as walnuts, cashew nuts, sunflower seeds, pumpkin seeds and sunflower oil on a monthly to bimonthly basis. The HND group was instructed to consume a diet lower in fat and higher in carbohydrates, with a focus on whole-grain rich food products. Focus points for the diet was to eat two portions of whole-grain from oat, barley or rye each day, eat at least one slice of whole-grain bread per day, limit fat-rich products, and to prioritize Nordic foods such as Swedish fruits and vegetables, salmon and mackerel and rapeseed oil. The HND group received key food items such as oat products, crisp bread, granola, raspberries, legumes, low-fat margarine and almonds. The UC group was instructed to follow the dietary guidelines presented in NNR 2012 edition with a focus on increasing fruits and vegetables to 600 grams per day (instead of 500 grams per day). This was done to partly mask the control group from knowing that they were in the control group. The UC group also received key food items including frozen vegetables, frozen fruit, carrots, crisp bread and muesli.

All three diet groups were instructed to increase fruits and vegetables, replace whole-fat dairy with low-fat dairy and butter with margarine or vegetable oils and to limit red and processed red meat, sugar-sweetened beverages, pastries and other energy dense sugar/fat-rich snacks.

Dietary adherence was assessed using 4-day weighed food diaries (WFD) at baseline, after 6 months and after 12 months. Dietary biomarkers such as plasma 18:2n-6 and 22:6n-3, and alkylresorcinols are currently being analyzed. Plasma fatty acids were analyzed in a similar way as in paper I and II,
with some differences, including using solid-phase extraction over TLC to separate lipid fractions [121].

Assessment of outcomes

Liver fat content in paper I and paper IV was assessed using MRI (Achieva, Philips Healthcare) as described in detail previously [122]. In paper I, energy metabolism was assessed in a fasted state using indirect calorimetry (Jaeger Oxycon Pro, Vyaire Medical). Whole-body RQ (measure of substrate utilization) was calculated by dividing the volume of CO₂ by the volume of O₂. REE was calculated using the abbreviated Weir equation [123]. REE was further divided by fat-free mass (FFM), assessed using bioelectrical impedance analysis (BIA) (Tanita BC-418).

In paper IV, secondary outcomes including glucose and lipid parameters were assessed using routine laboratory techniques in a fasted state (at least 10 hours). FIB-4, an index of liver fibrosis, was calculated using the following formula: (age x AST)/(platelet counts x ALT) and homeostatic model assessment for insulin resistance (HOMA-IR) was calculated using the following formula: (insulin x glucose)/22.5. Systolic (SBP) and diastolic blood pressure (DBP) were measured using an automated blood pressure device (Omron). Weight was assessed using BIA (Tanita).

In paper II, liver biopsies were taken guided by ultrasound in a fasted state (>6 hours) and evaluated using the steatosis, activity and fibrosis (SAF) scoring system by two liver pathologists [124]. Stages of fibrosis were classified accordingly: F0 (none), F1 (1a or 1b perisinusoidal zone 3 or 1c portal fibrosis), F2 (perisinusoidal and periporal fibrosis without bridging), F3 (bridging fibrosis) and F4 (cirrhosis). Fibrosis was dichotomized by combining F0 and F1 in one group (F0-1) (none to mild fibrosis) and F2, F3 and F4 in the other group (F2-4) (significant fibrosis).

In paper III, NAFLD cirrhosis and HCC were assessed using information from the Swedish National Patient Register (NPR), the Swedish National Cancer Register and the Swedish Cause of Death Register. The NPR includes all completed in-patient stays as well as out-patient visits from specialized care. All nationwide in-patient stays are included from 1987 whereas all out-patient visits are included from 2001. The National Cancer Register includes all reported cancer cases from six regional cancer centers in Sweden. The Cause of Death Register includes information about the causes of death in the Swedish population and is updated each year. All three registries contain International Classification of Diseases (ICD) codes. The validity and coverage of the three registries are generally considered good. NAFLD cirrhosis was defined as having an ICD-10 code for unspecified cirrhosis or any complications of cirrhosis such as portal hypertension, gastric/esophageal varices or hepatorenal syndrome without any concomitant liver diseases associated with cirrhosis.
(e.g. viral hepatitis or ALD) except for NAFLD or NASH. Positive predictive values (PPV) for these outcomes are generally high (e.g. 91% for unspecified cirrhosis and 96% for esophageal varices) [125]. HCC was defined as having an ICD of Oncology (ICD-O/2) or ICD-10 code of C22.0. PPV for HCC has been shown to be 84% in Swedish registries [125].

Assessment of confounders

In paper I, potential confounders were identified using the background literature. In the fully adjusted model, BMI, maximal aerobic capacity (VO$_{2\text{max}}$), carbohydrates, fat, vitamin E, alcohol and sex were included in the linear regression analyses. BMI was calculated as the weight (kg) divided by the height (m) squared. VO$_{2\text{max}}$ was measured using a bicycle ergometer coupled to indirect calorimetry. Diet was assessed using a FFQ consisting of 139 food items (including alcohol).

In paper II and III, potential confounders were identified using the background literature, expert knowledge, and directed acyclic graphs (DAGs), built in the online software tool: Dagitty (dagitty.net) [126]. DAGs are non-parametric graphs depicting one’s causal assumptions and are useful for identifying confounding paths as well as illustrating other biases [127]. DAGs are arranged in a temporal order and need to be both directed (meaning that the arrow (or arc) from one variable (or node) to another variable need to be in one direction) and acyclic (meaning that a variable cannot cause itself, either directly or indirectly). Arrows in a DAG are most commonly probabilistic but may in some circumstances be deterministic, i.e. a variable may be fully predicted by knowing the values of its parent variables (variables temporally preceding the variable in focus). As an example, total energy intake (child variable) is fully determined by the energy from carbohydrates, fat and protein (parent variables). Omitting an arrow assumes a lack of causal relationship between the two variables. A common cause of the exposure and the outcome is called a confounder. A confounding pathway is open and non-causal, thereby transmitting a statistical association between the exposure and the outcome if not conditioned on. An effect of the exposure but a cause of the outcome is called a mediator. A mediating pathway is an open causal pathway. Lastly, a common effect of two variables is called a collider. A collider pathway is a closed non-causal pathway. No statistical association is therefore transmitted between the variables if not conditioned on. For a more detailed description of DAGs, please see Figure 4 below. In paper II, BMI, age, sex, diet, PNPLA3 (I148M) and TM6SF2 (E167K) were identified as potential confounders (Figure 5). However, due to the small sample size only the most relevant confounders (decided upon a priori) were chosen. These were: BMI, age and PNPLA3 (I148M) genotype. BMI was calculated by dividing the weight (kg)
by the height (m) squared. PNPLA3 (I148M) was genotyped using the Taq-
Man® PCR method. Further sensitivity analyses were performed including sex
and the TM6SF2 (E167K) genotype. Diet was identified as an unobserved
confounder and was not able to be included in the regression model.

To estimate the relative substitution effect of increasing SFA or an SFA-
rich food for a decrease in another macronutrient or food, the following ad-
justment set was identified in paper III: age, sex, total energy intake (isocaloric
models) or total food mass (equal-mass models), dietary variables specified in
each leave-one-out model, current smoking status, education, family history
of CVD, physical activity, sleep, previous diet and BMI (Figure 6). As base-
line BMI might be depicted as an effect of diet and not a cause when both are
assessed at the same point in time, pre-baseline BMI was estimated from in-
formation provided in the 1997 questionnaire on weight at least five years
prior to baseline. Previous diet was identified as an unobserved confounder.
Sensitivity analyses were conducted to assess the impact of potential con-
foundering bias from previous dietary intake.
Figure 4. Directed acyclic graphs (DAG) illustrating the causal assumptions between the exposure (A), the outcome (Y), variable M, variable C1 and variable C2. In example A) variable C is identified as a common cause of both A and Y. Using the definition from causal graph theory, C is thus a confounder. M is identified as an effect of A and a cause of Y and hence a mediator on the path between A and Y. By conditioning (e.g. restricting, matching or adjusting) on C (indicated by the C-node becoming white), the following backdoor path “A ← C → Y” is closed (indicated by the pink path becoming black). The confounding path has thus been accounted for and the total effect of A on Y can now be estimated. In example B) the effect of variable C1 on outcome Y is fully mediated by its effect on C2. By conditioning on either C1 or C2, the following backdoor path “A ← C1 → C2 → Y” is closed and the total effect of A on Y can be estimated. In example C) C1 is identified as a common cause of A and Y (i.e. a confounder) but C2 is identified as a common cause of variable M and Y. Conditioning on C1 would close the following backdoor path “A ← C1 → Y” (as in example A)) and it would be sufficient to estimate the total effect of A on Y. The following path “A → M ← C2 → Y” is already closed (indicated by that path being black). Using so called d-separation rules of causal graph theory, this is due to the fact that the two arrows from A and C2 collide into variable M, making M a collider variable on that path. Conditioning on M (indicated by the M-node becoming white) opens up the following path “A → M ← C2 → Y” (indicated by that path becoming pink) and we have introduced a systematic bias called collider bias.
Figure 5. A directed acyclic graph (DAG) illustrating the causal framework for paper II. Potential confounders that were identified using this approach were: sex, patatin-like phospholipase domain-containing protein 3 (PNPLA3), transmembrane 6 superfamily member 2 (TM6SF2), age, body mass index (BMI) and diet. Diet is indicated as an unobserved variable. Picture retrieved from Fridén M, et al. Front Med. 2022 [128].
Figure 6. A directed acyclic graph (DAG) illustrating the causal framework for paper III. Potential confounders that were identified using this approach were: family history of CVD, smoking status, sleep, education, age, sex, physical activity, BMI, previous diet, nutrients/foods necessary for the specific substitution of interest and total energy intake or total food intake for food substitutions. Previous diet is depicted as an unobserved confounder. The DAG-example is given for macronutrient substitutions but applies to food substitutions as well.

Substitution modelling

Adjusting for total energy intake (TEI) when assessing exposure-outcome relationships in nutritional epidemiology is often justified to account for common determinant of dietary intake (e.g. body size and metabolic efficiency) [129]. However, including TEI transforms the estimand to a joint relative substitution effect of increasing the exposure of interest for a weighted average of the population-specific background distribution of the diet. The effect estimate obtained is therefore both difficult to interpret and to transport to other populations. More specified substitution models may help to mitigate some of these issues.

In paper III, isocaloric and equal-mass substitution analyses described by Tomova et al. in 2022 were conducted [130]. The leave-one-out method was used, which is mathematically equivalent to the energy partition method (another common method) and performs equally well when only one food item is replaced by another. The isocaloric leave-one-out model includes all nutrients/foods in kcal except for the one to be replaced, as well as total energy intake as a composite variable of these nutrients/foods. The equal-mass leave-
one-out model includes all nutrients/foods in gram except for the one to be replaced, as well as total food intake in grams as a composite variable of these nutrients/foods. A summary table describing the model specifications, mathematical derivations and illustrating DAGs for a one unit increase of food 1 in place of a one unit decrease of food 2 are provided below (Table 4). For simplicity we assume only three foods exist in the world. Mathematical equations are derived from the partition method of substitution modelling. The partition method includes all nutrients/foods in the model without any composite variable. The substitution effect is estimated from subtracting the coefficient from food 1 with the coefficient from food 2. The partition method is mathematically equivalent to the leave-one-out model.

Table 4. A summary of the two substitution modelling approaches used in paper III.

<table>
<thead>
<tr>
<th>Model</th>
<th>Model specification</th>
<th>Mathematical derivation</th>
<th>Directed acyclic graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isocaloric leave-one-out model</td>
<td>( F(Y) = \beta_0 + \beta_1\text{Food}_1(\text{kcal}) + \beta_2\text{Food}_3(\text{kcal}) + \beta_3\text{Totalenergyintake(\text{kcal})} )</td>
<td>( \text{Person 1:} ) ( B_0 + B_1(\text{Food}_1 + 1) + B_3\text{Food}_2 + B_3\text{Food}_3 + B_4\text{Totalenergyintake} )</td>
<td>![DAG for Isocaloric leave-one-out model]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{Person 2:} ) ( B_0 + B_1 + B_2(\text{Food}_2 + 1) + B_3\text{Food}_3 + B_4\text{Totalenergyintake} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{Person 1 – Person 2:} ) ( B_1 - B_2 )</td>
<td></td>
</tr>
<tr>
<td>Equal-mass leave-one-out model</td>
<td>( F(Y) = \beta_0 + \beta_1\text{Food}_1(\text{g}) + \beta_2\text{Food}_3(\text{g}) + \beta_3\text{Totalfoodintake(\text{g})} )</td>
<td>( \text{Person 1:} ) ( B_0 + B_1(\text{Food}_1 + 1) + B_3\text{Food}_2 + B_3\text{Food}_3 + B_4\text{Totalfoodintake} )</td>
<td>![DAG for Equal-mass leave-one-out model]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{Person 2:} ) ( B_0 + B_1 + B_2(\text{Food}_2 + 1) + B_3\text{Food}_3 + B_4\text{Totalfoodintake} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \text{Person 1 – Person 2:} ) ( B_1 - B_2 )</td>
<td></td>
</tr>
</tbody>
</table>
Statistical analyses

In paper I, crude and multivariable linear regression analyses were performed to address our primary research question. In paper II, multivariable logistic regression analyses were conducted. Regression models were adjusted for BMI (kg/m² [continuous]), VO₂max (ml/kg/min [continuous]), carbohydrates, fat, alcohol and vitamin E (E%, g and mg [continuous]) and sex (men/women [categorical]) in paper I and for age (years [continuous]), BMI (kg/m² [continuous]) and PNPLA3 (I148M) genotype (CC/(CG and GG) [categorical]) in paper II. For linear regression, estimates were expressed as β-coefficients with 95% confidence intervals (CI). For logistic regression, estimates were expressed as odds ratios (OR) with 95% CI. Further correlation analyses (Pearson’s correlation for normally distributed variables and Spearman’s rank correlation for non-normally distributed variables) were performed in paper I (secondary analyses). Additional multivariable linear regression analyses were performed between fatty acids and liver fat, adjusted for REE/FFM for those fatty acids that were associated with both liver fat and REE/FFM in the crude models in paper I (secondary analyses). Normal distribution was assessed using the Shapiro-Wilk W test, with a W-value > 0.95 indicating normal distribution. A P-value <0.05 was set as the significance level. All statistical analyses were performed using the JMP software (SAS Institute, Inc) version 13.1.0 (paper I) and version 15.1.0 (paper II).

In paper III, time-to-event analyses using multivariable Cox proportional hazard regression models were conducted. Hazard ratios (HR) with corresponding 95% CI were estimated. Time under follow-up was calculated from baseline (January 1, 1998) to outcome diagnosis, death or administrative end of follow-up (December 31, 2021), whichever occurred first, for both outcomes. For NAFLD cirrhosis, individuals that developed any other liver disease associated with cirrhosis under follow-up were censored (e.g. from ALD or viral hepatitis). Outcomes were identified through linkage to Swedish registries, i.e. the NPR, the Swedish Cancer Register and the Cause of Death Register. Cox regression models were adjusted for age (years [continuous]), sex (men/women [categorical]), total energy intake or total food mass (kcal or gram [continuous]), dietary variables specified in each leave-one-out model (kcal or gram [continuous]), current smoking status (never/former/current smoker [categorical]), education (primary school <= 9 years, high school 10-12 years or university > 12 years [categorical]), family history of CVD (yes/no if either mother or father have had a myocardial infarction before the age of 60 [categorical]), physical activity (<1, 1, 2-3, 4-5 or >5 hours/week of exercise [categorical]), sleep (hours/day [continuous]) and BMI (<18.5 kg/m², 18.5-24.9 kg/m², 25-30 kg/m² and >30 kg/m² [categorical]). Missing data on confounders were imputed using Multivariate Imputation using Chained Equations (MICE) (n=5 imputations) using the “mice” package in R [131].
Pooling of parameter estimates was done using Rubin’s rules. Several sensitivity analyses were conducted to assess the robustness of the primary findings, including 1) complete-case analyses, 2) exclusion of the first 2 years of follow-up, 3) further subcategorization of “other foods” in the food substitution models to investigate confounding bias by grouping foods with potentially different effects on the outcome together, 4) only counting NAFLD cirrhosis cases who had a history of T2D, hypertension and/or hyperlipidaemia or a BMI ≥25 kg/m² at baseline, 5) combining the Swedish Cancer Register and the Cause of Death Register with the NPR to define HCC as the Swedish Cancer Register has been shown to underreport cases [132], and 6) updating our baseline to 2009, adjusting for diet in 1997 to assess the impact of confounding bias from past dietary exposure (only for NAFLD cirrhosis). All analyses were conducted using R (R Core Team, Vienna, Austria) version 3.6.0.

In paper IV, general linear models (GLM) with change in the primary and secondary outcomes (month 12 – month 0) as the outcome, diet group as a fixed factor and T2D status, sex and baseline value of the outcome as covariates were conducted. The primary causal estimand of interest was the intention-to-treat (ITT) effect. To conserve baseline randomization and account for differential loss to follow-up, missing data of the outcomes were imputed using MICE (n=20 imputations) [131]. When the overall between-group GLM test was statistically significant (p<0.05), post-hoc linear models were performed to estimate marginal means (EMMs) with corresponding 95% CI for pairwise comparisons. Assumptions of the GLM were checked using residual plots and the Shapiro-Wilk test. When the assumptions of the GLM were violated, Willetts residual method was employed. Residuals were calculated by including the outcome, baseline value of the outcome, sex and T2D status in a GLM. Kruskal-Wallis (KW) tests were thereafter performed by regressing the residuals on diet group. If the overall between-group KW test was statistically significant (p<0.05), further post-hoc tests were performed to estimate marginal medians (EMM) with 95% confidence intervals. For retrieving pairwise median difference and 95% CI, bootstrapping (n=10 000) combined with MICE (n=20) was employed using the “bootimpute” package in R [133].

A priori decided subgroup analyses were conducted for sex (men/women), T2D status (yes/no), NAFLD status (yes/no) and I148M PNPLA3 genotype (CC/(CG+GG)). Further sensitivity analyses were conducted including 1) using a different imputation model and 2) excluding n=4 pairs that were co-randomized to the same diet group; co-randomization was done to avoid any spillover effects within the same families. Lastly, a post-hoc causal mediation analysis (CMA) was conducted to estimate the proportion mediated (PM) by weight change on liver fat between groups, using the regmedint package in R [134]. The original SAP can be found at Clinicaltrials.gov (NCT04527965). All analyses were conducted in R (R Core Team, Vienna, Austria) version 4.2.3 and IBM SPSS Statistics version 28.0.1.0 (142).
Main results

Paper I

The study population of \( n = 308 \) 50-year olds had an equal distribution of men and women, a median BMI of 25.8 (4.9) kg/m\(^2\), an LDL-cholesterol concentration of 3.4 ± 0.9 mmol/L and a NAFLD prevalence of 22.7% (determined by a liver fat content exceeding 5.5%). The primary aim of paper I was to investigate associations between fatty acids in CE and liver fat content. After adjusting for BMI, VO\(_{2\text{max}}\), dietary factors and sex, inverse associations were observed between 18:2n-6 (\( \beta = -0.03, 95\% \text{ CI:} -0.06, -0.001 \)) and D5D (\( \beta = -0.42, 95\% \text{ CI:} -0.76, -0.08 \)), whereas positive associations were observed for 16:1n-7 (\( \beta = 0.37, 95\% \text{ CI:} 0.04, 0.70 \)), 18:3n-6 (\( \beta = 0.59, 95\% \text{ CI:} 0.28, 0.90 \)), 20:3n-6 (\( \beta = 1.20, 95\% \text{ CI:} 0.65, 1.75 \)), 20:4n-6 (\( \beta = 0.08, 95\% \text{ CI:} 0.002, 0.16 \)), SCD-1 (\( \beta = 0.36, 95\% \text{ CI:} 0.02, 0.71 \)) and D6D (\( \beta = 25.63, 95\% \text{ CI:} 11.98, 39.28 \)) and liver fat content (Table 4).

Secondary aims were to investigate correlations between fatty acids and basal fat oxidation and REE/FFM and whether these associations could explain potential associations between fatty acids and liver fat. Weak inverse correlations were observed between 16:1n-7 (\( r = -0.11, P = < 0.05 \)) and SCD-1 (\( r = -0.10, P = < 0.05 \)) and RQ, a measure of basal fat oxidation. Weak inverse correlations were observed between 16:0 (\( r = -0.12, P = < 0.05 \)), 18:0 (\( r = -0.20, P = < 0.001 \)), 18:1n-9 (\( r = -0.14, P = < 0.05 \)) and 20:3n-6 (\( r = -0.17, P = < 0.001 \)) and REE/FFM whereas positive correlations were observed between 18:2n-6 (\( r = 0.09, P = < 0.05 \)), 22:6n-3 (\( r = 0.13, P = < 0.05 \)) and D5D (\( r = 0.11, P = < 0.05 \)) and REE/FFM. When further adjusting for REE/FFM, no associations except for the 22:6n-3-liver fat association (from \( \beta = -0.63, 95\% \text{ CI:} -1.24, -0.02 \) to \( \beta = -0.34, 95\% \text{ CI:} -0.95, 0.27 \)) were markedly altered.
Table 4. Multivariable linear regression analyses between fatty acids in serum cholesteryl esters and ln liver fat.

<table>
<thead>
<tr>
<th>Fatty Acid</th>
<th>β (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myristic acid (14:0)</td>
<td>0.41 (-0.13, 0.95)</td>
</tr>
<tr>
<td>Pentadecanoic acid (15:0)</td>
<td>-1.69 (-3.75, 0.37)</td>
</tr>
<tr>
<td>Palmitic acid (16:0)</td>
<td>0.09 (-0.06, 0.24)</td>
</tr>
<tr>
<td>ln Palmitoleic acid (16:1n-7)</td>
<td>0.37 (0.04, 0.70)</td>
</tr>
<tr>
<td>ln Stearic acid (18:0)</td>
<td>0.29 (-0.26, 0.83)</td>
</tr>
<tr>
<td>Oleic acid (18:1n-9)</td>
<td>0.009 (-0.05, 0.06)</td>
</tr>
<tr>
<td>Linoleic acid (18:2n-6)</td>
<td>-0.03 (-0.06, -0.001)</td>
</tr>
<tr>
<td>γ-linolenic acid (18:3n-6)</td>
<td>0.59 (0.28, 0.90)</td>
</tr>
<tr>
<td>α-linolenic acid (18:3n-3)</td>
<td>-0.04 (-0.42, 0.33)</td>
</tr>
<tr>
<td>Dihomo-γ-linolenic acid (20:3n-6)</td>
<td>1.20 (0.65, 1.75)</td>
</tr>
<tr>
<td>Arachidonic acid (20:4n-6)</td>
<td>0.08 (0.002, 0.16)</td>
</tr>
<tr>
<td>ln Eicosapentaenoic acid (20:5n-3)</td>
<td>-0.06 (-0.31, 0.20)</td>
</tr>
<tr>
<td>Docosahexaenoic acid (22:6n-3)</td>
<td>-0.24 (-0.75, 0.27)</td>
</tr>
<tr>
<td>ln Stearoyl CoA desaturase-1 (SCD1)</td>
<td>0.36 (0.02, 0.71)</td>
</tr>
<tr>
<td>ln Delta-5 desaturase (D5D)</td>
<td>-0.42 (-0.76, -0.08)</td>
</tr>
<tr>
<td>Delta-6 desaturase (D6D)</td>
<td>25.63 (11.98, 39.28)</td>
</tr>
</tbody>
</table>

1Model is adjusted for ln BMI, VO2max, ln carbohydrate, ln fat, ln vitamin E and sex.

BMI, Body Mass Index; VO2max, Maximal Aerobic Capacity.

Paper II

The study population of n=60 individuals with biopsy-proven NAFLD had an equal distribution of men and women between fibrosis groups F0-1 and F2-4. In addition, there were no statistically significant differences between fibrosis groups in any of the clinical (e.g. platelet concentration was 232.6±53.5 10⁹/L in F0-1 and 235.0±61.6 10⁹/L in F2-4) or histological variables (e.g. NASH prevalence of 69% in F0-1 and 88% in F2-4) related to disease severity, except for fibrosis. In the F0-1 group, 89% (n=32) had fibrosis stage 1 and in the F2-4 group, 79% (n=19) had fibrosis stage 2, 13% (n=3) had fibrosis stage 3 and 8% (n=2) had fibrosis stage 4.
The primary aim of paper II was to investigate associations between liver fatty acids in multiple lipid fractions and NAFLD fibrosis (Figure 7). A secondary aim was to investigate whether these fatty acid-NAFLD fibrosis associations would be replicated in the corresponding plasma lipid fractions as well. After adjusting for BMI, age and PNPLA3 (I148M) genotype, positive associations were observed for liver PL behenic acid (22:0) (OR: 1.86, 95 % CI: 1.0, 3.45) and liver PL SFA (OR: 2.13, 95 % CI: 1.10, 4.12) and NAFLD fibrosis, while inverse associations were observed for liver PL 22:6n-3 (OR: 0.45, 95 % CI: 0.23), liver TAG 18:1n-9 (OR: 0.52, 95 % CI: 0.28, 0.95), liver TAG 18:1 (OR: 0.52, 95 % CI: 0.28, 0.96), liver PL PUFA (OR: 0.39, 95 % CI: 0.20, 0.76) and liver TAG MUFA (OR: 0.52, 95 % CI: 0.28, 0.96). These associations were not markedly altered after further including the TM6SF2 (E167K) genotype and sex as potential confounders.

For secondary analyses, plasma PL 22:0 (OR: 0.46, 95 % CI: 0.25, 0.86), plasma TAG 18:1n-9 (OR: 0.55, 95 % CI: 0.31, 0.99), plasma TAG 18:1 (OR: 0.54, 95 % CI: 0.30, 0.97) and plasma TAG MUFA (OR: 0.50, 95 % CI: 0.27, 0.93) were inversely associated with NAFLD fibrosis.
Figure 7. Associations between liver fatty acids and significant liver fibrosis. Data are presented as adjusted odds ratios (OR) with 95% confidence intervals (CI) and P-values for each standard deviation change in liver fatty acid proportions. OR are adjusted for BMI, age and PNPLA3 (I148M) genotype. BMI, Body Mass Index; PNPLA3, Patatin-Like Phospholipase Domain-Containing Protein 3; 14:0, myristic acid; 15:0, pentadecanoic acid; 16:0, palmitic acid; 16:1n-7, palmitoleic acid; 17:0, heptadecanoic acid; 18:0, stearic acid; 18:1n-9, oleic acid; 18:1n-7, vaccenic acid; 18:1, oleic acid combined with vaccenic acid; 18:2n-6, linoleic acid; 18:3n-6, γ-linolenic acid; 18:3n-3, α-linolenic acid; 20:0, arachidic acid; 20:3n-6, dihomo-γ-linolenic acid; 20:4n-6, arachidonic acid; 20:5n-3, eicosapentaenoic acid; 22:0, behenic acid; 22:5n-3, docosahexaenoic acid; 24:0, lignoceric acid; SCD-1, stearoyl-coA desaturase; D5D, delta 5 desaturase; D6D, delta 6 desaturase; AA/LA, arachidonic acid/linoleic acid. SCD-1, D5D and D6D are estimated using fatty acid product-to-precursor ratios: 16:1n-7/16:0 (SCD-1), 20:4n-6/20:3n-6 (D5D), 18:3n-6/18:2n-6 (D6D). D5D in cholesteryl esters: n(F0-1)=19, n(F2-4)=16 due to 25 zero-values of 20:3n-6. Picture retrieved from Fridén M, et al. Front Med. 2022 [128].

Paper III

The study population of a somewhat equal proportion of men and women had an average age of 60 (53.0-69.0), 17% had a University degree and 24% were current smokers. The average intake of SFA was 14 (12-16) E% and the average intake of PUFA was 4 (3-4) E%. The median follow-up time was 24 years. Substituting SFA with carbohydrates showed a HR of 1.00 (95% CI: 0.89-1.11) for NAFLD cirrhosis and 0.87 (95% CI: 0.74-1.02) for HCC. Substituting butter with margarine and vegetable oils showed a HR of 1.03 (95% CI: 0.99-1.07) for NAFLD cirrhosis and 1.00 (95% CI: 0.95-1.06) for HCC. Substituting milk with fermented milk showed a HR of 0.97 (95% CI: 0.92-1.03)
for NAFLD cirrhosis and 0.93 (95% CI: 0.85-1.01) for HCC (Figure 8). Results were moderately robust in sensitivity analyses investigating the potential for reverse causation, selection bias, unmeasured confounding, residual confounding as well as outcome misclassification.

**Figure 8.** Associations between food and macronutrient substitutions and NAFLD cirrhosis and HCC. Hazard ratios (HR) with corresponding 95% confidence intervals (CI) are presented. Estimates are adjusted for age, sex, education, sleep, smoking, family history of CVD, BMI, physical activity, total energy intake (for macronutrients), total food intake (for foods) and other dietary variables necessary for estimating the joint substitution effect. The unit of exposure for each substitution is: SFA with PUFA (50 kcal), SFA with carbohydrates (50 kcal), SFA with fiber (20 kcal), butter with margarine and vegetable oil (5 gram), unprocessed red meat with fish (30 gram) and milk with fermented milk (100 gram). BMI, Body Mass Index; CVD, Cardiovascular Disease; SFA, Saturated Fatty Acids; PUFA, Polyunsaturated Fatty Acids. n=77059 with 566 events for NAFLD cirrhosis and n=79729 with 205 events for HCC.
Paper IV

A flow-chart of the NAFLDiet trial is depicted in Figure 9 below. A total of n=222 screening visits were scheduled, from which n=150 individuals were randomly allocated to one of the three diets. After randomization but before providing information of the assigned diets, n=2 individuals dropped out, leaving a final study sample of n=148. Of the n=148 individuals that received information of their allocated diet, n=8 dropped out (5% drop out rate). Study participants were 65 (10) years old at baseline, had a BMI of 30.1 ± 3.6 kg/m² (LCPUFA), 29.5 ± 3.7 kg/m² (HND), 30.3 ± 3.2 kg/m² (UC) and a distribution of women/men of 41/59 % (LCPUFA), 39/61 % (HND) and 37/63 % (UC). Median glucose levels were 6.7 (1.9) mmol/L (LCPUFA), 6.7 (1.6) mmol/L (HND) and 6.7 (1.0) mmol/L (UC) whereas median percentage liver fat was 6.3 (6.9) % (LCPUFA), 6.4 (5.5) % (HND) and 8.7 (10.8) % (UC).

Figure 9. Flow-chart of the NAFLDiet trial. BMI, Body Mass Index; HND, Healthy Nordic Diet; LCPUPA, Low-Carbohydrate Polyunsaturated Fatty Acids; UC, Usual Care.

Adherence to the diets was assessed using 4-days WFD and plasma PL 18:2n-6. For the overall test, between-group differences were observed for total energy intake, carbohydrates, fat, SFA, MUFA, PUFA, F&V including berries, nuts and seeds, oats, sunflower oil, butter and whole-grain bread but not for fiber, protein, fatty fish, red and processed red meat or rapeseed oil. In a subgroup of n=47, between-group differences were observed for PL 18:2n-6.

For pairwise comparisons, mean differences in PL 18:2n-6 were 1.84 (95% CI: 0.14, 3.54) % for LCPUPA vs HND, 2.76 (1.00, 4.52) for LCPUPA vs UC and 0.92 (-0.75, 2.58) for HND vs UC. Mean differences in carbohydrate intake were -12.2 (95% CI: -14.8, -9.7) E% for LCPUPA vs HND, -13.2 (-15.9,
-10.5) E% for LCPUFA vs UC and -1.0 (-3.8, 1.7) E% for HND vs UC whereas mean differences in fat intake were 13.8 (95% CI: 11.3, 16.2) E% for LCPUFA vs HND, 14.5 (11.8, 17.1) E% for LCPUFA vs UC and 0.7 (-2.0, 3.4) E% for HND vs UC. Mean differences in PUFA intake were were 7.7 (95% CI: 6.1, 9.2) E% for LCPUFA vs HND, 9.8 (8.2, 11.4) E% for LCPUFA vs UC and 2.2 (0.6, 3.8) E% for HND vs UC whereas mean differences in SFA intake were 0.6 (95% CI: -0.6, 1.8) E% for LCPUFA vs HND, -1.3 (-2.6, 0.0) E% for LCPUFA vs UC and -1.9 (-3.2, -0.6) E% for HND vs UC.

**Figure 10.** Within-group changes with corresponding 95% CIs in macronutrients (E%) and foods (E%) for those that were statistically significant in the overall general linear model test of between-group comparisons. Different letters above or below the error bars denote p<0.05 for pairwise group comparisons. Results are presented for the intention-to-treat (ITT) analysis. E%, Energy Percentage; F&V, Fruits And Vegetables; HND, Healthy Nordic Diet; LCPUFA, Low Carbohydrate Polyunsaturated Fatty Acids; MUFA, Monounsaturated Fatty Acids; PUFA, Polyunsaturated Fatty Acids; SFA, Saturated Fatty Acids; UC, Usual Care.

With regards to the primary and secondary outcomes of the study, between-group differences in the overall test were observed for the primary outcome; liver fat, as well as for body weight, HbA1c, LDL-cholesterol, total cholesterol, triglycerides and C-reactive protein (CRP). No statistically significant differences were observed for fasting glucose, HOMA-IR, high-density lipoprotein (HDL)-cholesterol, ApoA1, ApoB, insulin, FIB-4, SBP or DBP (Table 5).

Median differences in liver fat were 0.30 (95% CI: -0.52, 1.11) % for LCPUFA vs HND, -1.46 (-2.42, -0.51) % for LCPUFA vs UC and -1.76 (-2.96, -0.57) % for HND vs UC (Figure 10). Total cholesterol and LDL-cholesterol were lower in the LCPUFA group compared to the UC group whereas body weight, HbA1c, CRP, total cholesterol, LDL-cholesterol and triglycerides were lower in the HND group compared to UC (Table 5). Compared to LCPUFA, the HND group had a lower body weight and HbA1c after follow-
up (Table 5). Effect estimates for the primary and secondary outcomes were robust in sensitivity analyses.

![Figure 11](image-url). Within-group changes with corresponding 95% CI (left) and between-group ITT effect estimates with corresponding 95% CI (right) for the primary outcome. HND, Healthy Nordic Diet; LCPUFA, Low Carbohydrate Polyunsaturated Fatty Acids; UC, Usual Care.
<table>
<thead>
<tr>
<th>Outcome</th>
<th>LCPUFA</th>
<th>HND</th>
<th>UC</th>
<th>LCPUFA vs HND</th>
<th>LCPUFA vs UC</th>
<th>HND vs UC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body weight (kg)</td>
<td>-2.15</td>
<td>-4.61</td>
<td>-1.84</td>
<td>2.46</td>
<td>-0.31</td>
<td>-2.77</td>
</tr>
<tr>
<td></td>
<td>(-3.39, -0.91)</td>
<td>(-5.85, -3.36)</td>
<td>(-3.26, -0.42)</td>
<td>(0.71, 4.21)</td>
<td>(-2.18, 1.57)</td>
<td>(-4.64, -0.90)</td>
</tr>
<tr>
<td>HbA1c (mmol/mol)</td>
<td>0.36</td>
<td>-1.33</td>
<td>0.82</td>
<td>1.69</td>
<td>-0.46</td>
<td>-2.15</td>
</tr>
<tr>
<td></td>
<td>(-0.64, 1.36)</td>
<td>(-2.36, -0.29)</td>
<td>(-0.33, 1.97)</td>
<td>(0.24, 3.13)</td>
<td>(-1.96, 1.04)</td>
<td>(-3.67, -0.62)</td>
</tr>
<tr>
<td>CRP (mg/L)</td>
<td>-0.62</td>
<td>-0.88</td>
<td>-0.40</td>
<td>0.26</td>
<td>-0.22</td>
<td>-0.48</td>
</tr>
<tr>
<td></td>
<td>(-0.91, -0.33)</td>
<td>(-0.99, -0.76)</td>
<td>(-0.74, -0.05)</td>
<td>(-0.05, 0.57)</td>
<td>(-0.66, 0.22)</td>
<td>(-0.84, -0.12)</td>
</tr>
<tr>
<td>TC (mmol/L)</td>
<td>-0.30</td>
<td>-0.34</td>
<td>0.03</td>
<td>0.04</td>
<td>-0.32</td>
<td>-0.37</td>
</tr>
<tr>
<td></td>
<td>(-0.49, -0.11)</td>
<td>(-0.53, -0.15)</td>
<td>(-0.19, 0.24)</td>
<td>(-0.23, 0.32)</td>
<td>(-0.61, -0.04)</td>
<td>(-0.65, -0.08)</td>
</tr>
<tr>
<td>LDL-C (mmol/L)</td>
<td>-0.28</td>
<td>-0.30</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.28</td>
<td>-0.29</td>
</tr>
<tr>
<td></td>
<td>(-0.44, -0.12)</td>
<td>(-0.46, -0.14)</td>
<td>(-0.19, 0.17)</td>
<td>(-0.20, 0.24)</td>
<td>(-0.51, -0.04)</td>
<td>(-0.53, -0.06)</td>
</tr>
<tr>
<td>TG (mmol/L)</td>
<td>-0.30</td>
<td>-0.34</td>
<td>0.03</td>
<td>0.13</td>
<td>-0.24</td>
<td>-0.37</td>
</tr>
<tr>
<td></td>
<td>(-0.49, -0.11)</td>
<td>(-0.53, -0.15)</td>
<td>(-0.19, 0.24)</td>
<td>(-0.12, 0.38)</td>
<td>(-0.50, 0.02)</td>
<td>(-0.62, -0.11)</td>
</tr>
<tr>
<td>FPG (mmol/L)</td>
<td>-0.36</td>
<td>-0.60</td>
<td>-0.22</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(-0.59, -0.13)</td>
<td>(-0.83, -0.37)</td>
<td>(-0.48, 0.03)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>HOMA-IR</td>
<td>-0.79</td>
<td>-1.42</td>
<td>-0.48</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(-1.46, -0.12)</td>
<td>(-2.09, -0.74)</td>
<td>(-1.27, 0.31)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Metric</td>
<td>Estimated Mean</td>
<td>95% CI</td>
<td>Estimated Mean</td>
<td>95% CI</td>
<td>Estimated Mean</td>
<td>95% CI</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>HDL-C (mmol/L)</td>
<td>0.03 (0.00, 0.10)</td>
<td>0.04 (-0.01, 0.10)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ApoB (g/L)</td>
<td>-0.07 (-0.11, -0.02)</td>
<td>-0.02 (-0.07, 0.03)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ApoA1 (g/L)</td>
<td>-0.01 (-0.06, 0.02)</td>
<td>-0.02 (-0.07, 0.02)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Insulin (mE/L)</td>
<td>-3.24 (-5.02, -1.46)</td>
<td>-1.55 (-3.64, 0.54)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FIB-4</td>
<td>0.24 (0.08, 0.39)</td>
<td>0.07 (-0.10, 0.23)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SBP (mmHg)</td>
<td>-2.89 (-6.55, 0.77)</td>
<td>-5.82 (-9.91, -1.73)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DBP (mmHg)</td>
<td>-3.90 (-5.95, -1.86)</td>
<td>-3.92 (-6.24, -1.60)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

1Data are presented as estimated marginal means or medians with corresponding 95% CI. The table showcases results from secondary outcomes. For outcomes that were not statistically significant between groups in the overall general linear model test, no pairwise group intention-to-treat (ITT) effects were estimated. ApoA1, Apolipoprotein A1; ApoB, Apolipoprotein B; BW, Body Weight; CRP, C-Reactive Protein; DBP, Diastolic Blood Pressure; FPG, Fasting Plasma Glucose; HbA1c, Haemoglobin A1c; HDL, High-Density Lipoprotein; HOMA-IR, Homeostatic Model Assessment of Insulin Resistance; LDL-C, Low-Density Lipoprotein Cholesterol; TC, Total Cholesterol.
Discussion

Main findings
In paper I and II, serum CE 18:2n-6 and liver PL 22:6n-3, suggestive good biomarkers of dietary intake, were inversely associated with liver fat content and NAFLD fibrosis, respectively. In addition, several fatty acids and fatty acid ratios, primarily, but not exclusively, reflecting endogenous metabolism were associated with liver fat (i.e. serum CE 16:1n-7, 18:3n-6, 20:3n-6, 20:4n-6, D6D and SCD-1) and fibrosis (i.e. liver PL 22:0, plasma PL 22:0, liver TAG 18:1n-9 and plasma TAG 18:1n-9). Some degree of heterogeneity were noticeable over the different lipid fractions and between the compartments explored (blood vs liver) in paper II. Apart from 22:6n-3, associations between serum CE fatty acids and liver fat were not markedly altered when further adjusting for REE/FFM in paper I. In paper III, no clear (i.e. imprecise) associations were observed for any food or nutrient substitution. Replacing SFA with carbohydrates and milk with fermented milk were inversely associated with HCC, although with low precision and CIs covering one. In paper IV, liver fat content was reduced in both the HND group and the LCPUFA group compared to UC, with no between-group difference for HND versus LCPUFA. The HND improved several other cardiometabolic risk markers compared to both UC and the LCPUFA group whereas the LCPUFA group improved total cholesterol and LDL-cholesterol compared to UC.

Findings in relation to other studies
Fatty acids and liver fat
A few observational studies investigating associations between circulating fatty acids and liver fat content have been performed, largely in line with our findings [74,76,83]. However, these studies have generally been carried out in smaller samples and with less rigorous methods to assess fatty acid composition and liver fat content. To complement these biomarker-based studies, observational studies investigating the role of self-reported dietary fatty acids in NAFLD have also been conducted. Although findings on the composition of fatty acids on a nutrient-level is somewhat conflicting [135–139], SFA-rich foods such as red meat and processed red meat [139–141] and PUFA-rich
foods such as fish [142], nuts and seeds [143] have been associated with liver fat and NAFLD in both cross-sectional and longitudinal studies.

Whether the associations between PUFA and liver fat could be partially explained by an enhanced REE remains to be determined. In paper I, the association between 22:6n-3 and liver fat content was markedly attenuated after adjusting for REE/FFM, partly in line with human omega-3 supplementation studies showing an effect on energy expenditure [144].

Furthermore, shorter-term RCTs have repeatedly shown an increase in liver fat after consuming a SFA-rich diet compared with a PUFA-rich diet [7–9,86]. In line with results from paper IV, Holmer et al. reported a decrease in liver fat content over 12 weeks when contrasting a LCHF diet with standard of care in patients with NAFLD [103]. A large difference in body weight between the groups could have potentially explained the results. Noteworthy, the standard of care group decreased levels of fibrosis (using kPa from Fibroscan as a biomarker) in comparison to the LCHF diet group, suggesting that the LCHF diet may be beneficial for some aspects of the NAFLD spectra, but not others. In addition, Dalby-Hansen et al. reported an increase in LDL-cholesterol in the LCHF diet group when compared to a low-fat diet, although other cardiometabolic markers such as body weight and HbA1c improved [102]. In both of these trials, SFA increased dramatically compared to the comparator groups. As SFA have been shown to increase LDL-cholesterol and associate with more progressive forms of liver disease, such as HCC and chronic liver disease [145], substituting SFA with other fatty acid subclasses may be of relevance in the context of a LCHF diet. In paper IV, SFA decreased in the LCPUFA group by 1.3 E% compared to UC despite an increase in overall fat intake, possibly explaining the favourable results in LDL-cholesterol between the groups.

No study to date has investigated the long-term effects of a HND on liver fat content. However, whole-grain intake [110–112,146], rapeseed oil [147], fruits and vegetables [148] and fish intake [142,149], all of which are important components of a HND, have been suggested to play a beneficial role in NAFLD. In addition, comparing a HND with a habitual diet have shown beneficial effects on body weight and blood lipids, in line with findings from paper IV [107–109].

A 1.46-1.76 % absolute difference between the LCPUFA group and the UC group and the HND group and the UC group, respectively, corresponds to a relative reduction of 20-25% in this population. The clinical significance of such findings is important to discuss. In line with our results, Tamaki et al. demonstrated using paired liver biopsies in 100 NAFLD patients that a 25% decrease or more in MRI-quantified liver fat over 1.4 years of follow-up was associated with an increased odds of fibrosis regression compared to patients with less than a 25% decrease in liver fat [150]. This is further supported by a meta-analysis of RCTs showing that a 30% relative reduction in liver fat assessed using MRI is associated with improvements in NASH histology [151].
Extracting useful causal knowledge from these observations may however be difficult as change in liver fat may be influenced by various different treatments/exposures in a given population. This will be thoroughly discussed under methodological considerations below. Nevertheless, these findings may provide a rough indication of the importance of our findings. One should also note that isolated simple steatosis at baseline (defined by the presence of >5% liver fat in absolute numbers) is associated with a more rapid progression of fibrosis [152], increased risk of HCC and other cancer subtypes [153], all-cause mortality [21], cancer mortality [21], CVD mortality [21], death due to cirrhosis and HCC [21] as well as incidence of T2D [154]. An absolute reduction in liver fat of 1.46-1.76 %, as demonstrated in our study, may thus be of clinical importance for the prevention of such outcomes over long term, provided such differences are maintained.

Fatty acids and fibrosis, cirrhosis and HCC

With regards to fibrosis, cirrhosis and HCC, data on the composition of liver fatty acids are scarce. A general pattern observed is that patients with NAFL or NASH are characterized by PUFA-depleted and SFA-enriched livers, in line with our findings of the pooled fatty acids in paper II [77,79,155]. In addition, liver 22:6n-3 has been shown to be depleted in patients with NAFLD, further supporting our findings from paper II [77,79,155]. Data on the composition of dietary fatty acids in fibrosis are limited but a few randomized supplementation trials have been carried out using a combination of 20:5n-3 and 22:6n-3 in varying doses, showing conflicting results [156,157]. No study to date, of which we are aware of, has investigated 22:0 in relation to NAFLD. Interestingly, we found a positive association for liver PL 22:0 but an inverse association for plasma PL 22:0 and fibrosis. These findings are in part consistent with an inverse association of circulating 22:0 and T2D [158]. In addition, no correlation was observed between liver and plasma PL 22:0, suggesting that circulating very-long chain SFA might reflect extrahepatic tissue metabolism more so than the liver [82,159]. Lastly, although 18:1n-9 in both liver TAG and plasma TAG were inversely associated with liver fibrosis in our study, these associations are somewhat difficult to interpret. Circulating 18:1n-9 might be reflecting foods such as olive oil and rapeseed oil [67,160], but is generally regarded as a relatively poor biomarker of dietary intake and has been positively associated with both incidence of T2D and prevalence of liver fat in other studies [83,161]. On the other hand, these studies have assessed 18:1n-9 in either PL or CE and less so in TAG. Results may therefore differ depending on what lipid fraction (and compartment) is used to estimate these associations, and is important to reconcile before extrapolating results from one fraction to another [155,161].

In paper III, substituting SFA with carbohydrates showed a HR of 0.87 for HCC (although with a wide CI covering one (95% CI: 0.74-1.02), partly in
line with observational studies investigating high versus low consumption of SFA and RCTs investigating short-term effects on liver fat from such replacement [105,145]. Although accompanied with low precision, substituting milk with fermented milk showed a HR of 0.93 (95% CI: 0.85, 1.01) on the incidence of HCC, partly in line with observational studies comparing high versus low consumers of yoghurt [162]. The isolated comparison of milk with yoghurt on liver fat, a risk marker for HCC, have however shown conflicting results [163,164]. As of today, no other studies have investigated associations of diet in relation to NAFLD-specific cirrhosis, making it difficult to compare our findings to the existing literature. However, lack of an association of replacing unprocessed red meat with fish in our study aligns with another study on total NAFLD as the outcome [165]. In a recent RCT, replacing unprocessed red meat with freshwater fish reduced liver fat by 3% in absolute numbers in individuals with NAFLD over 84 days of follow-up, despite no difference in body weight between the groups [149]. No assessment of fibrosis progression was however conducted.

Methodological considerations

There are several important methodological considerations that need to be thoroughly discussed. DAGs will be used to illustrate some of the potential issues from each paper. These causal structures might be helpful when discussing complex issues such as confounding and selection bias. For a summary of DAGs, please see Figure 3.

Outcome assessment

MRI and histological grading from liver biopsies were used to assess liver fat and NAFLD fibrosis in paper I and IV and paper II, respectively. MRI is a non-invasive procedure which has been validated against liver histology and shown to be both reliable and accurate in detecting smaller quantities of liver fat [37,166]. Although non-invasive, MRI may be experienced as claustrophobic by some and lead to non-participation or loss-to follow up. Histological grading from liver biopsies on the other hand is considered the gold standard method for assessing NAFLD fibrosis. However, this method is both invasive and costly and may introduce some degree of misclassification due to sampling variability, as shown by Ratziu et al. in 2005, where discordance in fibrosis staging of one stage or more between two liver biopsies was 41% in patients with NAFLD [38]. Newer methods have been developed to evaluate fibrosis, such as magnetic resonance elastography (MRE), TE and clinical biomarker scores (e.g. FIB-4). However, liver biopsy grading remains the gold standard method [166].
In paper III, HCC was defined using ICD codes from the Swedish Cancer Register and the Cause of Death Register. The Swedish Cancer Register has high coverage of most cancer cases, although underreporting may vary depending on cancer site [167]. Assuming non-differential sensitivity with high specificity, any measurement error in the assessment of HCC would probably bias the effect estimates (on the relative scale) weakly towards the null [168]. As a caveat, HCC reporting in the Swedish Cancer Register has been shown to be underestimated when compared to the NPR and the Cause of Death Register [169]. Including codes from the NPR in a sensitivity analysis did not change the interpretation of our findings. The same reasoning on the direction of bias can be applied to the primary outcome; NAFLD cirrhosis, although with the risk of introducing somewhat larger bias towards the null due to potentially more false positives from alcohol-related cirrhosis not diagnosed as alcoholic liver disease (ALD). However, several approaches were taken to reduce these risks, including the exclusion of high alcohol-consumers at baseline and by only counting cirrhosis cases that had any cardiometabolic disease or a BMI>25 kg/m² at baseline as a sensitivity analysis, showing moderately robust results compared to the primary analysis. A combination of bias towards the null, arguably weak, and a relatively high proportion of undiagnosed cases of compensated cirrhosis may potentially have contributed to our null findings in this study, and should thus be considered when interpreting our findings. One should however not dismiss the fact that the specified isolated nutrient and/or food substitutions of moderate magnitude in this study may not have been enough of a contrast to observe differences in the risk of the outcomes.

In paper IV, any measurement error in liver fat content would probably be non-differential as outcome assessors were blinded to the assigned diet group, thereby making the risk of detection bias low when estimating the ITT effect [170]. Missing outcome data at follow-up is an issue as drop-outs may be differential with regards to the allocated diet. However, missing data was dealt with using multiple imputation, with the assumption that data was MAR and that the imputation model was correctly specified. A sensitivity analysis whereby another imputation model was used showed similar estimates, suggesting that missing outcomes were accurately accounted for. Missing outcome data for all secondary outcomes was low (5-7 %) but a bit higher for liver fat (14 %).

Exposure assessment
The composition of fatty acids in paper I and II was assessed using GC, which is considered the gold standard approach to measure fatty acids. In addition, multiple fatty acids over several lipid fractions were measured in both cohorts, making it somewhat easier to compare our findings to other studies. Analyzing fatty acids in separate lipid fractions instead of whole plasma/serum is further
regarded as a strength due to the high degree of variation of lipid fractions (and therefore fatty acid composition) among circulating lipoproteins [54].

A somewhat trivial, but nonetheless important, methodological matter to consider when interpreting findings from fatty acid biomarkers is their compositional nature. Tissue and plasma fatty acids are expressed as proportions (area percentages) of total fatty acids measured in the sample. Collectively therefore, fatty acids make up 100% in total, transforming it into a composite of all fatty acid proportions contributing to that pool. By conditioning on a composite variable (by restricting that variable to 100%), and hence a collider, all other backdoor paths are opened [171]. Hence, a joint relative substitution effect of increasing fatty acid x (e.g. 18:2n-6) relative to the weighted average effect of decreasing all other fatty acids (e.g. 15:0, 16:0, 22:6n-3 etc.) is estimated and must be considered when interpreting findings from our studies. This is illustrated in the DAG below (Figure 12).

Other potential methodological issues of using circulating fatty acids to reflect dietary intake in observational settings of aetiological questions include 1) vaguely defined causal questions, 2) confounding bias following from point 1 and 3) issues with interpretation and extension of effect estimates [172]. The proportion of circulating fatty acids, including essential fatty acids (although to a lesser degree than non-essential fatty acids), are influenced by many internal and external factors (including different foods, physical activity, smoking, underlying diseases, genes and unknowns). There are thus multiple possible interventions, with arguably different effects on the outcome, which could take place in a real life setting, leading to difficulties in articulating a properly formed causal research question that may inform public health decision making. From this follows the issue of confounding bias. If some of these factors are unknown (e.g. genes affecting lipid metabolism), confounding bias will be introduced if these same factors are causes of the outcome. This is certainly the case for more well-defined exposures as well, but the choice to eat 100 grams of fish instead of 100 grams of meat may not depend on complex genetic traits influencing e.g. weight, but rather downstream effects of such genes, i.e. increases in BMI. Adjusting for BMI would thus be sufficient to close those backdoor paths. Secondly, if the goal is to estimate the average causal effect on the outcome of interventions to increase a particular fatty acid in plasma by changing the diet, one would need to adjust for all common causes of diet and the outcome, in addition to all common causes of the fatty acid and the outcome (of which all are probably not known). The same reasoning applies to if one wishes to estimate the weighted average causal effect from all possible interventions affecting the proportion of the fatty acid in circulation/tissue [173]. Given conditional exchangeability holds, this estimand is a possibility to target for complex exposures representing multiple versions of treatment. The interpretation and generalizability of such estimate is however not straightforward.
Vaguely defined exposures make it more difficult to truly characterize all confounding factors necessary to estimate the effect of interest. This may be of particular concern when it comes to circulating fatty acids, due to their compositional nature (i.e. confounding factors for all fatty acids summing up to 100% must be identified and not just those for your fatty acid of interest). The above has been argued extensively before in the context of BMI as an exposure in observational studies [172].

How non-directly manipulative exposures should be viewed and/or used to inform decision making is under intense debate in the research community. Nevertheless, additional careful consideration as outlined above should be given when asking causal questions of complex exposures, such as fatty acid biomarkers, using observational data. As an additional caveat, even if we have done our best to estimate the effect from changing e.g. plasma 18:2n-6 by diet, we are still left with many of the issues that traditional nutritional epidemiology has been criticized for, namely what food(s) to intervene on (shifting the focus from nutrients to foods) and what the comparator food(s) would be. There is probably a difference in outcome if we compare cookies baked on margarine with carrots or walnuts with carrots. This is an area in the field of nutrition in which more discussion should take place. Viewing findings from causal questions of objective dietary biomarkers as a contributing piece to the larger evidence puzzle is probably a sensible approach to have.

Leaving this aside, biomarkers such as fatty acids may be used in other settings, such as disease prediction, causal mediation and as complementary to self-reported adherence measures in randomized trials. Further discussions regarding exposure assessment in paper III will be discussed under the section “measurement bias”.
Figure 12. A simplified directed acyclic graph (DAG) illustrating the compositional nature of fatty acids. By conditioning on the collider “100%” (by restricting that variable to 100%), the backdoor path “18:2n-6 (%) → 100% ← All other FA (%) → Outcome” opens up. Because of this, a joint average relative substitution effect is estimated. C1 is a confounder of 18:2n-6 and the outcome, C2 is a confounder for both 18:2n-6 and all other FA and the outcome and C3 is a confounder for all other FA and the outcome. C* refers to any factor influencing C1-C3 and the outcome. In this example C* is a common effect of C1-C3 and the outcome but C1-C3 may also have different factors causing them. This DAG illustrates the complexities of using vaguely defined exposures in observational studies, in particular compositional exposures such as circulating or tissue fatty acids.

What is the comparator?
Adjusting for total energy intake (TEI) when investigating diet-disease relations is often justified based on two primary reasons. Firstly, energy intake is strongly correlated with many physiological factors, such as body size and metabolic efficiency. TEI may therefore work as an overall proxy for these, most often, unmeasured but also unknown determinants of diet composition [174]. Secondly, separating the effect of single foods from overall energy intake requires isocaloric diet groups in RCTs. Adjusting for TEI in observational studies transforms the research question into a question of isocaloric substitutions. However, there are several practical and methodological issues with traditional adjustment strategies that need to be addressed. This was eloquently pointed out by Tomova, et al. in 2022 [175]. Due to the compositional nature of diet (as for fatty acids described above), conditioning on TEI opens up backdoor paths between the exposure, all other energy-providing foods/nutrients and the outcome of interest. An unspecified substitution effect of increasing the food/nutrient of interest by one unit while decreasing all other energy-providing foods/nutrients by one unit as a weighted average in the
given population is therefore estimated. The effect observed is thus highly dependent on the background distribution of the diet in that specific population. This will be further discussed under the headline “External validity”. A somewhat vague research question of this kind would probably not be that informative if asked in a strictly controlled RCT, so why are we persistently asking these research questions in observational settings? [176,177]. Secondly, adjusting for total energy intake as a composite variable may introduce both composite variable bias (a bias introduced when separate variables carrying information are aggregated into one) and residual confounding bias (bias introduced when each diet component has distinct effects on the outcome) [175]. Using other energy adjustment models may help to solve some of the issues described above.

In paper III, we investigated prespecified isocaloric food and nutrient substitutions, using the leave-one-out model [130]. The leave-one-out model has been shown to produce similar unbiased estimates as the energy partition model (another common model used in statistical substitution analyses), provided that no more than one dietary component is substituted and that the same unit is used for both the composite variable and all dietary components [130]. Using different units (e.g. grams for the exposure while adjusting for total energy intake in kcal) is common but was shown to produce biased estimates in simulation studies by Tomova et al., sometimes in the opposite direction of the true simulated causal effect [130]. Some ambiguity still exists on the degree of bias introduced when mixing units or when only adjusting for some but not all foods in real world cohort data [178]. Further studies are needed to elucidate the importance of such remaining questions related to substitution modelling approaches. However, despite these uncertainties, the bigger issue still remains - what are we comparing our exposure of interest with?

In 2022, Tobias DK. wrote an article in Circulation with the catchy title of “What Eggsactly Are We Asking Here? Unscrambling the Epidemiology of Eggs, Cholesterol, and Mortality”, emphasizing the importance of specifying the comparator food in nutritional epidemiology, as would have been done in a RCT [177]. Tobias argued that contrasting eggs with cinnamon rolls or with steel cut oatmeal with fruits would probably result in different magnitudes of effects. Rightly so, the article used data from a published cohort study of egg intake that showed different effect sizes (and even reversal of signs) when eggs were compared with all foods from the background diet versus when eggs were partly compared with other cholesterol-rich foods such as butter, red meat and dairy. The latter substitution was introduced when researchers further adjusted for total dietary cholesterol in their analyses. Clearly specifying the substitution of interest using statistical substitution models may help researchers to think about the comparator in greater detail [130,177,179]. The concept of defining your comparator food(s) has also been argued to be extended to the field of evidence synthesis [180]. Pooling estimates from similar research questions of diet and liver disease would certainly be of interest as
incidence rates of HCC and NAFLD cirrhosis are relatively low in the general population, leading to less precision, as suggested for the vast majority of substitutions in paper III. One should however bear in mind that substitution modelling approaches don’t solve every single problem in nutritional epidemiology. There are many methodological issues, discussed in detail below, that apply to them as well.

Reverse causation

Due to its temporal component, cohort studies are preferable study designs when trying to estimate causal effects using observational data. In paper I and II, we used cross-sectional data to investigate associations between circulating and liver fatty acids and liver fat and NAFLD fibrosis. This particular study design limits us from making causal interpretations from our findings and is regarded as the biggest limitation of the two studies. Cross-sectional data are prone to reverse causation and must be carefully considered.

The issue of reverse causation may still be present in longitudinal studies, if the exposure is influenced by an underlying subclinical disease state that is also associated with the clinical outcome. Reverse causation is here a type of confounding bias. In paper III, we investigated this by excluding the first two years of follow-up, providing similar estimates as for the primary analysis. In paper IV, we estimated the ITT-effect, thereby preventing reverse causation to occur as any underlying disease is unrelated to randomization at baseline.

Confounding

Confounding is introduced when backdoor paths, that should otherwise be closed, remain open [181]. This may lead to the wrong conclusions being drawn on the causal relationship between the exposure and the outcome. Confounding paths are closed when conditioning on confounding variables and there are several ways of which these may be identified. In paper I, we identified potential confounders based on subject-matter knowledge and the background literature. In paper II and III (and IV but only for the causal mediation analysis), we identified confounding paths based on subject-matter knowledge and DAGs [127]. It is possible that other variables would have been adjusted for in paper I if a more causal framework would have been used.

Although causal diagrams are useful for the identification of biasing paths, unknown or unmeasured variables may still be present, potentially introducing unmeasured confounding bias. In paper II, we did not have data on diet and diet was therefore indicated as an unobserved variable in our DAG. Due to this, unmeasured confounding may have biased some of the associations between liver and plasma fatty acids and NAFLD fibrosis. However, as BMI was depicted as an effect of diet and a cause of fibrosis in our DAG, some of the variation in the outcome explained by diet might have been captured
through BMI. Furthermore, whether dietary components should be adjusted for or not depend on the specific research question under investigation (please see the reasoning under exposure assessment). Confounding structures probably also differ between the different fatty acids as well (e.g. fish intake might be a confounder for 16:0 and must be adjusted for but maybe not for 22:6n-3 if 22:6n-3 is used to reflect fish intake). Using biomarkers as exposure variables in observational studies is a complex matter which deserves a lot more discussion in the nutrition space.

Several sensitivity analyses approaches may be employed to investigate the potential for unobserved or residual confounding bias. In paper III, previous diet was identified as an unobserved confounder on the pathway of baseline diet and NAFLD cirrhosis and HCC. To address this question we chose to update our baseline to 2009, further adjusting for pre-baseline diet in 1997. As it turns out, adjusting for previous diet had little impact on the estimates. There are at least three potential explanations for this: 1) confounding from previous diet is not that strong for this particular research question, as indicated by other studies [182,183], 2) the time interval between previous diet and baseline diet was far too long to capture the confounding, as argued before [182,183] and 3) effect estimates may be biased due to measurement error in self-reported diet, hence introducing residual confounding bias. In paper III, we also further subcategorized the category named “other foods” in order to examine the potential for residual confounding bias by grouping foods with different effects on the outcome together. No change in estimates were observed.

Selection bias

Selection bias may be introduced when a selected group of participants are either recruited into the study, are lost to follow-up or are restricted upon after outcome ascertainment. Another way of defining selection bias (at least under the null) is the bias introduced when conditioning on a common effect of two variables, of which one is either the exposure or a cause of the exposure and the other is either the outcome or a cause of the outcome [184]. Selection bias may lead to distorted non-causal associations between the exposure and outcome that are different from the population eligible to participate (i.e. the source population). In some cases, this may lead to the direction of association being inverted [185]. In paper I, n=2008 individuals were randomly invited to participate in the POEM study one month after their 50th birthday. Of these, only n=502 individuals signed a written informed consent (25%) and n=310 underwent MRI scanning. This high rate of non-participation may have introduced selection bias in the form of participation bias/self-selection bias [184,186]. The same applies to paper II and III. Available data on participants who did not participate might have been useful to tackle these issues by weighting the sample to the distribution of important covariates determining participation in the source population [186]. However, these data were not
available for either cohort. Although selection bias is perhaps inevitable, the important question is not if but by how much this bias distorts the associations. In the UK biobank cohort, low participation rates have been shown to cause severe selection bias for certain exposure-outcome associations, but not for others [187]. One should however bear in mind that the UK biobank cohort had a 5.5% response rate which is considerably lower than participation rates in paper I-III. Further adjusting for variables potentially associated with participation (e.g. age, BMI and education) might have helped reduce some of this bias.

Prospective cohort studies, such as in paper III, are also prone to differential loss to follow-up, although this was mitigated by linkage of each individual’s personal identification number to different patient registries. In paper IV, differential loss to follow-up was dealt with using multiple imputation as described under outcome assessment. Loss to follow-up was also low in paper IV.

Missing data on confounders may also pose a problem of selection bias. Restricting the analysis to those with complete data (i.e. using listwise deletion), as in paper I, is a very common approach but may lead to biased estimates in either direction if the missingness mechanism is not completely at random (MCAR) [188–190]. Even if operating under the assumption of MCAR, efficiency may be hampered if the proportion of missing data is substantial. In paper II, no missing data on confounders were present. In paper III, we used multiple imputation to account for this missing data, assuming that missingness was MAR (or MCAR) and that the imputation model was correctly specified.

**Mediation**

In paper I, we investigated potential associations between fatty acids and liver fat, mediated through REE/FFM, using the principles of the difference method [191]. The relationship between 22:6n-3 and liver fat was found to be partly mediated through an increase in REE/FFM, as indicated by the change in the \( \beta \)-coefficient when including the mediator in the regression model. However, there are several potential issues of using conventional mediation approaches in this way, of which were not known or considered at the time of paper I. First, conditioning on a mediator (in our case REE/FFM) might lead to collider bias if there is mediator-outcome confounding present. Second, if the mediator-outcome confounders are themselves influenced by the exposure, this will lead to both incorrect estimates of the direct and indirect effects and potential collider bias if there are unknown or unmeasured confounders (U) of the mediator-outcome confounders present. As such, there are several assumptions that need to hold when performing mediation analyses [191]. Although this was considered a secondary analysis of the study and no strong conclusions were drawn based on the findings, a more *a priori* defined causal structure
would have been helpful to elucidate these assumptions and perhaps mitigate some of the potential biases. One must also bear in mind that this was a cross-sectional analysis for which temporality between the exposure, mediator and outcome were not known, limiting our ability to draw any firm conclusions of mediation from these findings.

In paper IV, we used a counterfactual approach to mediation whereby we estimated the proportion of the total effect of diet on liver fat that was mediated through a change in body weight [134]. This was done by dividing the natural indirect effect by the total effect. We included potential mediator-outcome confounders and assumed no treatment-confounding feedback by diet and mediator-outcome confounders. We furthermore specified an interaction between the diet groups and change in body weight. Although accompanied by very low precision, we estimated that 21-74 % of the total effects from the three different diet comparisons was explained by changes in body weight. We also compared these estimates to more traditional approaches to mediation commonly used in dietary RCT (i.e. only including the mediator in the regression model or including both the mediator and mediator-outcome confounders but not an interaction between the exposure and the mediator), demonstrating relatively minor departures from the estimates of the counterfactual approach.

Measurement bias

Some errors in the measurement of the exposure, outcome and/or confounders are mostly unavoidable when conducting research [192]. Errors are often categorized as dependent or independent and differential or non-differential. Dependent errors refer to errors of two variables that are correlated. Differential measurement error refers to an error in one variable that is dependent on the true value of a second variable. Measurement errors generally introduce bias, except in cases where errors are independent and non-differential under the null [192]. Measurement bias has been touched upon under “Exposure assessment”, “Outcome assessment” and “Confounding” but new perspectives will be introduced and discussed thoroughly under this subheading.

Using biomarkers to reflect dietary intake may have mitigated some issues of measurement error from self-reported data in paper I and II. Importantly however, measurement errors can also occur for the confounding variables and hence lead to bias. In paper I, we adjusted our model for carbohydrates, fat, alcohol and vitamin E. Assuming the errors are independent and non-differential, any measurement error in these variables would bias the estimates in the direction of the crude association. Using simple rules of signed causal DAGs, if carbohydrates negatively impact the proportion of CE 18:2n-6 and negatively impact liver fat content, any independent non-differential mismeasurement in self-reported carbohydrate intake would probably bias the estimate upwards. The strength of bias from residual confounding depends on the degree of mismeasurement, of which remains unknown. Using validated FFQs
may however help reduce the impact of mismeasured dietary variables, as was
done in paper I (at least for an earlier version of the same FFQ) and III. An
important strength in paper I was the adjustment for thoroughly measured var-
iables such as VO₂\text{max} (a proxy for exercise intensity) and BMI (assessed by
research personnel), thereby eliminating residual confounding from self-re-
ported data.

In paper II, fibrosis was assessed from liver biopsies which are prone to
sampling variability and hence measurement error as discussed under “Out-
come assessment”. Furthermore, using two independent pathologists to clas-
sify significant fibrosis may have helped reduce any variability associated
with the histological evaluation of the biopsies. Lastly, as the exposure (liver
and plasma fatty acids), outcome (fibrosis from liver biopsies) and confound-
ers were measured using different assessment methods/tools, dependent mis-
classification of disease may possibly be ruled out.

In paper III, we adjusted for total energy intake as well as other nutri-
ents/food items to mitigate some of the potential issues with dependent meas-
urement error [193]. We however treated diet as a time-fixed exposure using
only baseline diet from 1997. This may have introduced some degree of meas-
urement error of long-term adherence as individuals may change diet over
time. Nevertheless, the use of traditional regression-based approaches to ac-
count for time-varying exposures and confounders may introduce collider bias
in the presence of exposure-confounding feedback (e.g. from diet and BMI)
[194]. As of this, while exposure mismeasurement may indeed be reduced by
incorporating time-updated information, bias from conditioning on a post-
baseline confounder influenced by past exposure may instead be introduced
[194]. Using more rigorous statistical approaches to properly deal with time-
varying covariates in the setting of exposure-confounding feedback (i.e. g-
methods) is warranted in future studies. The latter methods are however more
complex and time-consuming.

BMI and other confounders were self-reported in paper III, leading to the
potential of measurement error and residual confounding bias. As for paper I
discussed above, assuming independent non-differential mismeasurement of
self-reported confounders would probably bias the estimates in the direction
of the crude associations. However, whether independent measurement errors
can be expected given that the exposure and all confounders were captured
using self-reported questionnaires, is only speculative. Dependent measure-
ment errors of the confounders and the exposure may bias the estimate in any
direction [195]. One may argue that individuals who perceive themselves as
less health-conscious underestimate both their weight as well as their intake
of certain snack foods, when assessed using self-reported questionnaires.
Whether the degree of underreporting from foods used in paper III (e.g. fer-
mented milk) is as high as argued for foods rich in added sugar and fat is
perhaps less likely. Using different assessment methods or tools to measure
the exposure, outcome and the confounders are generally advised, although sometimes unfeasible in larger cohort studies as those in paper III.

In paper IV, self-reported data from 4-days WFD were complemented with data from biomarkers of dietary intake in a subgroup of participants, showing concordant results. Outcome misclassification and mismeasurement for paper III and IV has been discussed under “Outcome assessment”.

External validity

External validity is an umbrella term encompassing both generalizability and transportability [196]. Generalizability refers to the ability to extrapolate your estimates to the a priori defined target population from which the study sample was drawn (e.g. middle-aged to elderly men and women living in Sweden in paper III). Transportability refers to the ability of applying your estimates to other target populations of which your sample is not part of (e.g. middle-aged to elderly men and women living in Norway in paper III). The ability to extend ones findings to other populations are mainly determined by the distribution of effect modifiers, version of treatments and interference patterns between populations. As discussed previously, proportions of circulating and tissue fatty acids are influenced by many different internal and external factors, contributing to the possibility of multiple versions of treatments/interventions. If the distribution of these “interventions” in paper I and II differ between populations, estimates may not be generalizable or transportable. As fatty acids are also expressed in proportions, different proportional distributions may also hamper our ability to extrapolate our findings from paper I and II to other populations.

On the contrary, in paper III we investigated clearly specified nutrient and food substitutions which may have mitigated some of the issues of external validity. This is particularly the case for food substitutions but may not apply to nutrient substitutions as nutrients reflect many different foods that may vary among populations. Due to few events, we were however not able to investigate any effect measure modification in this study. Nevertheless, participants in the SIMPLER cohorts were comparable to the general Swedish population at baseline with regards to age, education and BMI [114].

In paper IV, potential effect modification on the additive scale was examined using single-variable stratification, showing some differences between strata of T2D status, NAFLD status and PNPLA3 I148M polymorphism. Differences in liver fat between groups were statistically significant for individuals with T2D, those without NAFLD and those with the CC genotype in the PNPLA3 I148M polymorphism, with pairwise comparison estimates in the same direction as the primary analyses. These findings suggest that extension of effect estimates from the trial to the target population or other populations may be somewhat hampered, contingent that the distribution of effect modifiers differ between the two. Unfortunately, we do not have comparable data on
all these measures on individuals with dysregulated glucose metabolism in Sweden. Nevertheless, T2D prevalence will arguably be lower in a broader Swedish population of individuals with dysregulated glucose metabolism, and hence lead to attenuated differences in liver fat in that population. Importantly, effect heterogeneity can only be investigated for those variables that have been measured and determined a priori, although new machine learning techniques have been developed to capture effect modifying strata defined by many variables [197]. The latter is important as individuals may belong to different strata with different effects, which may have implications for how we choose to treat that individual. As such, effect modification from other variables may still be present, potentially impacting generalizability and transportability of our effect estimates. One should also notice that potential differences between strata may be explained by differences in study sample sizes, thereby limiting statistical power. Furthermore, as the intensities of the interventions in the trial likely differ from interventions implemented in clinical care, absolute differences between groups may be larger than for the general patient group. The latter could have been prevented if a strict pragmatic trial would have been carried out. However, benefits of a pragmatic trial with regards to external validity should be weighed against the drawbacks of lower adherence and higher drop-out rates. Extension of ITT effect estimates, which are dependent on study-specific adherence, will be addressed below.

ITT vs per-protocol – what do we really care about?

In paper IV, the primary a priori determined estimand of interest was the ITT effect. The ITT effect estimate is an estimate of the effect of diet assignment (or baseline randomization) and not the actual diet received and adhered to over time (Figure 13). As baseline randomization is conserved, given no differential loss to follow-up, the ITT effect is most often the primary effect of interest in trials of dietary exposures. An additional argument for the ITT effect over the per-protocol effect is that the ITT estimate reflects the effect in real world conditions whereas the per-protocol estimate reflects the effect under ideal circumstances. However, one should bear in mind that the ITT effect is the effect of assignment under study-specific adherence. If adherence to the diet is different outside of the research clinic (due to e.g. different intervention intensities or if patients learn that a specific diet works from the results of that study), the effect will differ from the effect observed in the trial [198]. The ability to transport ITT effect estimates to other populations may thus be hampered. In addition, from a patient perspective, knowing the effect on the outcome had you fully complied to the protocol may be of more practical interest than knowing what would have happened had you been randomized to a specific diet. However, estimating the non-naïve per-protocol effect in a RCT is difficult and relies on several strong assumptions, such as no unmeasured confounding pre- and post-baseline (Figure 13). An analysis whereby the per-
protocol effect is estimated should therefore be regarded as an observational study and thus needs careful consideration. Using appropriate statistical methods to account for factors influencing censoring due to non-adherence would make it possible to estimate the effect had everybody adhered to their assigned diet group over follow-up. This was done in the updated version of the PREDIMED study, showing a larger per-protocol effect than ITT-effect [199] as well as planning to be done in the PREDIMED-plus trial [200]. In paper IV, this was not done as our pre-registered protocol did not specify any cut-offs for adherence on an individual level. In addition, statistical methods to validly estimate the per-protocol effect were not known at the time of study planning. The “per-protocol” effect in our study should thus be interpreted as a complete-case analysis. However, on a group-level, we observed large exposure contrasts in nutrients and foods between groups, suggesting good average adherence to the diets. If completely adherent to the protocol, then the ITT effect is equal to the per-protocol effect [198]. Whether or not this is the case requires a priori determined cut-offs for adherence (preferably a combination of self-reported measures and more objective biomarkers), strong assumptions, collection of pre- and post-baseline variables predictive of protocol adherence and prognostic of the outcome and appropriate statistical methods to account for treatment-confounder feedback.
Figure 13. A directed acyclic graph (DAG) illustrating the two different estimands that are often specified in RCTs. In figure A) the estimand is the intention-to-treat (ITT) effect. This estimand targets the effect of assignment to diets, i.e. baseline randomization. Because of this random element, exchangeability is expected to hold and no arrows are directed into the node “Randomization”. In figure B) the estimand is the per-protocol effect. This estimand targets the effect of sustained adherence to diets over the entire study period. As adherence to diet at month 0 (M0) and month 6 (M6) are probably not completely random, common causes of these nodes are included (i.e. Confounders (M0) and Confounders (M6)). In addition, adherence to diets at baseline may affect confounders at month 6, which is why an arrow from “Adherence to diet (M0)” to “Confounders (M6)” is included. This is a case of treatment-confounder feedback. If traditional regression-based approaches to adjustment would be carried out, by adjusting for the confounders at M0 and M6, one will block the open pathway from “Adherence to diet (M0)” to “Confounders (M6)” to “Liver fat (M12)”. In addition, collider bias would be introduced from opening the pathway “Adherence to diet (M0)” to “Confounders (M6)” to “U” to “Liver fat (M12)”. U could be any variable influencing the outcome and confounders at month 0 and month 6. For a valid unbiased estimate of the per-protocol effect, one would need other statistical methods, such as g-methods to account for the issue of treatment-confounder feedback.
Conclusion

Findings from this thesis suggest that higher dietary intakes of unsaturated fatty acids (in particular PUFA) and lower intakes of SFA may be of importance for the prevention and treatment of NAFLD (at least for liver fat and fibrosis). The conclusion is supported by the inverse association between 18:2n-6 and liver fat content in paper I, liver 22:6n-3 and NAFLD fibrosis in paper II and the effects on liver fat reduction from being randomized to a LCPUFA diet or a HND, compared to UC, in paper IV. Effects from the latter are probably results from a combination of dietary changes, but dietary fat quality may be an important component. Whether isolated substitutions of nutrients or foods may impact more severe stages of NAFLD or consequences of NAFLD requires well designed and large cohort studies to be conducted. Mechanisms by which dietary fatty acids impact the different stages of NAFLD remain elusive, but changes in energy metabolism is probably not the most important mechanism (with the possible exception of 22:6n-3), as indicated in a population-based sample in paper I and supported by other short-term randomized studies in the field. Further, circulating and liver tissue fatty acids and fatty acid ratios that reflect both diet and endogenous metabolism were associated with both liver fat and NAFLD fibrosis. Whether proportions from the latter fatty acids are causes of NAFLD or consequences of an altered metabolism in NAFLD remain a question of interest.
Clinical and societal implications

As a whole, findings from this thesis contribute to potentially important information that may be used as an evidence base for dietary guidelines to both the wider population as well as the patient population. The use of dietary biomarkers from paper I and II give some support to the idea that dietary PUFA play an important role in NAFLD. Results from paper IV provide novel information with important clinical implications of a low-carbohydrate high PUFA diet and a whole-grain rich Nordic diet that may be used as a foundation for new updated dietary guidelines for patients with T2D and NAFLD. Results from paper III are however more difficult to interpret but potentially indicate that SFA (in place of carbohydrates) may be less healthful for the development of HCC, although this may vary among SFA sources (i.e. milk versus fermented milk).

Findings from this thesis also lay the ground for future studies to be conducted to further disentangle underlying mechanisms of both dietary and non-dietary fatty acids in NAFLD pathogenesis. Would increasing peanuts that are relatively rich in 22:0 enrich the plasma and livers of this fatty acid in patients with NAFLD? Would this translate to improvements in fibrosis or should we focus on targeting pathways involved in the metabolism of 22:0 instead, using pharmacological treatments? This is just one example of the potential for clinical implications following from results from paper II in this thesis.
Future perspectives

Observational population-based studies investigating dietary fat quality in relation to NAFLD and NAFLD-related complications/diseases such as HCC are necessary. Although statistical substitution models may be one way forward to clearly specifying the comparator food or nutrient of interest, a transition to a more wholesome dietary pattern approach (such as the healthy Nordic diet) may be of higher relevance for future public health decision-making. Using conventional approaches of comparing high versus low adherers to different diets may be of some utility, but these approaches are limited by at least three key points: 1) relevant public health exposure comparisons 2) estimation of the effect of sustained adherence to the diet over time and 3) allowing for adjusting for pre-baseline dietary exposures. The first key point is important as traditional approaches to the epidemiology of dietary patterns compare high versus low consumers and not sustained adherence to a dietary pattern versus the habitual diet or another dietary pattern over time. If the goal of Swedish public health recommendations is to shift the Swedish population from their current habitual diet to a more healthy diet, comparing high versus low consumers is probably not the most efficient way to approach observational studies. The second point refers to the estimation of the observational analogue of the per-protocol effect (i.e. the effect of sustained adherence over follow-up) that cannot be validly estimated using traditional regression-based approaches to adjustment, as outlined above in the thesis. The third key point refers to the fact that as diet has no exposure-naive individuals (as compared to medications in pharmacological epidemiology), time zero is difficult to define and previous diet may therefore confound any association between the exposure and the outcome. The same limitations (at least point 2 and 3) apply to statistical substitution models as well. Using a target trial framework whereby one specifies the hypothetical trial protocol one wishes to conduct and then use observational data to emulate such protocol may be a way to deal with some of the issues above [201]. This approach also clearly specifies the population of interest (through eligibility criteria), exposure contrasts and the causal estimand one wishes to target. A target trial approach has not been used in the setting of diet and liver diseases, although a few have been performed in the context of CVD [182,202].
RCTs of longer-term follow-up than one year on dietary fat quality with outcome data on liver inflammation and fibrosis, using liver biopsies or radiological assessment of fibrosis, would be of interest to conduct. This is a less studied area in the field of nutrition and liver disease. Different stages in the NAFLD spectra may (or may not) be influenced by dietary changes differently. The latter is an interesting aspect and may pave the way for a more targeted approach to diet and liver disease. Furthermore, as adherence to assigned diets wean over time (especially if long-term pragmatic studies on NASH development/regression are carried out), influencing the size of the ITT-effect, it would also be of interest for future studies to estimate the per-protocol effect as a complement to the ITT-effect.

In conclusion, although novel findings from this thesis contribute to new knowledge about the relation between fatty acids and NAFLD, important research questions remain, as outlined above. Combining these questions with methodological advances in the field of causal inference would progress the field of diet and NAFLD, hopefully leading to strong evidence-based dietary guidelines for both the wider population and the patient population.
Icke-alkoholinducerad fettleversjukdom (NAFLD) är ett spektrum av leversjukdomar som innefattar fettlever utan inflammation till fettlever med inflammation, med eller utan tillhörande ärrbildning (s.k. fibros). NAFLD kan i sin tur utvecklas till skrumplever (s.k. cirros) och levercancer (HCC). NAFLD definieras av en grad leverfett som överstiger 5 % och förekommer idag hos en tredjedel av befolkningen globalt. Idag finns inga godkända läkemedel mot NAFLD. Behandling yttrar sig istället i livsstilsförändringar med fokus på viktminskning. Samtidigt har det visat sig i välkontrollerade kortsiktiga studier att om man byter mättat fett mot fleromättat fett i kosten så minskar nivåer av leverfett oberoende av en skillnad i vikt mellan grupperna. Detta talar för att kvaliteten av fett i kosten har en central roll i uppbyggnaden av fett i levern, men vetenskapligt underlag för mer allvarlig NAFLD (cirros och HCC) är begränsat. Vetenskapligt underlag saknas också för hur kostmönster där fettkvaliteten är en viktig del påverkar nivåer av leverfett hos grupper med ökad risk för mer allvarlig NAFLD över tid (ex. diabetiker och prediabetiker). Såväl randomiserade kontrollerade studier som observationella studier behövs för att reda ut dessa kvarstående frågor. I synnerhet behövs observationella studier av god kvalitet för att studera samband mellan kostfaktorer och utfall som tar årtonden att utveckla, nämligen cirros och HCC. Observationella studier med självrapporterad kostdata möter dock många begränsningar, såsom felrapportering. Ett sätt att delvis komma runt detta problem är att använda objektiva biomarkörer för kostintag, såsom cirkulerande fettsyror i blodet eller fettsyror i vävnad. I min avhandling försöker jag reda ut vilken roll fettsyror har vid NAFLD. Fokus för avhandlingen ligger på dietära fettsyror men icke-dietära fettsyror som speglar en blandning av livsstilsfaktorer och metabolism diskuteras såväl.

I det första delarbetet var huvudsyftet att studera kopplingen mellan fettsyror i cirkulerande kolesteryl estrar och leverfett i ett populationsbaserat urval av 50-åriga män och kvinnor bosatta i Sverige. Studien var av tvärnittlig karaktär och innefattade n=308 individer. Sekundärt syfte var att titta på om samma fettsyror och energiomtebolismen och huruvida dessa samband kunde förklara en del av fynden mellan fettsyror och leverfett. Linolsyra (18:2n-6), som anses vara en god markör för livsmedel rika på denna fettsyra, var invers kopplad till energiomtebolism och huruvida dessa samband kunde förklara en del av fynden mellan fettsyror och leverfett. Mer av 18:2n-6 innebar alltså en lägre mängd leverfett. Likaså var förhållandet mellan arakidonsyra...
(20:4n-6) och dihomo-gamma-linolensyra (20:3n-6), en markör för aktiviteten av enzymet delta-5-desaturas (D5D), invers kopplat till mängden leverfett. Ett flertal fettsyror som i synnerhet speglar metabolismen (16:1n-7, 20:3n-6, 20:4n-6, 18:3n-6, 16:1n-7/16:0 (SCD-1) och 18:3n-6/18:6n-6 (D6D)) var positivt kopplade till mängden leverfett. Sekundärt föreläg svaga till moderat starka korrelationer mellan åtskilliga fettsyror i kolesterololstrar och fettoxidation och energiförbrukning i vila. Kopplingen mellan dokosahexaensyra (22:6n-3) och leverfett minskade kraftigt när justering för en ökad energiförbrukning i vila genomfördes.


I delarbete fyra var huvudsyftet att undersöka effekterna av tre olika kostbehandlingar på förändringen i leverfett över ett års uppföljning hos personer med typ-2 diabetes eller förstadium till typ-2 diabetes, s.k. prediabetes. I denna RCT-studie randomiserades 150 individer till tre olika koster som varierade i mängd kolhydrater och fettkvalitet. Den ena kosten (LCPUFA) utgjordes av ett lägre kolhydratsintag med högre mängd PUFA i form av nötter, frön, solrosolja och andra livsmedel. Den andra kosten (hällosom Nordisk kost
(HND)) utgjordes av ett högre kolhydratsintag och lägre fettintag där fokus låg på fullkornsprodukter i form av havre och råg, rapsolja, rikligt med svenska frukter och grönsaker och fet fisk. Den tredje gruppen (UC) var en kontrollgrupp som instruerades att följa de Nordiska näringsrekommendationerna från 2012. Samtliga deltagare besökte kliniken totalt 13 gånger på ett år där livsmedelskassar med nyckellivsmedel delades ut 10 gånger. Leverfett mättes med magnetkamera och följsamhet till kosterna mättes med hjälp av 4-dagars vägd kostdagbok samt biomarkörer i blod. Resultaten visade att följsamheten till kosterna var bra. Leverfett minskade i LCPUFA jämfört med UC samt i HND jämfört med UC. Ingen skillnad i leverfett noterades mellan HND och LCPUFA. Såväl totalkolesterol som LDL-kolesterol minskade i LCPUFA jämfört med UC medan kroppsvikt, långtidssocker (HbA1c), totalkolesterol, LDL-kolesterol, triglycerider och inflammationsmarkören CRP minskade i HND jämfört med UC. När HND jämfördes med LCPUFA observerades en skillnad i kroppsvikt och HbA1c, till favor för HND.

Sammanfattningsvis styrker fynden från denna avhandling tidigare studier på området om att kvaliteten på fettet i maten och i kroppen är av vikt vid NAFLD, i synnerhet för leverfett och fibros. Att öka intaget av linolsyra-rika och omega-3-rika livsmedel (i synnerhet 22:6n-3) och lägga fokus på ett kostmönster som betonar PUFA och ett minskat intag av SFA tycks spela roll vid NAFLD. Framtida studier behöver i sin tur utarbeta vilken roll icke-dietära fettsyror har i uppkomsten av NAFLD samt hur kosten påverkar mer allvarlig NAFLD, såsom cirros och HCC.
I would like to extend my full gratitude to the many who have contributed to making this thesis possible.

First of all, I would like to direct a massive thanks to my main supervisor Ulf Risérus and co-supervisor Fredrik Rosqvist for creating a safe and inspiring environment for me to develop as a PhD student. You instantly made me feel like I was part of the team and I couldn’t have wished for better supervisors to guide me through these four years. Allowing me to pave my own way, but letting me know when I have gone a bit too far, has been the most valuable insight of this journey, and for that I am forever grateful. Your extensive knowledge of the scientific literature will forever inspire me to read more and the never-ending curiosity for science that you both exude has been a joy to witness. With that said, you are not only great supervisors, you are also tremendous researchers.

I would like to thank all of you who have not yet gone tired of me asking lengthy, and sometimes confusing questions. Liisa Byberg for introducing the fields of epidemiology and causal inference to me. The seminars I have attended and the discussions we have had over the years have truly been eye-opening; Lars Lind for your extensive knowledge in the field of epidemiology and for always getting back with quick and thought-provoking answers to my questions; Lars Berglund for some of the more interesting discussions I have had on trial methodology and statistics; Daniel Ibsen and Conor Mac-Donald for the valuable discussions on food substitution models, the target trial framework and causal inference methods; Johan Vessby and Fredrik Rorsman for your valuable clinical input to the papers; Erika Olsson and Eva Warensjö Lemming for your expertise in dietary assessment methods and epidemiology.

A warm gratitude to all the collaborators that I had the fortune to work with during the NAFLDiet trial. Kerstin Marttala for your directness and professionalism. I had such a wonderful time working with you and I just hope we will work together again in the future; Sara Hansson for your help with taking care of the participants in the beginning of the trial; Lillemor Källström for your expertise in fatty acid analysis and laboratory work (and for doing the heavy lifting (no pun intended) in carrying all of the food bags to the clinic); Anders Lundberg and all the colleagues at the MR unit for great collaboration; Hans-Erik Johansson for your professionalism and for being deeply
involved in many parts of the trial; Joel Kullberg and Håkan Ahlström for your expertise in radiology and Andréa Flores for your extraordinary work with the food diaries. I would like to thank all the collaborators involved in other parts of the trial; you have been many and your work has been immensely valuable. I would also like to direct a massive thanks to all the participants in the NAFLDiet trial for your tremendous effort. Without you, none of this would have been possible. Furthermore, I wish to acknowledge and thank all the participants and personnel involved in the POEM cohort, the AM-02 NASH cohort and the SMC and COSM cohorts.

I would also like to thank all the former and present colleagues at the unit of Clinical Nutrition and Metabolism for all the discussions we have had over the years. Ulf Holmbäck, David Iggman, Mikael Karlsson, Roger Olsson, Matti Marklund, Mohammed Bajahzer, Tommy Cederholm and Wulf Becker, your input have really been nothing less than extremely valuable.

To my family and dear friends for always believing in me, pushing me and encouraging me throughout, although most of you probably had no idea what I was doing during these four years.

To you Seyran for being the love of my life, my best friend and my biggest supporter. Sharing the adventures of life with you is a bliss.
References


Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr 1997;65:1220S-1228S.


Alten S van, Domingue BW, Galama T, Marees AT. Reweighting the UK Biobank to reflect its underlying sampling population substantially reduces pervasive selection bias due to volunteering. MedRxiv 2022;2022.05.16.22275048.


Dahabreh IJ, Kazi DS. Toward Personalizing Care. JAMA 2023;329:1063.


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 2030

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)