The role of ethnicity in the increased susceptibility to non-alcoholic fatty liver disease in people of Indian origin

Jane Chalmers

Doctor of Philosophy

The University of Nottingham

June 2020

Abstract

Introduction

Non-alcoholic fatty liver disease (NAFLD) is fast becoming a global health concern. It is closely associated with obesity, diabetes and the metabolic syndrome, has a global pooled prevalence of 25% and is the leading cause of chronic liver disease in Europe and the United States.

People of Indian ethnicity are at increased risk of diabetes and metabolic complications at a lower body mass index (BMI) than Caucasians. This predisposition for NAFLD is further compounded by Westernisation of Asian culture resulting in increased intake of sugar-rich, energy-dense foods and decreased levels of physical activity.

Despite this well documented propensity to metabolic disease (including NAFLD), there is a lack of knowledge about the true disease prevalence in India. The studies that have been published are biased towards urban, tertiary centres, where access to healthcare and appropriate diagnostic tools is more readily available. In addition to this, there is conflicting data about the role of changing lifestyle habits on NAFLD risk – particularly in relation to diet. Reduced brown adipose tissue (BAT) activity has also been linked to obesity and diabetes, however there have been no prospective studies done to examine whether reduced BAT activity could also contribute to the increased NAFLD risk within this ethnic group.

Aims

The aim of this thesis is therefore to fill the above gaps in knowledge, to estimate accurately the NAFLD prevalence within a large Indian population, to identify the impact of different NAFLD risk factors within this population and compare NAFLD risk profile of native Indians with their UK migrant counterparts. It will also investigate the impact of BAT activity on NAFLD risk and understand the ethnic differences in BAT activity.

Results

Through population-level sampling of a large Southern Indian district, NAFLD prevalence (as diagnosed by ultrasound) was shown to be 49.8%.

Risk factors for NAFLD within India were the same as those seen worldwide, namely male gender (adjusted OR 2.29 1.86-2.83, $p<0.001$), obesity (adjusted OR 2.81 2.015-3.68, $p<0.001$) – in particular central obesity – and components of the metabolic syndrome (diabetes adjusted OR 1.76 1.40-2.21, $p<0.001$). Dietary fat intake was also independently associated with NAFLD within India (adjusted OR 1.02 1.00-1.03, $p=0.019$). These risk factors appear unchanged by migration to the UK. Although there were no significant differences in dietary habits following UK migration, there appeared to be an element of dietary acculturation with decreased consumption of

carbohydrate within the UK-migrant Indian cohort. This may result in changes to NAFLD phenotype over time.

BAT activity was lower in native Indians (ΔTrel 0.36°C) compared to UK Caucasians (ΔTrel 0.50°C, p=0.010) and UK South Asian migrants (ΔTrel 0.57°C, p<0.001). This difference was however due to environment, as there was no difference in BAT activity between control groups of different ethnicity living in the same country. BAT activity does not directly influence risk of NAFLD (adjusted OR 0.47 0.07-3.20, p=0.444). Any differences in BAT activity were related to increasing BMI, which is itself a risk factor for NAFLD.

Conclusion

The prevalence of NAFLD in India is significantly higher than current national and global estimates. There is a commonality of risk between India and the rest of the world – namely obesity, diabetes and a diet high in fat. Within India, increased consumption of saturated fat in the form of edible oils and meat appears to impact NAFLD risk additionally. Whilst BAT activity is lower in Indians, it is not a direct cause for the increased NAFLD risk in this ethnic group. This work highlights the need for NAFLD screening for people of Indian ethnicity, both in India and in migrant populations. Further research needs to focus on education and interventional strategies to reduce prevalence of obesity and diabetes, which may be achieved through dietary manipulation.

Declaration

I declare that this thesis is of my own composition, and that the research contained within it is entirely my own unless stated otherwise. No part of the work has been submitted for any other degree or professional qualification.

Dr Jane Chalmers

Nottingham, June 2020

Acknowledgements

I would like to thank my supervisory team, Prof Guru Aithal, Prof Kim Edwards and Prof Mike Symonds for their enthusiasm, encouragement and support throughout this PhD. I am particularly grateful to Prof Aithal for taking a chance on a non-academic trainee, for his never-ending supply of patience, and his willingness to find time whenever I needed it.

The whole team at the Population Health and Research Institute, Trivandrum, who made me so welcome whilst out collecting data, and whose hard work and dedication made this work possible.

I would like to acknowledge all those who have been part of the “liver lobule” during my time in Nottingham. For limitless coffee and offers of advice.

Dr Neil Guha, for his guidance throughout this research, both in terms of academia and life. His continued support has helped shape my future career.

Above all, my husband Jonny, for enabling me to pursue this opportunity, and for his patience and understanding during the writing of this thesis.

Relevant publications, presentations and awards

Papers

Law, J., Chalmers J et al., The use of infrared thermography in the measurement and characterization of brown adipose tissue activation. Temperature (Austin), 2018. 5(2): p. 147-161.

Chalmers J, Ban L, Leena KB, et al Cohort profile: the Trivandrum non-alcoholic fatty liver disease (NAFLD) cohort BMJ Open 2019;9:e027244.

Abstracts

Chalmers J, Law J, Ban L, Leena KB, Symonds M, Shenoy KT, Aithal GP 'Does brown adipose tissue activity underlie ethnic variation in susceptibility to non-alcoholic fatty liver disease (NAFLD)?' - Poster presented at AASLD (Nov 2019) and BASL (Sept 2019)

Chalmers J, Wayne D, Grove JI, Symonds M, Leena KB, Shenoy KT, Aithal GP 'The influence of TRPM8 variant on brown adipose tissue activity and contribution to increased susceptibility to non-alcoholic fatty liver disease among South Asians' – Poster accepted for presentation at EASL (April 2020)

Table of Contents

Abbreviations

ALP	Alkaline phosphatase
ALT	Alanine transferase
AST	Aspartate aminotransferase
BAT	Brown adipose tissue
BMI	Body mass index
CRF	Case report form
CI	Confidence interval
CT	Computed tomography
CVD	Cardiovascular disease
EE	Energy expenditure
ELF	Enhanced liver fibrosis
FFQ	Food frequency questionnaire
GGT	Gamma-glutamyl transpeptidase
GPAQ	Global physical activity questionnaire
GWAS	Genome-wide association study
HCC	Hepatocellular carcinoma
HDL	High-density lipoprotein
HOMA-IR	Homeostatic model assessment of insulin resistance
HR	Hazard ratio
IDF	International Diabetes Federation
IGT	Impaired glucose tolerance
IHD	Ischaemic heart disease
IPAQ	International physical activity questionnaire
IRT	Infrared thermography
LDL	Low-density lipoproteins
LDL-C	Low-density lipoprotein cholesterol
LFT	Liver function tests
MAF	Mean allele frequency
MET	Metabolic equivalent

MRI	Magnetic resonance imaging
MRS	Magnetic resonance spectroscopy
NAFLD	Non-alcoholic fatty liver disease
NAS	NAFLD activity score
NASH	Non-alcoholic steatohepatitis
OR	Odds ratio
PUFA	Polyunsaturated fatty acids
PET	Positron emission tomography
PVD	Peripheral vascular disease
ROI	Region of interest
SNP	Single nucleotide polymorphism
TE	Transient elastography
TG	Triglycerides
UCP1	Uncoupling protein 1
UK	United Kingdom
USA	United States of America
VIF	Variants inflation factor
VLDL	Very-low density lipoproteins
WAT	White adipose tissue
WHO	World Health Organisation
WHR	Waist-hip ratio

List of tables

List of figures

CHAPTER 1: Introduction

Non-alcoholic fatty liver disease (NAFLD) is a condition defined by fat deposition within the liver (>5% of hepatocytes), in the absence of an alternative cause of liver disease or excess alcohol intake (greater than 20g/day women, greater than 30g/day men). It is a disease that is closely associated with obesity, diabetes and the metabolic syndrome[1]. Non-alcoholic steatohepatitis (NASH) is a subtype of NAFLD that may result in progressive inflammation and fibrosis within the liver, culminating in cirrhosis, hepatocellular carcinoma (HCC) and necessity for transplantation. NAFLD has an estimated global prevalence of 25%[1] and is the leading cause of chronic liver disease in the Europe and the United States of America (USA)[2]. NAFLD is associated with increased morbidity and mortality. Those with NAFLD have a higher 10-year mortality than controls, and NAFLD itself is a risk factor for incident metabolic disease[3].

People of Indian ethnicity are at increased risk of diabetes and metabolic complications at a lower body mass index (BMI) than Caucasians[4, 5]. This has been compounded by Westernisation of Asia-Pacific culture[6], resulting in higher rates of NAFLD in this ethnic group[7, 8].

The purpose of this thesis is to add to this evidence base, to estimate accurately the NAFLD prevalence across a large Indian district and to confirm which factors influence disease risk, to investigate whether decreased brown adipose tissue (BAT) activity contributes to the adverse metabolic phenotype seen in this group, and to understand the impact of culture and environment on the NAFLD risk profile of people of Indian origin.

1.2 Non-alcoholic fatty liver disease (NAFLD)

1.2.1 Introduction

NAFLD is defined by the presence of fat within >5% of hepatocytes within the liver, in the absence of an alternative cause of liver disease or excess alcohol consumption (less than 20g/day women, less than 30g/day men)[9, 10]. NAFLD represents a spectrum of disease from simple steatosis (fat accumulation within hepatocytes) to NASH (liver cell injury and death), which is a sub-type which is more likely to lead to increasing stages of fibrosis to cirrhosis[11] resulting in increased risk of HCC[12] and liver transplantation requirement[13]. NAFLD is widely considered the hepatic manifestation of the metabolic syndrome, a constellation of metabolic abnormalities including glucose intolerance, obesity (particularly central obesity), dyslipidaemia and hypertension[14]. These metabolic complications often coincide and have been shown to increase risk of cardiovascular disease and mortality overall[15]. Thus, with the global pandemic of obesity and the metabolic syndrome in both developed and developing countries, comes a dramatic rise in prevalence of NAFLD[16, 17].

1.2.2 Prevalence and disease burden

The prevalence of NAFLD is increasing at a similar rate to obesity[18]. NAFLD has a current estimated global prevalence of 25%[1], whilst prevalence of NASH is 3-5%[2, 19]. It is important however to highlight the difference in NAFLD prevalence across the world. For example, prevalence of NAFLD has been shown to be significantly higher in South America, than North America[20], despite having lower rates of obesity[2]. This is likely to be due to a combination of lifestyle and genetic factors. Data on the burden of NAFLD within Asia is evolving. Urbanisation of developing countries has resulted in increased sedentary lifestyle and over-nutrition, resulting in increased rates of metabolic disease[21]. Prevalence of NAFLD in China has doubled over the last 20 years[22]. Here in the UK, rates of NAFLD also appear to be slightly higher than global estimates, ranging from 26.4%-32% depending on the diagnostic criteria used[23-25] - which is unsurprising given that 25% of the UK population is obese[26]. Global rates of NAFLD appear to be increasing in line with rates of obesity and diabetes, with data modelling suggesting increasing cases of advanced liver disease and liver-related mortality over the coming decade[27].

A diagnosis of NAFLD is associated with an increase in both morbidity and mortality. Population-based data on NAFLD has shown that ten-year mortality is higher in those with NAFLD compared with controls (10.2% vs. 7.6%), and that NAFLD itself is a risk factor for incident metabolic disease[3]. In those with NAFLD, an increasing number of metabolic complications is associated with increasing mortality, and as such, development of NAFLD in those with obesity and diabetes predisposes to further metabolic disease and further increases risk of death. This is almost a self-fulfilling prophecy, when 80% of those with NASH are known to be obese and 44% to have

type 2 diabetes[1]. Presence of diabetes in those with NAFLD is of particular concern, as it is associated with increased risk of fibrosis and thus progression of disease[28]. Whilst the most common cause of mortality in those with NAFLD is cardiovascular disease[29], NAFLD itself only independently increases risk of cardiovascular disease in those without the metabolic syndrome i.e. incident metabolic complications annuls impact of NAFLD on cardiovascular disease[3].

Alongside the significant clinical burden of NAFLD, there is also huge socioeconomic burden. Data from a large national administrative claims database has shown that the long-term annual cost to a patient with NAFLD in the USA is $3,789 (IQR $1,176-10,539), compared to a patient with metabolic syndrome alone $2,298 (IQR $681-6,580)[30]. The global economic burden of NAFLD has been estimated through Markov modelling, with the USA having 39 million cases of NAFLD costing their economy $62 billion, whilst in Europe, 30 million people are estimated to have NAFLD costing their economy $19 billion[31].

In addition to significant economic burden, there is also emerging data on the burden of NAFLD to patients themselves. One study showed that 27.2% of patients with NAFLD report symptoms of depression, a rate that is four times higher than in the general population[32]. Another study showed that 20% of patients with NAFLD reported their health as “fair or poor”, compared to 10% of healthy controls[33]. Interestingly, loss of 5% of body weight in those with NAFLD appears to improve not only disease-related outcomes, but also measures of health-related quality of life[34].

Given the rapidly growing global burden of NAFLD and NASH, further research is needed to identify those at risk, the reasons for their increased risk and to enable future interventional studies to focus on those factors that influence risk the most.

1.2.3 Diagnosis

The gold standard for diagnosis of NAFLD is through histological assessment of liver tissue via liver biopsy. Disease activity and stage are in general assessed using the Kleiner classification (Table 1-1)[11]. Due to its invasive nature and inherent sampling error[35], liver biopsy is not a suitable diagnostic tool for a disease with such high prevalence. Imaging modalities such as ultrasound and magnetic resonance imaging (MRI) techniques are now more widely used[36], although these modalities are also limited by their sensitivity/specificity (ultrasound has a sensitivity of 43% and specificity of 73%[37]), cost and availability.

Histological Feature	Score	Category Definition
Steatosis	0	< 5%
	1	5-33%
	2	34-66%
	3	>66%
Ballooning	0	None
	1	Few
	2	Many
Inflammation	0	None
	1	1-2 foci per x20 field
	2	2-4 foci per x20 field
	3	>4 foci per x20 field
Total NAS score = 0-8 Score ≥5 with steatosis and ballooning = NASH		
Fibrosis	0	None
	1a	Zone 3 mild perisinusoidal
	1b	Zone 3 moderate perisinusoidal
	1c	Periportal/portal
	2	Zone 3 + periportal/portal
	3	Bridging
	4	Cirrhosis
Fibrosis grade = 0-4		

Table 1-1. NAFLD activity score (NAS)[11]

In addition to identifying presence of NAFLD, it is also important to stage disease accurately. This is because numerous natural history studies have demonstrated that severity of liver fibrosis most accurately predicts liver-related outcomes in people with NAFLD[38, 39]. 70% of patients with NASH develop fibrosis progression[1], and studies show that HCC may develop in 1.1% of those with NASH cirrhosis over a 10 year period[40]. This demonstrates the importance of risk stratifying those diagnosed with NAFLD, identifying those at risk of significant fibrosis and those with advanced liver disease. This can be done utilising a number of non-invasive tools developed over the last decade. These include blood-based biomarkers such as the NAFLD fibrosis score, FIB-4, ELF and Fibrotest®, as well as imaging tools such as TE and magnetic resonance elastography[28, 41, 42].

1.2.4 NAFLD risk factors

Given the high prevalence of disease and the metabolic/cardiovascular implications of the diagnosis, screening for NAFLD could be considered important both clinically and economically. However, high upfront costs of screening, low yield for significant, advanced disease and lack of effective treatments mean that it is not currently appropriate to screen for NAFLD within the general population[9]. Attempts to identify advanced disease (presence of significant fibrosis) should nevertheless be

made in those at risk of NAFLD, such as those with diabetes or the metabolic syndrome, as its presence is important in prognostic terms for both liver-related and cardiovascular outcomes[43].

Age and sex

NAFLD prevalence is known to increase with age[2]. Studies have also demonstrated that there is an increased risk of advanced fibrosis and HCC in NAFLD with advancing age[44]. These findings may however be in part due to the increased prevalence of metabolic disease in older age. Data relating to NAFLD risk in relation to gender is conflicting. Some studies suggest a female preponderance to NAFLD[45, 46], whilst others report higher rates in men[47, 48]. Neither age nor gender form part of current risk stratification models in the UK[9, 43].

Metabolic risk factors

NAFLD is strongly linked to obesity, diabetes and the metabolic syndrome[1, 49]. The physiological links between obesity, insulin resistance and NAFLD have been studied in detail – a process outlined by Shulman et al in 2014. Essentially, intake of dietary lipid that exceeds the storage capacity of adipose tissue leads to spill-over and ectopic lipid deposition in skeletal muscle and liver. Lipid deposition within the muscle leads to local insulin resistance, which further drives lipid accumulation in the liver. When carbohydrate is ingested, skeletal muscle usually acts as a "sink", utilising the energy from the glucose. However, intramyocellular lipid disrupts uptake of glucose, which is diverted to the liver where it is metabolised to free fatty acids. The liver, therefore, accumulates fat from both spill-over from adipose tissue, and de novo lipogenesis from glucose diversion from insulin resistant muscle [50].

In addition to this process, adipocytokines (TNF α, IL-6, IL-8, Resistin, Leptin and Adiponectin) also play a role in regulation of hepatic and peripheral glucose and lipid metabolism. The balance between these adipocytokines may be important in the pathogenesis of NAFLD [51]. One theory is that visceral adiposity causes release of adipocytokines, which are transported to the liver via the portal vein [52].

Evidence shows that increased visceral adipose tissue is a risk factor for hepatic steatosis and increased severity of disease [53, 54], whereas subcutaneous fat distribution and peripheral adiposity is negatively associated with liver fat [55] – a distribution more commonly seen in African American populations, who have the lowest rates of NAFLD. Visceral adiposity is also directly associated with liver inflammation and fibrosis [56].

Therefore, central obesity, impaired glucose tolerance or type 2 diabetes, hypertension and dyslipidaemia all form part of the NAFLD risk profile, and current guidelines recommend assessing for presence and severity of NAFLD if any of these are present[9, 57].

Lifestyle risk factors

The rising prevalence of NAFLD may be linked to global changes in lifestyle, where increased intake of an energy-dense, high sugar diet and decreased levels of physical

activity are associated with contemporaneous epidemics of obesity, diabetes and the metabolic syndrome. Diet in particular has been strongly linked to the development of obesity and subsequent insulin resistance[58], as well as directly resulting in fat deposition, altered lipid metabolism and oxidative stress within the liver[50, 59, 60].

Animal models have shown that a high fat diet induces steatosis of the liver as well causing steatohepatitis and fibrosis through oxidative damage[61, 62]. The same was shown in humans through a dietary intervention crossover study, where obese women followed diets composed of 16% (low fat) and 56% (high fat) of total energy intake as fat. Liver fat as measured by magnetic resonance spectroscopy (MRS) decreased 20% on the low fat diet, and increased 35% on the high fat diet in the absence of change in body composition[63].

Type of food consumed also has an impact on insulin sensitivity and glucose handling. It is suggested that foods of high glycaemic index are associated with insulin resistance[64], in particular consumption of fructose-rich soft drinks lead to increased hepatic synthesis of triglycerides[65].

Physical activity also has an effect on the pathophysiology of NAFLD even in the absence of change in weight, through improvement of insulin sensitivity and glucose homeostasis with higher physical activity levels. Evidence shows that exercise training causes upregulation of insulin receptors in muscle tissue leading to increased delivery of glucose and insulin to the muscle[66]. Through strength training, increased muscle mass further increases insulin sensitivity by increasing glucose storage capacity[67]. Physical activity also has a positive impact on lipid metabolism, improving whole-body lipid oxidation and decreasing hepatic free fatty acid uptake[68].

In the absence of pharmacological treatments, interventions targeted at diet and exercise have long been the hallmark of NAFLD management, with evidence showing that weight reduction and increased physical activity can lead to reduced steatosis and degree of inflammation and fibrosis[69-71]. Less is known however about the impact of lifestyle on the development of NAFLD. This section presents detailed evidence on the impact of lifestyle on presence of NAFLD.

Diet

Multiple case-control and cross-sectional studies have been undertaken to identify differences in diet between people with NAFLD and those without. Some studies analysed the differences in macronutrients, whilst others look at dietary patterns to enable assessment of overall effects of diet through quantification of the cumulative effect of multiple nutrients. These findings are presented in Table 1-2.

Study (Country)	Patient groups	Study type	Results
Musso et al. (Italy) [59]	25 NASH 25 Control	Case control study	Dietary intake of NASH higher in saturated fat and cholesterol, lower in polyunsaturated fats (PUFA).
Toshimitsu et al. (Japan) [72]	28 NASH 18 Simple steatosis	Case control study	Higher consumption of simple carbohydrates and lower consumption of protein in NASH
Solga et al. (USA) [73]	74 NAFLD undergoing bariatric surgery	Cross-sectional study	Diet with higher proportion of carbohydrate increased odds of histological inflammation
Machado et al. (Portugal) [74]	43 NASH 33 Control	Case control study	Negative correlation between total/saturated fat intake and measures oxidative stress. Converse for carbohydrate and PUFA
Li et al. (China) [75]	78 NASH	Randomised controlled trial	Improved histological markers of NAFLD severity with PUFA therapy
Abdelmalek et al. (USA) [76]	341 NASH	Cross-sectional study	Daily fructose consumption associated with higher fibrosis grade
Zelber-Sagi et al. (Israel) [77]	108 NAFLD 241 Control	Case control study	Higher consumption of soft drinks and meat associated with increased risk of NAFLD
Cortez-Pinto et al. (Portugal) [78]	45 NASH 856 Control	Case control study	Carbohydrate consumption lower in cases, fat intake higher (including omega-3)
Oddy et al. (Australia) [79]	151 NAFLD 844 Control	Cohort study	“Western” diet associated with increased risk of NAFLD compared to “Healthy” diet
Yang et al. (China) [80]	345 NAFLD 654 Control	Case control study	“Animal food” dietary pattern associated with greatest odds of NAFLD, “Grains-vegetables” pattern the lowest
Chung et al. (South Korea) [81]	331 NAFLD 859 Control	Case control study	“Traditional” dietary pattern associated with greatest odds of NAFLD, “Simple” pattern the lowest

Table 1-2. Studies assessing the role of diet in NAFLD

Musso et al compared 25 patients with NASH to 25 matched controls using a seven-day dietary record and postprandial lipid metabolism testing[59]. They found that there was no difference in total energy, carbohydrate, protein or fat intake between cases and controls, but that people with NASH consumed more cholesterol, saturated fat, and less polyunsaturated fat and dietary fibre (likely due to reduced intake of fruit and vegetables). Those with NASH had increased postprandial triglyceride levels. This work suggested that whilst total calories and macronutrient intake were the same between groups, it is the quality of fat intake that differs, which may lead to altered lipid metabolism.

A similar study was conducted by Toshimitsu et al, who compared three days of dietary intake between 28 patients with NASH and 18 with simple steatosis[72]. There was no difference in total calorie or fat intake between groups, but it appeared that those with NASH had a higher carbohydrate intake and lower protein intake. However, these findings were only seen within specific age groups, and, as such, are limited by small sample size.

Dietary composition may also influence severity of NAFLD. Solga et al analysed the diet of 74 morbidly obese women undergoing bariatric surgery, of whom 90% had NAFLD[73]. Food intake was analysed in conjunction with liver histology from intraoperative biopsy. They showed that a higher carbohydrate intake (>54% of total calories) was associated with increased odds of inflammation, whilst the converse was true of higher fat intake (as proportion of diet consumed as fat increases, proportion of carbohydrate decreases). The study however was underpowered to detect differences between those with/without inflammation, and did not give detail about the differences in specific lipid subtypes. The same result was not seen by Machado et al, who examined differences in blood markers of oxidative stress between NAFLD cases and controls in relation to dietary intake[74]. They demonstrated a pro-oxidant effect with increased total and saturated fat intake and the reverse with increased carbohydrate, fibre and polyunsaturated fat intake when adjusted for total energy consumption. This link between increased poly-unsaturated fatty acid intake and improvement of NASH was further demonstrated by Li et al[75].

Other studies have demonstrated links between specific nutrients and NAFLD. The consumption of sugar-sweetened beverages has been linked to development of obesity, diabetes and may result in fat deposition within the liver[65]. A sugar-rich diet is also associated with higher fibrosis grade in those with NASH[76]. Evidence also suggest that fructose in particular is linked to NAFLD, where hepatic metabolism of fructose increases lipogenesis (NAFLD) and formation of reactive oxygen species resulting in increased risk inflammation (NASH)[82, 83]. Omega-3 consumption has been linked to NAFLD too, although observational studies are conflicting as to whether patients with NAFLD consume more [78] or less [77] in their diet. Interventional studies do however appear to support a protective role of Omega-3 supplementation in those with disease through improving hepatic steatosis[84, 85].

In addition to studies examining macro and micronutrient dietary composition and presence of NAFLD, more recent studies have focussed on dietary patterns and disease. This is of particular relevance when there is such difficulty in achieving and maintaining sustained weight loss long-term in patients with NAFLD, and where a change in diet composition without reducing calorie intake may be a more realistic alternative. One study showed that following a "Western" diet (high in processed food, red meat, refined cereals and confectionary) resulted in increased risk of NAFLD over three years (OR 1.59, $p<0.005$) when compared to a "healthy" diet (high in wholegrain cereals, fruit, vegetables and fish)[79]. Other studies have examined the effect of a Mediterranean diet on NAFLD[86, 87]. The Mediterranean diet is traditionally low in saturated fat and cholesterol, and high in monounsaturated fatty acids with a balanced PUFA omega-6 to omega-3 ratio. The protective effect of this balanced fatty acid consumption on NAFLD further strengthens the findings outlined in the previous paragraph. Large studies in Asia have looked at the impact of a "Traditional diet" compared to a Western diet, or a diet composed of more grains, fruit and vegetables[80, 81]. These studies appear to show that "Simple diets" composed principally of vegetables and grains infer the least risk of NAFLD, but are conflicting in which diet patterns confer the most risk, mainly due to differences in local dietary habits.

Exercise

Numerous studies have looked at the impact of physical activity in relation to NAFLD, in terms of amount or type of exercise, or in relation to amount of time spent inactive. Perseghin et al showed that compared to controls, people with NAFLD did less exercise overall[88], and Zelber-Sagi et al showed that people with NAFLD engage in less leisure-time physical activity - although only differences in resistance training remained significant when adjusted for insulin resistance, diet and BMI[89]. Kistler et al examined the role of exercise in relation to severity of disease, and showed that in patients with biopsy-proven NAFLD, meeting the recommendations for vigorous activity (75 minutes per week) was the only way to reduce the risk of NASH and significant fibrosis[90]. Church et al demonstrated an inverse association between cardiorespiratory fitness in general and NAFLD prevalence, regardless of BMI[91]. In addition to studies demonstrating the link between exercise and NAFLD, a British group looked at the link between sedentary time and presence of NAFLD and found that, compared to matched controls, those with NAFLD took fewer steps per day, had a lower overall energy expenditure and spent more time inactive[92].

The largest and most detailed longitudinal study of the effect of exercise on development of NAFLD was undertaken in South Korea, utilising data from regular health check-ups of over 126,000 staff members from a large corporation over an average period of five years[93]. They demonstrated that any amount of regular exercise improved the hazard ratio (HR) for NAFLD development, independent of BMI (HR 0.86, $p<0.001$) and that NAFLD risk decreased with increasing amount of exercise during the follow-up period. They also showed that in those with NAFLD,

any amount of regular exercise increased the chance of NAFLD resolution (HR 1.4, p<0.001).

Brown adipose tissue activity

As a result of changing lifestyle habits and increasing rates of metabolic syndrome and NAFLD, much work has been done to try to identify the intricate mechanisms that trigger these metabolic diseases. With this, there has been a resurgence of interest in BAT, its relationship with obesity, and its ability to increase energy expenditure on activation[94, 95].

BAT was first documented in 1551 by the naturalist Conrad Gesner, who described tissue found in the interscapular region of the marmot as "neither fat nor flesh – but something in between"[96]. Having initially been identified solely by its colour (due to its granular, mitochondria-rich cytoplasm), BAT was only conclusively demonstrated as being from a separate cell lineage to white adipose tissue (WAT) in the early 2000's – where it was shown that bone morphogenic proteins were involved in the differentiation of adipocytes into white or brown adipose tissue [97]. BAT is unique, in that it contains uncoupling protein 1 (UCP1) on the inner surface of its mitochondrial membrane, which, when activated, allows rapid dissipation of energy by the free flow of electrons, producing heat [94] and playing a pivotal role in thermoregulation, particularly in infancy. Historically, BAT was thought to disappear with age, however the persistence into adulthood of BAT in small depots, at sites similar to those found in infants, was noted by the pathologist Heaton in the 1970s [98]. The largest BAT depots are found surrounding the vasculature of the neck [99], where they are thought to play a role in thermoregulation of the blood to and from the brain [100]. It is the activity of these depots, noticed on positron emission tomography/computed tomography (PET-CT) imaging of adults [95, 101], which promoted the resurgence of interest into BAT activity early in the millennium.

BAT activity is controlled by the action of norepinephrine (from the sympathetic nervous system) on β_3-adrenoreceptors[102], and is inhibited by vagal nerve stimulation [103] (from the parasympathetic nervous system). Norepinephrine has a number of actions on BAT to increase its activity: promoting proliferation of preadipocytes, differentiation of mature adipocytes[104], directly upregulating the expression of genes coding for UCP1 [105], increasing mitochondrial mass [106] and preventing apoptosis [107]. The activation of BAT through this pathway can be acute (i.e. in response to abrupt cold exposure or a meal [108]) or can be a result of enhanced BAT recruitment (through prolonged, mild cooling), which can increase thermogenic capacity over a period of time.

One theory proposed within the literature is that activation of BAT may be one way to control whole-body energy expenditure as well as amount of body fat. Mice who overexpress UCP1 are notably leaner than wild-type mice and further activation of BAT (using cold or beta-agonists) results in a reduction in body fat and BAT hyperplasia [109]. In mice, UCP1 ablation induces obesity even on a controlled diet, with an increase in body weight by 50% more than wild-type ($p<0.05$), as well as

augmenting diet-induced obesity [110]. Further work utilised mouse-models to examine the link between BAT activity and NAFLD. Giles et al demonstrated that thermoneutral housing (22-24°C) augmented proinflammatory immune responses and exacerbated high fat diet-induced NAFLD in mice – implying lack of thermogenic energy expenditure from BAT [111]. In addition, Poekes et al [112] showed that foz/foz mice (prone to obesity, insulin resistance and progressing fibrosing NASH) have severely impaired thermogenic adaptability to diet or cold exposure associated with decreased sympathetic tone in BAT. Intermittent cold exposure however did restore BAT function, improve glucose tolerance and reduce fat mass and liver steatosis. This study concluded that failure of BAT adaptation might play an important role in development of hepatic steatosis.

Early work looking at BAT activity in adult humans was done through retrospective review of large numbers PET-CT scans performed for clinical purposes. These demonstrated that BAT activity is inversely proportional to body weight[113] and is linked to insulin sensitivity[114] and plasma lipid profile[115]. This promoted further detailed, prospective studies looking at the impact of BAT activity on the metabolic profile of humans. Chondronikola et al. examined the role of cold-activated BAT on glucose homeostasis in healthy men and showed that, when activated, BAT increased whole-body glucose disposal and insulin sensitivity[116]. The same was not seen in those participants without active BAT on PET-CT (deemed "BAT-negative"). It was therefore hypothesised that activation of BAT could similarly improve glucose homeostasis in people with diabetes. Hanssen et al. looked at the impact of cold acclimation in men with type 2 diabetes, where participants were exposed to temperatures of 14-15°C for 6 hours a day for 10 days. They demonstrated that through activation of BAT, peripheral insulin sensitivity was improved by 43%[117], exceeding the improvements seen in long-term exercise training[118] – even in the presence of low levels of BAT activity overall. The same group examined the impact of 10-day cold acclimation in healthy obese men. In those men who were "BAT-positive" on PET-CT, cold acclimation did increase BAT activity[119]. There were however no changes in energy expenditure or plasma biochemistry. A further study compared BAT activity via PET-CT between lean and obese healthy men[120]. Following activation through cooling and infusion of insulin, glucose uptake of BAT and overall insulin sensitivity was higher in lean compared to obese (in whom rates of "BAT-positivity" were low).

Although there are a number of studies that have examined BAT activity in obese and diabetic patients, there is a paucity of data on BAT activity in relation to presence of NAFLD. Yilmaz et al. retrospectively analysed clinical PET-CT scans of 1832 individuals, of which only 2% were "BAT-positive" within a thermoneutral, non-activated state. When compared with matched "BAT-negative" individuals, the odds of having NAFLD was significantly higher in the "BAT-negative" group (odds ratio – OR 3.12, 1.03-9.88, $p<0.05$)[121]. More recently, Zhang et al examined the prevalence and predictors of active BAT within a retrospective cohort of over 31,000 patients undergoing PET-CT, either as a routine medical check-up or for cancer

surveillance. Again, the rate of detection of BAT from these clinical scans was only 1.3%, but there was evidence that BAT was more likely to be detected in those who had no fatty liver ($p<0.001$)[122].

Whilst mechanistically it has been shown that BAT activity is linked to obesity, diabetes and the metabolic syndrome, and that through activation of BAT, insulin resistance and lipid profiles can be improved, translation of this into human studies is more difficult. Although rates of "BAT positivity" using PET-CT seem to be low in those with an adverse metabolic phenotype, cold-activation of BAT appears to improve metabolic parameters[117]. This suggests that it is the tool used to measure BAT activity that may be a limiting factor. In addition to glucose, lipid is also utilised by BAT as a substrate. Less is known about the rates of lipid hydrolysis within the BAT active state, particularly within a phenotype that have a lipid surplus. Therefore, it may be that glucose uptake via PET-CT is not the most appropriate modality for quantifying BAT activity. The methods of assessing BAT activity in clinical studies is described further in Chapter 5.

Genetic risk factors

In addition to environmental factors, genetic factors also play an important role in the pathogenesis of NAFLD. Numerous studies have been done looking into genetic variants that may predispose a person to NAFLD. These can be broadly split into candidate-gene studies and genome-wide association studies (GWAS). Candidate gene studies look at the frequency of a specific genetic variant in relation to prevalence of disease (one that has been previously implicated in that process), whereas a GWAS investigates the whole (or large proportion of) genome to identify variants that are more common in people with a certain disease. To date, there have been 10 GWASs that have identified various genetic variants that confer risk of NAFLD [123-132] including PNPLA3[133], PPAR[134], APOC3[135], GCKR[136] and TM6SF2[131]. The role of genetics on NAFLD risk falls outside the remit of this body of work, and as such are not described in detail here.

1.2.5 NAFLD and ethnicity

Several studies have demonstrated differences in prevalence and severity of NAFLD between ethnic groups. The majority of the large, population-based studies of NAFLD prevalence were conducted in America, with data collected from participants of the main ethnic groups in the USA, namely Caucasians, Hispanics and African-Americans[55, 137, 138]. These studies showed that Hispanics are at greatest risk of NAFLD compared to Caucasians, with African-Americans having the lowest rates of disease prevalence and severity, despite having higher rates of obesity and diabetes. It was hypothesised therefore that ethnic variation in body composition – body fat distribution in particular – might influence NAFLD risk. African-Americans were shown to have less visceral fat and more lower-extremity fat than the other ethnicities[55]. When controlled for visceral fat, the differences in liver fat were abolished between groups, indicating a close association between liver fat and visceral fat, regardless of ethnicity[55]. Other studies have looked at differences in

NAFLD risk factors between different ethnic groups, but are limited by sample size and NAFLD diagnostic tools used[139, 140]. The findings of a large number of studies examining the NAFLD prevalence in different countries was summarised by Younossi et al. in their meta-analysis[1]. They demonstrated that the highest prevalence of NAFLD was seen in South America (30.45%, 22.74-39.44) and the Middle East (31.79%, 13.48-58.23), whilst the lowest prevalence was in Africa (13.48%, 5.69-28.36). These data suggest the interaction of genetic and environment factors may influence the prevalence of NAFLD.

1.3 NAFLD and Indian ethnicity

People of Indian ethnicity are at increased risk of diabetes and metabolic complications at a lower body mass index (BMI) than Caucasians[4, 5]. This has been compounded by Westernisation of Asia-Pacific culture[6], resulting in higher rates of NAFLD in this ethnic group[7, 8]. The causes of the increased risk of NAFLD amongst Indians have been studied widely throughout the literature, a summary of which is presented below.

1.3.1 Indian NAFLD prevalence

As outlined in section 1.1.2, the prevalence of NAFLD is rising on a global scale, with a recent meta-analysis showing a pooled global prevalence of 25.24% [1]. The issue with this generalisation is that it is mainly based on large population-based studies from the USA, which include relatively small numbers of people of Indian ethnicity. In addition, the sub-meta-analysis of Asian NAFLD prevalence within this study (27.37%) was calculated from studies undertaken in East Asian populations such as Japan and China, and is therefore not representative of the Asian population as a whole.

The documented prevalence of NAFLD also varies depending on the diagnostic tool used. In low-middle income countries (like India), it is difficult to perform large population-based studies, with those that are published being biased towards urban, tertiary centres where ultrasound and biopsy are more readily available[7, 141]. This is likely to give an overestimate of disease prevalence in India due to the significant variation in degrees of urbanisation and socioeconomic class.

Studies published on the prevalence of NAFLD in India are often small in sample size and heterogeneous in their diagnostic modality, with very few utilising biopsy. Documented prevalence ranges from 8.7% in rural West Bengal [142], where the majority of people are manual workers and are economically poor, to 87% in an urban tertiary diabetes centre [143]. Other large population-based studies include the Chennai Urban Population study, which included a random sample of 541 participants (from the total population cohort of 26,000) demonstrating prevalence of sonographic NAFLD as 32%[7] - which is higher than the global estimate. A study

of two suburban railway colonies (n=1168) reported NAFLD prevalence as 16.6% [144]. Table 1-3 shows the variation in documented prevalence according to setting, diagnostic tool used and study group.

Study	Setting	Diagnostic tool	NAFLD prevalence (%)	Sample size	Population
Mohan et al. [7]	Urban	Ultrasound	32	541	Subset of Chennai Urban Population Study
Das et al. [142]	Rural	Ultrasound (USS) and CT	8.7 (9.9 via USS)	1911	Adults from a rural administrative unit
Prashanth et al. [143]	Urban Tertiary	Ultrasound	87	204	Tertiary diabetes centre
Amarapurkar et al. [144]	Urban Suburban	Ultrasound	16.6	1168	Railway colonies
Chan et al. [145]	Urban	Ultrasound	33	73	Medical students
Singh et al. [146]	Urban Tertiary	Ultrasound	24.5	159	Healthy attendants accompanying patients to clinic
Praveenraj et al. [147]	Urban Tertiary	Biopsy	65.7	134	Morbidly obese patients undergoing bariatric surgery
Majumdar et al. [8]	Rural	Ultrasound	30.7	176	Random selection from 8/28 villages
Anurag et al. [141]	Rural	Ultrasound	28.1	302	Subjects attending a tertiary centre from villages

(continued)

Study	Setting	Diagnostic tool	NAFLD prevalence (%)	Sample size	Population
Gupte et al. [148]	Urban Tertiary	Ultrasound	49	100	Consecutive patients with diabetes
Agarwal et al. [149]	Urban Tertiary	Ultrasound	57.2	124	Consecutive patients with diabetes
Hajong et al. [150]	Urban Tertiary	Biopsy	31	200	Patients undergoing cholecystectomy
Duseja et al. [151]	Urban Tertiary	Ultrasound	54.3	971	Voluntary blood donors
Uchil et al. [152]	Urban Tertiary	Ultrasound	22.6	1003	Consecutive patients attending routine health check-up
Bajaj et al. [153]	Urban	Ultrasound	32.2	121	Consecutive healthy individuals

Table 1-3. Studies of NAFLD prevalence in India

This emerging data shows that NAFLD is no longer a disease confined to the Western world. The World Health Organisation (WHO) Global Burden of Disease study shows that hypertension and raised fasting blood glucose are the leading cause of non-communicable disease in India, unlike other low-middle income countries where childhood malnutrition and unsafe water predominate [154]. This rapid rise in non-communicable disease is predicted to impede poverty reduction initiatives in low-income countries by increasing household costs associated with healthcare. It is for this reason that we need to identify and address causes of the increased prevalence of the metabolic syndrome and its sequelae in this population.

1.3.2 Indian NAFLD Risk factors

Anthropometry

The links between obesity and NAFLD in general have been outlined in section 1.1.4. Whilst the average BMI has plateaued in high-income countries, dietary globalisation has led to accelerated rates of obesity across Asia, which is likely to be a contributing factor to the increased prevalence of NAFLD within these populations[155, 156]. This is a particular concern as evidence shows that South Asians are at higher risk of obesity-related morbidity and mortality compared to other ethnic groups[157] and at lower BMI [7] - the so-called Asian paradox. This resulted in the reclassification of BMI cut-offs for Asian populations in the year 2000 [4]. Data shows that for every increase in BMI of 1kg/m^2, Asians have an increased risk of diabetes/hypertension of 15% compared to 11% in Caucasians and 8% in Black ethnic groups [158].

Within the literature, it has often been suggested that Asian populations have high rates of “lean NAFLD”. However, a review of these studies suggest it is failure to utilise the Asia-specific cut-offs for BMI that give increased prevalence of disease in those considered “normal” weight[159]. Published data examining differences in prevalence of lean NAFLD suggest that rates are higher in Asia (19%)[18, 160] than in the USA (7%)[161].

Other studies looking at risk factors for biopsy-proven NAFLD in Indian populations specifically show varying levels of obesity (12-69%) [162-164] but are significantly limited by their small sample sizes. The two large Indian population-based cross-sectional studies of NAFLD within the literature also give varying rates of obesity, which is likely to be related to differences in degree of urbanisation and socioeconomic class. The Chennai Urban Population study showed that patients with NAFLD had an obesity prevalence of 36.5-76.9% which was linked to the grade of steatosis [7], whereas the rural population-based study from West Bengal showed an obesity prevalence of 25% in patients with NAFLD when only 7% of this population as a whole was classified as overweight when using Asia-specific BMI cut-offs[142]. The implications of lean NAFLD are yet unknown. Some studies suggest that those with lean NAFLD have lower rates of other metabolic components[161], whilst others suggest that those with lean NAFLD share a common metabolic and cardiovascular risk profile[165]. One longitudinal study suggests that, despite having

less metabolic complications, cumulative survival was worse in lean NAFLD than non-lean NAFLD[166].

Nevertheless, the disproportionate effect of weight gain on risk of NAFLD in the Indian population also contributes to the theory that the distribution of fat rather than overall obesity has a greater influence on NAFLD risk. Indians are documented as having more visceral adipose tissue and have a higher tendency to central adiposity compared to Caucasians [167, 168], and that waist-hip ratio correlates with risk of metabolic complications within this ethnic group[169].

Increased visceral fat is also related to increased frequency of insulin resistance, which is thought to be a precursor for NAFLD. This may be part of the "missing link" between tendency to central adiposity and increased NAFLD risk in Indians. This has been demonstrated in a number of studies, including one that examined the relationship between insulin sensitivity and visceral fat between age, gender, BMI and diet-matched healthy volunteers who were Caucasian and Indian [167]. Indians had significantly lower glucose disposal rates compared to Caucasians and had significantly greater total abdominal and visceral fat (despite being matched for BMI). There were no differences in waist hip ratio (WHR), suggesting that this method may not be sensitive enough to detect important differences in body composition. In both groups, insulin-mediated glucose disposal (insulin sensitivity) was inversely correlated with all fat compartments, suggesting that, for any increase in abdominal fat, there is a proportional change in insulin sensitivity. This work corroborated the findings of Banerji et al, who also found that healthy non-obese Indian men were more insulin resistant with high percentage body fat relative to BMI and muscle mass than matched Caucasian men [170]. Although not significant in these papers, WHR has been shown to correlate with diabetes within an Indian population, with a UK group finding that increased abdominal adiposity in Indians resulted in a 4 fold increased risk of diabetes as compared to Caucasians [169].

All the above literature suggests that Indians have more central adipose tissue, which in itself correlates to high levels of insulin resistance and therefore NAFLD.

Insulin resistance

Increasing prevalence of diabetes is not limited to Western populations, with India showing a dramatic rise in estimated diabetes prevalence from 2011 to 2030 (mean annual increment of diabetes diagnoses UK=31,000, India=2,102,000) [5]. Rates of diabetes and the metabolic syndrome in India vary according to degree of urbanisation and affluence with the odds ratio for impaired glucose tolerance (IGT) and diabetes increasing with increasing income [171]. Interestingly, this is not the case for developed countries, where higher socioeconomic status results in lower prevalence of the metabolic syndrome – which is likely to be due to better education and health awareness [172]. Data from the Chennai Urban Population Study shows that prevalence of NAFLD is higher in those with diabetes (54.5%), than those with prediabetes (33%) compared with those with normal glucose tolerance (22.5%)[7].

Multiple studies have shown that even healthy, non-obese Indians are more insulin resistant than other ethnicities[167, 170, 173]. Peterson et al. demonstrated that the prevalence of insulin resistance in healthy, young, lean Indian men is 3-4 times greater than other ethnicities and that they had higher levels of hepatic triglyceride and blood levels of the adipocytokines IL-6 and leptin, even when matched for BMI, dietary intake and activity levels[173].

The increased prevalence of insulin resistance in Indians has been confirmed in cohorts of patients with NAFLD. The NAFLD cohort from the population-based study in West Bengal had higher rates of insulin resistance (measured using the homeostatic model assessment of insulin resistance - HOMA-IR, mean 2.24 vs 1.44) despite being largely non-obese, suggesting that the strength of the physiological link between IR and NAFLD transcends anthropometric phenotypes [142]. A retrospective review of 71 cases of biopsy-proven NAFLD showed that 60% were insulin resistant[162] (although this was not an independent risk factor for NAFLD) whilst another cohort study of 51 biopsy-proven NAFLD showed 80% were insulin resistant [163]. Both studies were small, retrospective and had significant selection bias, meaning these results could be a significant over-estimate.

This wealth of evidence demonstrates that Indian populations are more insulin resistant, even at a normal BMI, which may be related to their tendency to visceral adiposity.

Lifestyle factors

Changes in dietary composition and exercise habits over the last century are thought to be contributing to the rising burden of obesity and diabetes. This trend in changing lifestyle habits is not consistent between socioeconomic classes across the globe. In developed countries, higher socioeconomic status is often associated with healthier lifestyle habits, whereas in developing countries, higher income usually equates to a more sedentary job and poorer dietary habits[172]. Within an urban population in India, those within the middle-income group (as opposed to low-income) were shown to consume more total calories, fat and sugar [171].

Evidence shows that the adverse cardiovascular risk profile related to the metabolic syndrome is associated with a diet high in carbohydrate, low in unsaturated fat and with low levels of physical activity [174], which unfortunately describes the typical Indian diet. Sudha et al. describes the Indian diet as "*predominately vegetarian, rich in complex carbohydrate and high in visible fat*", and with increasing urbanisation and affluence, the use of fat is increasing across the country [175]. As outlined in section 1.1.4, a diet rich in carbohydrate in persons with excess adipose tissue increases hepatic fat deposition and insulin resistance. This means that within the Indian population, which is already more insulin resistant (even when young, lean and healthy), their high carbohydrate diet is likely to increase risk of NAFLD further.

There are similar findings in Indian studies. One case control study conducted of sequential people attending a routine health appointment in Hyderbad, India,

including 98 NAFLD cases (diagnosed via ultrasound) and 102 controls, confirmed that the total calorie intake, percent carbohydrate and fat intake was significantly higher in cases than controls [176]. There was no difference in physical activity levels between groups. A larger case-control study done through a gastroenterology clinic in Orissa, India compared 464 consecutive patients diagnosed with NAFLD to 181 age-matched controls (who underwent ultrasonography for other reasons) [177]. NAFLD patients were more likely to consume a non-vegetarian diet (35.3% vs 22.9% p=0.002). Frequent consumption of fried food was seen more in the NAFLD group (34.9% vs 9.2% p<0.0001) as was drinking tea, which often contains large amounts of sugar (54.9% vs 39.1% p<0.001). Through logistic regression, only consumption of non-vegetarian diet and fried food was associated with NAFLD. This study was interesting in that encompassed patients from both urban and rural communities. 9.7% of the NAFLD group were manual labourers, with 45.9% having a sedentary job (57% cases middle class, 59% controls lower class). This again highlights the differing lifestyle habits in relation to socioeconomic class and their impact on NAFLD risk. Despite low levels of obesity overall within this study, they found that those with NAFLD were richer, less active and had a higher BMI (even when within the "normal" range) indicating that income and lifestyle play a large role in development of NAFLD, potentially through the increased rates of obesity in those of higher socioeconomic class (Table 1-4).

	NAFLD adjusted OR (95% CI)	**P value**
Family income		
<$1.00/day	1.0	
$1.00-$2.00/day	1.8 (1.0-3.2)	0.05
>$2/day	2.4 (1.2-5.0)	0.01
BMI		
<18.5kg/m^2	1.0	
18.5-24.9kg/m^2	2.0 (1.1-3.8)	0.03
>25kg/m^2	4.3 (1.6-11.5)	<0.001

Table 1-4. Multiple logistic regression for risk factors for NAFLD in a rural Indian population [142]

There have only been two studies looking at whether differences in BAT activity account for differences in rate of metabolic disease between Indians and Caucasians. In 2011, a Dutch group used PET-CT to examine the BAT activity of 20 healthy, lean men (10 European Caucasian, 10 South Asians) and found that there were no differences between the groups [178]. A similar study was conducted in 2013 with similar results; Bakker et al showed that, although BAT activity measured as standardized uptake values of ^{18}F-FDG via PET-CT did not differ between healthy

Caucasians and South Asians, the total BAT volume and daily energy expenditure was reduced in the South Asian group[179]. However, both these studies were conducted using healthy, lean participants, and may not give a true representative of the impact of BAT activity in those with components of the metabolic syndrome.

It is this paucity of data on the differences in BAT activity between South Asians and Caucasians and the possible link between reduced BAT activity and NAFLD that has prompted the work of Chapter 5.

Genetics

As outlined in section 1.1.4, through genome-wide association studies, a number of genetic variants have been linked to NAFLD risk. These studies include a wide variety of ethnicities, including some South Asian populations (mainly Japan), but to date, no GWAS has been done looking at genetic susceptibility for NAFLD in Indian groups. There have been however, a number of candidate-gene studies looking at some of the most well documented single-nucleotide polymorphisms (SNPs) and their link to NAFLD in this group as summarised in Table 1-4.

Study	Variant	Sample size		Diagnostic modality	Result	P value
Zain et al. [180]	PNPLA3	Case 144 Control 198		Histology	OR 2.34 (1.69-3.24)	<0.0001
Kanth et al. [181]	PNPLA3	Case 156 Control 150		Ultrasound	OR 12.66 (5.45-29.38)	0.001
Bhatt et al. [182]	PPAR	Case 162 Control 173		Ultrasound	OR 1.64 (1.09-2.45)	0.05
Peterson et al. [183]	APOC3	n = 95		H-MRS	Variant: Mean HTG 7.5 ±10.3% WT: Mean HTG 1.5±1.3%	<0.001
Puppala et al. [184]	APOC3	Case 150 Control 150		Ultrasound	OR 1.706 (1.22-2.37)	0.001
Kanth et al. [181]	APOC3	Case 156 Control 150		Ultrasound	OR 1.83 (1.02-3.28)	0.12
Tan et al. [185]	GCKR	Case 144 Control 198		Histology	OR 2.64	0.012
Kanth et al. [181]	GCKR	Case 156 Control 150		Ultrasound	OR 1.22 (0.69-2.15)	0.62
Bale et al. [186]	TM6SF2	South	Case 93 Control 138	Ultrasound	OR 2.7 (1.37-5.3)	0.0004
		North-East	Case 163 Control 109		OR 1.51 (0.86-2.66)	0.31
Kanth et al. [181]	FDFT	Case 156 Control 150		Ultrasound	OR 0.69 (0.36-1.36)	0.40

Table 1-5. Characteristics of genetic studies of NAFLD in Indian populations

Due to the limitations in sample size, and lack of whole-genome examination, these studies merely confirm association of specific single-nucleotide polymorphisms with NAFLD within Indian populations. This highlights the need for large, robust, deeply phenotyped population-based studies of NAFLD in India to enable better understanding of the links between genetic, phenotypic and lifestyle habits of this population, which results in higher rates of disease.

1.4 Conclusion

The evidence presented in this chapter highlights the need for further research to estimate accurately the prevalence of NAFLD in India, to confirm which factors influence disease risk, and to understand the impact of culture and environment on NAFLD risk profile within people of Indian origin. This thesis will therefore address the following hypotheses, using the subsequent aims and objectives.

1.5 Hypotheses, aims and objectives

- The NAFLD prevalence in India is higher than global pooled estimates
- NAFLD risk factors in India are similar to those described globally – namely obesity, insulin resistance and components of the metabolic syndrome
- Dietary composition influences NAFLD risk within India – with increased consumption of fat and carbohydrate increasing NAFLD risk
- Migration of South Asians to the UK results in alteration of NAFLD risk profile through dietary acculturation
- Indians have lower brown adipose tissue activity than Caucasians
- Reduced brown adipose tissue activity increases NAFLD risk

AIM 1: To estimate accurately the population prevalence of NAFLD in a large Indian district

Objective A) Clean and analyse data from the Trivandrum population-based study forming a nested NAFLD case-control study

Objective B) Calculate NAFLD prevalence within Trivandrum from the nested NAFLD cohort

AIM 2: To identify the impact of different NAFLD risk factors within this population

Objective C) Perform case-control analysis of NAFLD risk factors within the Trivandrum NAFLD cohort, identifying which factors are independently associated with NAFLD risk

Objective D) Analyse transient elastography data from the Trivandrum NAFLD cohort to identify which factors influence severity of disease

Objective E) Clean and analyse dietary data from the Trivandrum NAFLD cohort, creating a database which outlines the average macronutrient intake as grams per day and percentage of total calorie intake

Objective F) Summarise the dietary intake of the Trivandrum NAFLD cohort with regard to the 12 basic food groups

Objective E) Summarise the physical activity of the Trivandrum NAFLD cohort as metabolic equivalent hours per week

Objective F) Perform case-control analysis of NAFLD risk with regard to dietary intake and physical activity within the Trivandrum NAFLD cohort, identifying which lifestyle factors are independently associated with NAFLD risk

AIM 3: To compare NAFLD risk profile of native Indians with their UK migrant counterparts

Objective G) Design and create a multi-ethnic study comprised of native Indian, UK-migrant South Asian and UK-Caucasian participants with and without NAFLD, collecting data regarding NAFLD risk profile

Objective H) Compare NAFLD phenotype between native Indian, UK-migrant Indian and UK-Caucasian cohorts

Objective I) Perform case-control analysis of NAFLD risk across the multi-ethnic study

Objective J) Analyse the impact of ethnicity on NAFLD risk within the multi-ethnic study

AIM 4: To investigate the ethnic differences in BAT activity and its impact on NAFLD risk

Objective K) Assess BAT activity across the multi-ethnic study

Objective L) Compare BAT activity between native Indian, UK-migrant Indian and UK-Caucasian cohorts

Objective M) Perform case-control analysis of NAFLD risk with regard to BAT activity across the multi-ethnic study

Objective N) Analyse the impact of ethnicity on BAT activity and NAFLD risk within the multi-ethnic study

Objective O) Analyse the impact of BAT activity on diabetes risk within the multi-ethnic study

CHAPTER 2: NAFLD in a South Indian population

2.1 Introduction

As outlined in Chapter 1, NAFLD has long been considered a disease of the Western world. However, Westernisation of Asian culture has led to more sedentary behaviour and increased consumption of energy-dense foods [187] resulting in a dramatic rise in rates of obesity, diabetes and NAFLD in South Asia [18]. This is reflected in the fact that hypertension and raised fasting blood glucose are now the leading cause of non-communicable disease in India [154].

Within India, there is dramatic variation in NAFLD prevalence across different geographic, cultural and socioeconomic regions, making it hard to estimate the true population prevalence accurately. Evidence suggests that NAFLD prevalence in India is anywhere between 9-32% depending on the population of interest, study setting and method of NAFLD diagnosis utilised (section 1.2.1, Table 1-3). To date there have been no large, cross-regional population-based studies to assess NAFLD prevalence in India. Studies of this size are difficult to conduct, particularly within this lower-middle income country, which has a large rural contingent and where access to healthcare or research opportunities is limited. Studies of this nature may not even be appropriate, considering the apparent variation in disease prevalence.

In addition to the limitations of available studies in relation to sample size and availability of diagnostic tools, the documented variation in NAFLD prevalence across India is likely to be due to differences in risk profile across different regions. This may include: i) cultural differences in diet including access to energy-dense, high fat snacks ii) geographic setting - whether rural or urban domicile - which may be an indicator of socioeconomic group iii) level of physical activity in relation to work or iv) genetic predisposition based on ethnic background. For this reason, it is important not only to identify disease, but also to identify those factors that most influence risk of disease within a population. Thus, population-based studies should include collection of data on the different variables attributed to NAFLD risk in order to focus future studies and interventions. Furthermore, given the difficulties identifying population prevalence within India as a whole and likely variance in regional risk profiles, focussed interventions targeted at areas of higher rates of disease prevalence and in those with specific at-risk profiles may be more appropriate.

2.2 Hypotheses, aims and objectives

This chapter will address the following hypotheses using the subsequent aims and objectives:

- The NAFLD prevalence in India is higher than global pooled estimates

- NAFLD risk factors in India are similar to those described globally – namely obesity, insulin resistance and components of the metabolic syndrome

AIM 1: To estimate accurately the population prevalence of NAFLD in a large Indian district

Objective A) Clean and analyse data from the Trivandrum population-based study forming a nested NAFLD case-control study

Objective B) Calculate NAFLD prevalence within Trivandrum from the nested NAFLD cohort

AIM 2: To identify the impact of different NAFLD risk factors within this population

Objective C) Perform case-control analysis of NAFLD risk factors within the Trivandrum NAFLD cohort, identifying which factors are independently associated with NAFLD risk

Objective D) Analyse transient elastography data from the Trivandrum NAFLD cohort to identify which factors influence severity of disease

2.3 Methods

2.3.1 Study design

The Trivandrum NAFLD cohort was originally designed and set up in 2013 to examine the interaction between genetics and lifestyles factors that result in increased risk of NAFLD within this population. The cohort also gives an accurate estimation of population prevalence of NAFLD and enables analysis of impact of different variables on NAFLD risk. The development of this cohort also facilitates further disease-specific studies, including interventional studies, and has the capability of generating longitudinal follow-up data.

Trivandrum (Thiruvanathapuram) is the southern-most district in the state of Kerala, located on the south-west coast of India . At the time of the Indian census in 2011, it had a population of 3.3 million divided into urban and rural domiciles (urban 54%, rural 46%) [188]. The Trivandrum NAFLD cohort was created between February 2013 and July 2016 through population-based sampling of all inhabitants over the age of 25 years.

As described, Trivandrum has both urban and rural districts. The rural districts are divided into 12 blocks, which are themselves formed of 78 panchayats (originally denoting a village council). Panchayats are then subdivided further into wards (originally denoting a local authority area – usually indicating a neighbourhood). The urban districts are divided into wards only.

The enrolment of study participants was through un-weighted multi-stage cluster sampling. Using random number tables, 3 out of 12 rural blocks were selected (made up of 22 panchayats), 4 rural panchayats were selected out of 22 (made up of 71 wards), and finally 40 out of 71 wards were selected. Using the same technique, 7 out of 81 urban wards were selected (Figure 2-1).

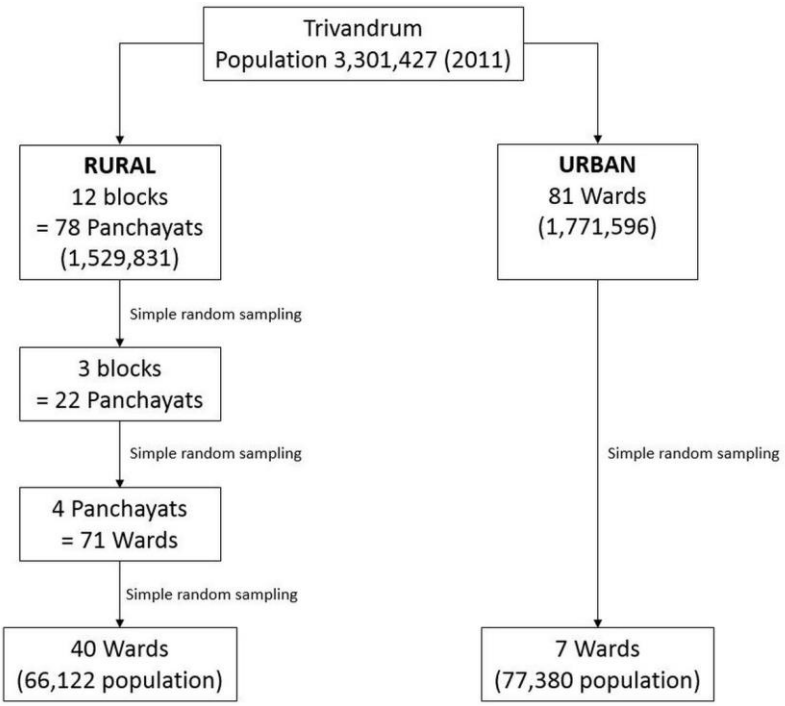


Figure 2-1. Trivandrum NAFLD cohort sampling framework

Using the Trivandrum electoral roll from 2011, households within these wards were grouped into clusters of seven. Using random number tables, 112 of 2,420 rural clusters and 128 of 2,631 urban clusters were identified for sampling. Each household within the chosen clusters was sampled and, through house-to-house survey by local social workers, all eligible inhabitants were invited to enrol. For some households, this meant repeated visits to ensure all those who were eligible (anyone over the age of 25 years) were given the opportunity to participate. Originally, the aim was to recruit around 1000 cases of NAFLD, enabling a genome-wide association study with 80% power to detect an odds ratio of 1.6 for genetic variants with a mean allele frequency (MAF) of at least 20% (OR 1.8 for MAF of 10%, OR 1.5 for MAF of 35%) with a genome-wide significance of $p<1x10^{\wedge-6}$, in line with previous NAFLD GWAS[189, 190]. Based on an estimated NAFLD prevalence of 31.8% from a previous pilot study [191], this meant recruiting a total cohort of 3,125 participants. Due to time constraints, cohort recruitment at household-level was ceased at 2,222 participants (Figure 2-2), which, based on estimated prevalence, would result in 711 NAFLD cases.

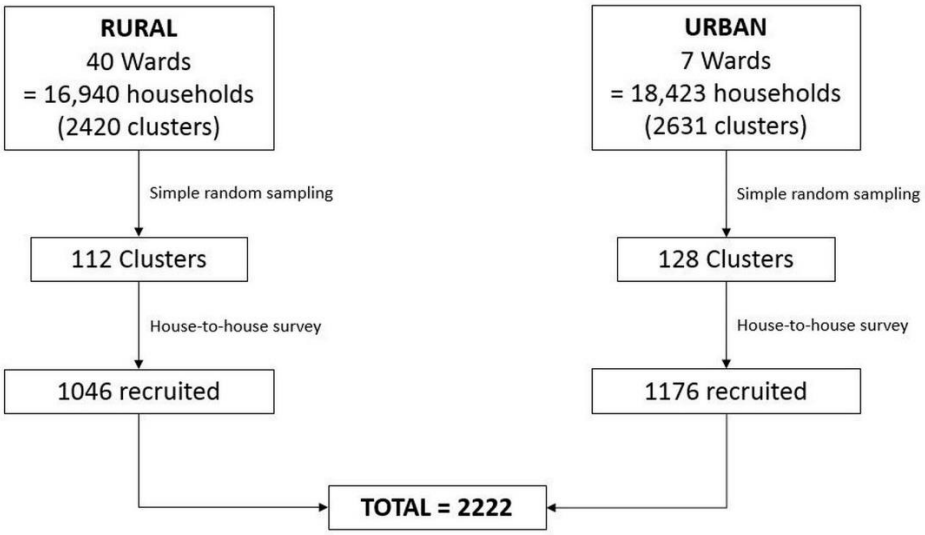


Figure 2-2. Trivandrum NAFLD cohort recruitment diagram

2.3.2 Ethical approval

Ethical approval for the study was granted by the Sree Gokulam Medical College and Research Foundation, Venjaramoodu, Trivandrum ethics committee. This study and all relevant documentation received approval from the University of Nottingham Faculty of Medicine and Health Sciences Research Ethics Committee (REC: 26/299/05/2017, 14/06/17) and the Nottingham University Hospitals NHS Trust Research and Development department. As participants were not NHS patients, no NHS ethics was required or obtained.

3.3.3 Consent

All participants provided written informed consent at the time of recruitment, having had the opportunity to discuss the study in full with those taking consent (particularly those that were illiterate) and to receive written information about the study. Participants had the opportunity to ask questions prior to signing consent and before any data was collected.

2.3.4 Data collection

Clinical data

Potential participants were approached and consented to the cohort through house-to-house survey by social workers within their wards. Initial clinical data were collected at the time of consent within the participant's home. For clarity, variables were grouped into categories and data were collected using the following definitions:

<u>Demographics</u>

- Date of birth (for age)
- Gender
- Domicile (urban or rural)

- Level of education:
    i. Illiterate
    ii. Just literate
    iii. Primary school
    iv. Secondary school
    v. College (but not graduate)
    vi. Graduate or professional[ŧ]
    vii. Postgraduate[ŧ]
    viii. Unknown
- Work status
    i. Student
    ii. Not currently working for pay
    iii. Working full time
    iv. Unable to work due to disability
    v. Retired
- Religion
- Marital status
    i. Single
    ii. Married
    iii. Divorced
    iv. Widowed

ŧ Entry into higher education (graduate or above) was used as an indicator of higher socioeconomic position. The use of education as a socioeconomic indicator originates from Weberian theory[192]. It captures knowledge-related assets of a person and is a strong determinant of future employment and income[193].

Anthropometry

- Height: measured using a wall-mounted vertical rule, standing height was measured to the nearest centimetre (cm), standing without shoes.
- Weight: measured using mechanical standing-scales to the nearest kilogram (kg) without outdoor clothes or shoes.
- Waist circumference: measured using a non-expandable tape measure midway between the lower rib margin and iliac crest during exhalation. The average of two readings were taken to the nearest cm. Central obesity was defined as ≥80cm in women and ≥90cm in men (as per Indian consensus statement and International Diabetes Federation (IDF) ethnic-specific metabolic syndrome definition)[194].

Medical history

Participants were asked if they had a history of the following conditions, and if so, whether they had undergone treatment (no treatment, previously treated, under treatment)

a) Bronchitis
b) Asthma
c) Tuberculosis
d) Hypertension
e) Ischaemic heart disease (IHD) – history of chest pain or myocardial infarct
f) Cerebrovascular disease (CVD) – history of cerebrovascular accident
g) Peripheral vascular disease (PVD)
h) Hyperlipidaemia
i) Diabetes
j) Thyroid dysfunction
k) Peptic ulcer disease
l) Cancer

Lifestyle

- Tobacco smoking (including pack/years)
    i. Current smoker
    ii. Past smoker
    iii. Never smoked
- Tobacco chewing
    i. Current chewer
    ii. Past chewer
    iii. Never chewed
- Alcohol intake (alcohol excess defined as >21units per week)
    i. Never
    ii. Drank alcohol in last year
    iii. Average units per week

Pathological data

To collect pathological data, study camps were held within the area local to the selected wards and those who had been recruited were invited to attend.

Participants who attended the study camps had fasting bloods taken for the following:

a) Liver function tests

- Alanine transferase (ALT)
- Aspartate aminotransferase (AST)
- Gamma glutamyl transferase (GGT)

b) Platelets

c) Lipids

- Cholesterol
- Triglycerides (TG)
- Low density lipoproteins (LDL)

- High density lipoproteins (HDL)
- Very-low density lipoproteins (VLDL)

d) Metabolic parameters

- Fasting glucose
- Fasting insulin

Participants also underwent abdominal ultrasonography to assess presence and grade of fatty infiltration (Grade 1: Liver echogenicity is increased. Grade 2: Echogenic liver obscures the echogenic walls of portal vein branches. Grade 3: Echogenic liver obscures the diaphragmatic outline) [36]. A qualified radiographer using a portable ultrasound machine undertook ultrasound at the study camps. Due to logistical constraints, a specific ultrasonographer was not used at each site, and the ultrasound scans were not crosschecked.

Those who had evidence of NAFLD (of any grade) subsequently underwent TE using FibroScan® (Echosens, Paris, France) as a surrogate marker to stage degree of fibrosis. Based on a paper by Das et al, which analysed liver stiffness measurements compared to histology within an Indian population, a reading of ≤8.4kPa is considered low risk of fibrosis, 8.5-11.6kPa intermediate risk and ≥11.7kPa high risk (NAFLD fibrosis stage >2, AUROC 0.965)[195]. Whilst this study utilised a sample from a population-based study within India, its sample size was reasonably small (n=625), reducing confidence in the results. It is however the only study examining the accuracy of TE in an Indian population and its cut-offs were therefore considered the most appropriate for this study, rather than using data from Caucasian populations. FibroScans® were undertaken by an operator who had been formally trained by Echosens. The operator obtained ten valid liver measurements and the median liver stiffness was calculated. The IQR was calculated and the TE reading was considered valid if the IQR was less than 30% of the median stiffness[196].

2.3.5 Data management

All study staff and investigators endeavoured to protect the rights of the study participants to privacy and informed consent, and adhered to the Data Protection Act, 1998.

Detailed identifiable data were collected for each participant including name (including father's name – which is taken as surname), address and date of birth. Clinical data were collected during the household visit using paper case report forms (CRF) and pathological data were collected from laboratory or imaging reports. Data from these forms were then transcribed onto an electronic Access database (Microsoft®, Redmond, USA). The minimum required information for the purposes of the study were recorded on the CRF and held securely, in a locked room, or locked cupboard. Access to the information was limited to the study staff, investigators and relevant regulatory authorities. Electronic data including the study database were held securely and password protected.

2.3.6 Data cleaning

Data from the original Access database were converted into STATA format (version 15, Statacorp, Texas, USA). Data were examined for duplicates, using name, fathers name, gender and postcode. Any true duplicates were deleted. Each participant was then allocated a new unique identifier number for purposes of data analysis and any identifiable data was removed from the database.

The distribution of continuous variables was examined for normality and outliers. Those outlier results that were considered implausible were deleted and treated as missing, but plausible data were allowed to remain and its use decided upon dependant on the individual research question.

2.3.7 Data analysis

Due to the large sample size and the use of random sampling, central limit theorem suggests that continuous data are presumed to be normally-distributed[197] and as such were presented as mean (SD). For variables with clear non-normality of data when plotted via histogram [198] (e.g. ALT and age), data were presented as median (IQR). Categorical data were presented as number (%).

Univariate analysis was undertaken using unpaired Students t-test of equal variance (normal continuous), Mann-Whitney U test (non-normal continuous) and Chi squared test (categorical) to compare the following: i) cohort characteristics to population data ii) numbers of cases and controls (population prevalence) iii) characteristics of the cohort with/without NAFLD iv) characteristics of urban and rural districts and v) characteristics of those with/without fibrosis who had NAFLD. The analyses were not adjusted to account for missing data points, as these numbered less than ten for any given data set.

Univariate logistic regression was used to determine the odds for NAFLD in relation to each variable. This was then used to calculate unadjusted odds ratios (OR), presented with a 95% confidence interval (CI) and p value. BMI and waist circumference were presented as both continuous and categorical variables, (as per BMI cut-offs and categorised as presence or absence of central obesity), and components of the metabolic syndrome were presented individually. Forward, stepwise logistic regression was then performed with significance set to $p<0.05$ to identify which variables independently influenced NAFLD risk. All variables that increased NAFLD risk on univariate analysis were included in the model, with BMI and central obesity as categorical variables. Baseline categories for ordinal variables were as presented e.g. male sex, normal BMI and fibrosis risk (in those with NAFLD).

2.4 Results

Of the original 2222 population sample, 2161 participants attended the camps – demonstrating a high attendance rate and only 3% dropout. Three participants did not have an ultrasound so were excluded, leaving the total number of study sample participants as 2158. Analysis of the demographic/anthropometric data of those who did and did not have an ultrasound shows that a higher proportion of men (61% vs

40.6% p<0.001) and a higher proportion of people within the rural setting (73.4% vs 46.3% p<0.001) did not attend. This was likely to be due to difficulty in getting to the rural camps, which covered a larger geographical area than the urban camps, and that men were often limited by their need to work.

Demographic characteristics of the study sample (n=2158) are presented in Table 2-1 to enable comparison with data from the Trivandrum population census of 2011. There was a higher proportion of women in the study sample compared to the population census. The proportion of rural and urban population sampled was the same as the population census, demonstrating adequate sampling within the study. Literacy rates were higher in the study sample, as was the proportion of each main religion.

	CENSUS n=3,301,427	STUDY SAMPLE n=2158	P value
Age (years, mean, SD)	N/A	47.22 (11.56)	
Sex (n, %)			
- Male	1,581,678 (47.9)	876 (40.6)	<0.001
- Female	1,719,749 (52.1)	1282 (59.4)	
Domicile (n, %)			
- Rural	1,529,831 (46.3)	999 (46.3)	0.966
- Urban	1,771,597 (53.7)	1159 (53.7)	
Education (n, %)			
- Illiterate		31 (1.4)	
- Just literate		13 (0.6)	
- Primary school		118 (5.5)	
- Secondary school		398 (18.4)	
- High school		861 (39.9)	
- College		363 (16.8)	
- Graduate/professional		314 (14.5)	
- Postgraduate		59 (2.7)	
- Unknown		1 (0.05)	
Literacy (%)	93%	98%	<0.001
Work status (n, %)			
- Student	N/A	19 (0.88)	
- Not working		1073 (48.29)	
- Working full time		940 (42.30)	
- Unable to work		62 (2.87)	
- Retired		64 (2.97)	
Religion (n, %)			
- Hindu	2,194,057 (66.5)	1765 (81.79)	<0.001
- Christian	630,573 (19.1)	226 (10.47)	<0.001
- Muslim	452,915 (13.7)	167 (7.74)	<0.001
Marital status (n, %)			
- Single	N/A	88 (4.08)	
- Married		1966 (91.10)	
- Divorced		8 (0.37)	
- Widowed		95 (4.40)	

Table 2-1. Comparison of the study sample demographics to the Trivandrum population census.

The purpose of the creation of the Trivandrum NAFLD cohort was to enable accurate estimation of NAFLD prevalence within this population. Through population-level sampling, the most representative cohort possible was created. Through ultrasound assessment of liver echogenicity, the prevalence of an echo-bright liver within this population was 49.8%.

To enable analysis of variables that may influence risk of NAFLD within this cohort, it was necessary to ensure that those with NAFLD had not been identified in error due to the presence of an alternative cause of an echogenic liver. For this reason, those participants who had a documented cause of liver disease (e.g. viral hepatitis) or who were at risk of alcohol-related fatty liver (anyone who consumed >21 units of alcohol per week) were excluded leaving a final cohort of n=2089 (NAFLD n=1040, Control n=1049, NAFLD prevalence 49.8%)(Figure 2-3).

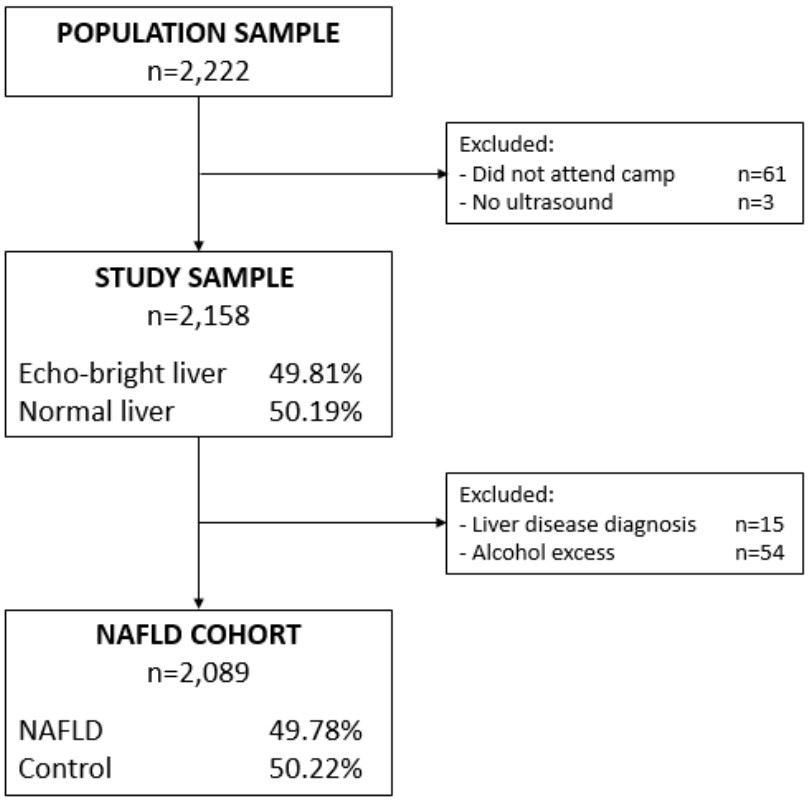


Figure 2-3. NAFLD prevalence of study sample (49.81%) and creation of NAFLD cohort (n=2,089, NAFLD prevalence 49.78%).

2.4.1 Clinical data

Table 2-2 outlines the anthropometric, medical history and social data of the NAFLD cohort, with statistical testing to demonstrate whether any differences between NAFLD and control groups were statistically significant ($p<0.05$).

	NAFLD n=1040	Control n=1049	P value
Demography			
Age (years, mean SD)	48.34 (10.74)	45.98 (12.31)	<0.001
Gender (n, %)			
- Male	473 (45.5)	339 (32.3)	<0.001
- Female	567 (54.5)	710 (67.7)	
Domicile (n, %)			
- Rural	417 (40.1)	543 (51.8)	
- Urban	623 (59.9)	506 (48.2)	<0.001
Higher education (n, %)	200 (19.2)	166 (15.8)	0.04
Anthropometry			
Height (cm, mean SD)	159.34 (10.23)	158.12 (8.48)	<0.001
Weight (kg, mean SD)	68.35 (11.79)	60.77 (11.20)	<0.001
BMI (kg/m^2, mean SD)	26.90 (4.16)	24.30 (4.08)	<0.001
BMI category (n, %)			
- Underweight ($<18.5kg/m^2$)	4 (0.38)	68 (6.48)	<0.001
- Normal weight ($18.5\text{-}22.9kg/m^2$)	165 (15.87)	336 (32.03)	<0.001
- Over weight ($23\text{-}24.9kg/m^2$)	195 (18.75)	222 (21.16)	0.17
- Obese ($>25kg/m^2$)	675 (64.90)	423 (40.32)	<0.001
- Missing	1 (0.10)	0 (0)	0.32
Waist circumference (cm, mean SD)	95.85 (9.45)	88.87 (9.94)	<0.001

(continued)

	NAFLD n=1040	Control n=1049	P value
Past Medical History			
Metabolic (n, %)			
- Diabetes	355 (34.13)	195 (18.59)	<0.001
- Hypertension	419 (40.29)	310 (29.55)	<0.001
- Dyslipidaemia	727 (69.9)	565 (53.86)	<0.001
Cardiovascular disease (n, %)			
- IHD	31 (2.98)	29 (2.76)	0.77
- CVD	4 (0.38)	2 (0.19)	0.41
- PVD	2 (0.19)	3 (0.29)	0.66
Thyroid disease (n, %)	63 (6.06)	62 (5.92)	0.89
Lung disease (n, %)			
- Asthma	36 (3.46)	31 (2.96)	0.51
- Bronchitis	3 (0.29)	3 (0.29)	1.50
- Tuberculosis	1 (0.01)	1 (0.01)	1.95
Peptic ulcer disease (n, %)	4 (0.39)	10 (0.96)	0.11
Cancer (n, %)	3 (0.29)	2 (0.19)	0.65
Social history			
Tobacco (smoking n, %)			
- Never	911 (87.60)	940 (89.61)	0.15
- Past	112 (10.77)	95 (9.06)	0.19
- Current	17 (1.63)	14 (1.33)	0.57
Tobacco (chewing n, %)			
- Never	1007 (96.83)	1015 (96.76)	0.94
- Past	1 (0.10)	2 (0.19)	0.57
- Current	32 (3.08)	32 (3.05)	1.07
Alcohol (n, %)			
- Never	874 (84.04)	947 (90.28)	<0.001
- Drank in past year	166 (15.96)	102 (9.72)	<0.001

Table 2-2. Baseline anthropometric, medical history and social data of the NAFLD cohort

Abbreviations: BMI, body mass index; IHD, ischaemic heart disease; CVD, cerebrovascular disease; PVD, peripheral vascular disease.
Higher education = graduate or above; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy; hypertension = history of hypertension, antihypertensives, systolic blood pressure >130, diastolic blood pressure >85. .

Overall, females outnumbered males and constituted 1277 of the total 2089 participants (61.1%). 55.2% of the urban population had NAFLD compared to 43.4% of the rural population. Those with NAFLD were taller, heavier and had a higher BMI than the control group, although the proportion of participants who were categorised as overweight was the same between NAFLD and control groups. Waist circumference was also larger in the NAFLD group. There were higher rates of all components of the metabolic syndrome within the NAFLD group. Despite this, rates of IHD, CVD and PVD are the same. Rates of thyroid disease, lung disease, peptic ulcer disease and cancer were the same between groups, as were rates of smoking. The proportion of participants who had never drunk alcohol was higher in the control group with rates of alcohol intake over the last year being lower.

On analysis of the impact of these variables on NAFLD risk, living within an urban domicile was associated with increased risk of NAFLD with an unadjusted OR of 1.60 (1.35-1.91). An increase in BMI was associated with increased odds of NAFLD (unadjusted OR 1.79 if overweight; OR 3.25 if obese) and, for every unit increase in BMI, the odds of having NAFLD appeared to increase by 17% (unadjusted OR 1.17, 1.15-1.20).

Presence of diabetes (unadjusted OR 2.27, 1.85-2.79), dyslipidaemia (unadjusted OR 1.99, 1.66-2.79) and hypertension (unadjusted OR 1.61, 1.34-1.93) were also associated with increased odds of disease. A waist circumference above that outlined in the Asian-specific IDF definition for metabolic syndrome appeared to increase odds of NAFLD (unadjusted OR 2.47, 1.98-3.10), and for every cm increase in waist circumference, the odds of NAFLD appeared to increase by 7% (unadjusted OR 1.07, 1.07-1.09). Following multivariate logistic regression adjusting for all variables as potential confounders, living within an urban domicile and presence of hypertension were no longer associated with an increased risk of NAFLD. These data are presented in Table 2-3.

		Unadjusted OR (95% CI, p value)	Adjusted OR (95% CI, p value)
Age (years)	1.02	(1.01, 1.03 p<0.001)	1.01 (1.00, 1.02 p=0.029)
Male sex	1.75	(1.46, 2.09 p<0.001)	2.29 (1.86, 2.83 p<0.001)
Urban domicile	1.60	(1.35, 1.91 p<0.001)	1.18 (0.98, 1.44 p=0.084)
Higher education	1.27	(1.01, 1.59 p=0.036)	1.30 (1.00, 1.68 p=0.047)
BMI (kg/m^2)	1.17	(1.15, 1.20 p<0.001)	
- Underweight *<18.5kg/m^2*	0.12	(0.04, 0.34 p<0.001)	0.18 (0.06, 0.52 p=0.001)
- Overweight *23-24.9kg/m^2*	1.79	(1.36, 2.34 p<0.001)	1.55 (1.15, 2.08 p=0.004)
- Obese *>25kg/m^2*	3.25	(2.58, 4.10 p<0.001)	2.81 (2.15, 3.68 p<0.001)
Waist circumference (cm)	1.07	(1.07, 1.09 p<0.001)	
- Central obesity	2.47	(1.98, 3.10 p<0.001)	1.60 (1.19, 2.15 p=0.002)
Diabetes	2.27	(1.85, 2.79 p<0.001)	1.76 (1.40, 2.21 p<0.001)
Dyslipidaemia	1.99	(1.66, 2.39 p<0.001)	1.70 (1.39, 2.06 p<0.001)
Hypertension	1.61	(1.34, 1.93 p<0.001)	1.14 (0.93, 1.41 p=0.204)

Table 2-3. Unadjusted and adjusted odds ratios for NAFLD risk factors within the cohort

Abbreviations: BMI, body mass index
Higher education = graduate or above; central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy; hypertension = history of hypertension, antihypertensives, systolic blood pressure >130, diastolic blood pressure >85.

To understand the interaction between urban-living and other risk factors for NAFLD, the rates of obesity, diabetes and levels of higher education for each domicile are presented in Table 2-4.

	OVERALL n=2089	URBAN n=1129	RURAL n=960	P value
BMI >25kg/m^2 (n, %)	1098 (53.58)	696 (61.65)	402 (41.88)	<0.001
Diabetes (n, %)	550 (26.84)	326 (28.88)	224 (23.33)	0.004
Higher education (n, %)	366 (17.52)	237 (20.99)	129 (13.44)	<0.001

Table 2-4. NAFLD risk factors as per domicile

Abbreviations: BMI, body mass index
Diabetes = Diagnosis of diabetes &/or fasting glucose >126mg/dl; higher education = graduate or above. P value between urban and rural groups.

Within the urban domicile, there were higher rates of obesity (61.7% vs 41.9%), higher rates of diabetes (28.9% vs 23.3%) and higher rates of people with higher education/of higher socioeconomic class (21% vs 13.4%).

2.4.2 Pathological data

Pathological data for the NAFLD cohort are presented in Table 2-5.

	NAFLD n=1040	Control n=1049	P value
Liver function tests			
ALT (U/l median IQR)	31 (25)	19 (12)	<0001
AST (U/l median IQR)	29 (15)	25 (10)	<0.001
GGT (U/l median IQR)	25.7 (18.4)	17.25 (9.8)	<0.001
Platelets			
Platelets (10^9/l mean SD)	249 (70)	247 (66)	0.68
Lipids			
Cholesterol mg/dlmean (SD)	219.19 (44.05)	212.26 (44.34)	1.00
Triglycerides mg/dl median (IQR)	133 (85)	94.5 (59)	<0.001
HDL mg/dl median (IQR)	46 (14)	49 (15)	<0.001
LDL mg/dl mean (SD)	141.02 (37.50)	138.46 (38.24)	0.94
LDL-C mg/dl mean (SD)	140.71 (39.04)	140.03 (38.87)	0.69
VLDL mg/dl median (IQR)	27 (16)	19 (12)	<0.001
Metabolic parameters			
Glucose mg/dl median (IQR)	107 (39)	98 (11)	<0.001
Insulin mU/l median (IQR)	7.8 (7.7)	4.8 (4.8)	<0.001
HOMA-IR median (IQR)	2.14 (2.14)	1.17 (1.2)	<0.001

Table 2-5. Pathological data for NAFLD cohort

*Abbreviations: ALT = alanine transferase, AST = aspartate aminotransferase, GGT = Gamma-glutamyl transpeptidase, HDL = high-density lipoproteins, LDL = low-density lipoproteins, LDL-C = low-density lipoprotein cholesterol, VLDL = very low-density lipoproteins, HOMA-IR = homeostatic model assessment of insulin resistance HOMA-IR ((fasting insulin*fasting glucose)/405) calculated for those not on treatment for diabetes. LDL-C = (total cholesterol-HDL cholesterol-(triglycerides/5))*

All liver function tests were higher in the NAFLD group compared to controls. Platelet counts were similar between groups. Triglyceride levels were higher in the NAFLD group with corresponding lower levels of HDL. Both fasting glucose and fasting insulin levels were higher in the NAFLD group. In those without diabetes, levels of insulin resistance were higher in the NAFLD group as demonstrated by the increased HOMA-IR.

Originally, only participants with NAFLD were to be assessed for degree of fibrosis using FibroScan®. Due to cultural influences, a number of control group members were also scanned. The results of these data are presented in Table 2-6.

Fibrosis risk	**NAFLD n=1040**	**Control n=1049**
Low risk (≤8.4kPa)	543	427
Intermediate risk (8.5-11.6kPa)	114	27
High risk (≥11.7kPa)	47	16
No result	336	579

Table 2-6. Transient elastography results for the NAFLD cohort

Of those with NAFLD who underwent TE, 161 (15.5%) had evidence of potentially significant liver disease (>8.4kPa), of which 47 (4.5% of those with NAFLD) had evidence of advanced chronic liver disease (≥11.4kPa). Interestingly, 4.1% of controls also had raised TE with 1.5% potentially having advanced chronic liver disease in the absence of NAFLD. This data should however be interpreted with caution, as there may be a degree of selection bias toward those who thought, or were perceived to be at higher risk of liver disease for whatever reason.

Subsequent analysis of TE data validity revealed that 8.0% of FibroScan® readings had an IQR <30% of the median reading. This number dropped to 3.25% when an IQR cut-off of <40% was used.

Further analysis was undertaken to identify which of the contributory risk factors identified in Table 2-3 were associated with presence of fibrosis (TE>8.4kPa) in those with NAFLD who underwent TE (n=704). These findings are presented in Table 2-7.

	NAFLD with TE result n=704		
	No Fibrosis (≤8.4kPa) n=543	Fibrosis (>8.4kPa) n=161	P value
Age (years) Mean (SD)	48.4 (10.6)	50.4 (10.3)	0.04
Male sex n (%)	221 (40.7)	86 (53.4)	0.004
BMI (kg/m2) Mean (SD)	26.6 (3.9)	28.5 (4.8)	<0.001
Central obesity n (%)	464 (85.5)	148 (91.9)	0.032
Diabetes n (%)	174 (32.0)	75 (46.6)	<0.001
Dyslipidaemia n (%)	385 (70.9)	117 (72.7)	0.66

Table 2-7. Comparison of NAFLD patients with and without significant fibrosis

Abbreviations: BMI = body mass index
Central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy

Patients with evidence of fibrosis were older, had a higher BMI, were more likely to have central obesity, be male and diabetic. Those variables that appeared to influence risk of fibrosis within this group were combined within a logistic regression model to give adjusted odds ratios (Table 2-8).

	Adjusted OR (95% confidence interval)	**P value**
Age (years)	1.02 (0.99, 1.04)	0.075
BMI (kg/m^2)	1.13 (1.08, 1.18)	<0.001
Central obesity	1.75 (0.89, 3.44)	0.103
Male sex	2.60 (1.75, 3.86)	<0.001
Diabetes	1.80 (1.22, 2.64)	<0.001

Table 2-8. Adjusted odds ratios for presence of fibrosis

Abbreviations: BMI = body mass index
Central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl

Table 2-8 shows that increasing age was not independently associated with increased NAFLD risk (adjusted OR 1.02 0.99-1.04), neither was presence of central obesity (adjusted OR 1.75 0.89-3.44). Male sex was associated with the highest risk of fibrosis (adjusted OR 2.60 1.75-3.86 p<0.001), followed by presence of diabetes (adjusted OR 1.80 1.22-2.64 p<0.001), then increasing BMI (adjusted OR 1.13 1.08-1.18 p<0.001).

2.5 Discussion

2.5.1 Principal findings

The principal finding of this body of work was that the prevalence of NAFLD within this cohort was 49.8%, which through population level sampling (Table 2-1) is likely to be representative of the population as a whole. Univariate analysis demonstrated an increased odds of NAFLD in those of increased age, male sex, with higher education living in the urban domicile, of higher BMI and with components of the metabolic syndrome. Through adjusted multivariate analysis, male sex, increased measures of body composition (BMI and waist circumference) and presence of diabetes or hyperlipidaemia appeared to be independent risk factors for NAFLD. Those living within the urban domicile were more likely to have higher education, have a higher BMI, and have diabetes but interestingly did not have increased rates of NAFLD as is seen within the literature (Table 1-3). Of those with NAFLD, odds of fibrosis were higher in those of increased age, male sex, increased BMI and with diabetes, although increased age was no longer a significant contributing factor when data was analysed through adjusted logistic regression.

These findings demonstrate a commonality of risk factors with NAFLD studies from India and across the world, with the most significant risk being conferred by male sex (adjusted OR 2.29, 1.86-2.83, p<0.001) and presence of obesity (adjusted OR 2.81, 2.15-3.68, p<0.001).

Prevalence

As described, the NAFLD prevalence, as measured by ultrasound within this population, was 49.8%. This is significantly higher than both the global pooled NAFLD prevalence of 25%[1] (p<0.001) and estimates from other population-based studies in India, which range from 9-32%[7, 142, 144].

To the author's knowledge, this is the largest study of NAFLD prevalence within India. The size of this study and robust population-level sampling framework mean this estimate is likely to be accurate within this adult population. In addition to these strengths, this study also sampled both urban and rural domiciles at a population level when compared to the Indian census of 2011 (Table 2-1). Whilst other large studies in India also utilised population-based sampling through electoral rolls[142], or sampling of all adult inhabitants of a particular area[144], they were conducted in

either rural or urban regions respectively and, as such, show significant variation in reported NAFLD prevalence. As expected, the NAFLD prevalence from this study was higher than those conducted in rural regions (prevalence 8.7%-30.7%) and lower than studies conducted in urban tertiary units (prevalence 49-87%) (Table 1-2). Of note, all of these large population-based studies also sample adults only – although cut-offs vary from 18 years[142, 144] to 20 years[7] – meaning this study may have overestimated prevalence due to a marginally older group (age >25 years).

The NAFLD prevalence reported in this work is based on the use of ultrasound as a diagnostic tool, and whilst this is not the gold standard, its use within this study does however make it comparable to the other large population-based studies in India[7, 142, 144].

Body composition

Obesity has long been established as a risk factor for NAFLD. These pathological processes have been outlined in detail within the literature, and demonstrated clinically through observational studies that show that NAFLD prevalence is up to 90% in obese adult populations[199]. This finding is highlighted within this cohort, with 64.9% of those with NAFLD being obese (BMI>25kgm^2) compared to 40.3% of the control group. This gave an adjusted odds ratio of 2.81 (2.15-3.68, p<0.001). Rates of obesity were higher within the urban domicile compared to rural domicile (61.7% vs 49.5%), which could be thought to increase NAFLD risk within these areas. Through multivariate analysis however, urban living was no longer associated with increased NAFLD risk, as there were similar rates of obesity in those with NAFLD in both urban and rural regions (62.8% vs 58.5%).

Interestingly, rates of obesity within this study population were much higher than those quoted by the WHO Non-Communicable Disease Collaboration. 52.6% of the Trivandrum NAFLD cohort (both NAFLD and controls) were obese compared to the Indian national estimate of 3.9%[200] - although national data incorporates all ages, compared to this pre-selected cohort of adults. Data from the National Family Health Report in India also show that rates of obesity vary according to domicile, with more adults being obese in urban areas (27-31%) compared to rural areas (14-15%)[201], a difference which was not apparent within this cohort (see paragraph above). Kerala is specifically quoted as one of the regions with highest obesity rates (29-32%), although this is still lower than the rates seen within this cohort.

Although obesity rates were higher within this population compared to national averages, rates were similar to those quoted by other population-based studies in India. Mohan et al. (urban) quoted rates of obesity as 34.4% and reported an unadjusted NAFLD odds ratio of 2.0[7]. Amarapukar et al. (also urban) had an obesity prevalence of 32%[144]. Das et al. showed significantly lower rates of obesity within their rural population, with only 0.9% of people being obese[142]. However, the odds ratio for NAFLD in relation to obesity in their cohort was significantly higher (OR 4.3 1.6-11.5) even in the context of a normal BMI (OR 2.5 1.4-4.6).

Similarly, increasing waist circumference (unadjusted OR 1.07, 1.07-1.09, p<0.001) and central adiposity (adjusted OR 1.60, 1.19-2.15, p=0.002) were also associated with NAFLD risk within this cohort. This finding is in line with data from the literature which suggests that tendency to central adiposity in Indians may increase insulin resistance and predisposition to NAFLD[169].

Data from this cohort, along with that from the other large works published in the literature, further highlight the impact of the Asian-paradox – where people of Asian origin are at increased risk of metabolic and cardiac complications in relation to smaller increases in BMI. In addition, these data add to the increasing body of evidence that rates of obesity and NAFLD may be higher in urban regions, which could be related to income and socioeconomic position.

Demographics

Within this cohort, when history of higher education was used as a surrogate for socioeconomic position[192], those that lived within the urban domicile were of higher socioeconomic status and as such were likely to have higher income[193]. When analysed in isolation, living in an urban environment was associated with increased risk of NAFLD, but the significance of this was lost when adjusted for confounders through stepwise logistic regression, and the association may instead be related to an increased BMI in those living within the urban domicile.

Evidence from other large population-based studies suggest that rates of obesity, diabetes and NAFLD are closely linked to socioeconomic status. It has already been noted that within the rural district sampled by Das et al. rates of obesity were low, but that small changes in BMI equated to large increases in NAFLD risk[142]. The same was demonstrated for socioeconomic position, which was measured using household income (<$2 per day was considered economically poor). This work showed that earning more than $2 per day resulted in a more than two-fold increase in risk of NAFLD (adjusted OR 2.4, 1.2-5.0, p=0.01). Amarapukar et al. surveyed the slum settlements along the Mumbai railway lines[144]. Whilst economically poor, these areas are urban, and rates of obesity, diabetes and NAFLD were higher than the rural-poor. This is likely to be due to the availability of high energy-dense foods and the less active working environments. Kerala, as a state, has a higher per capita income (88,527 Rupees) than the average for India as a whole (74,380 Rupees), which may also influence rates of obesity and its metabolic consequences in this region.

Insulin resistance

Presence of diabetes was significantly associated with NAFLD risk within this cohort (adjusted OR 1.76 1.40-2.21, p<0.001), with rates being higher in the urban domicile compared to rural. The physiological link between obesity, insulin resistance and NAFLD are well documented[50], and as such this finding is in line with worldwide literature.

Of note, are the high rates of diabetes within this cohort (26.8% overall), which may be in some part linked to the higher rates of obesity as outlined above. 31.1% of NAFLD patients and 18.6% of the control group had diabetes – which is significantly higher than those documented in the IDF atlas (8.3%-9.9%)[5]. Again, rates of diabetes in relation to domicile and socioeconomic position reflect those found in the other large Indian studies of NAFLD – with Mohan's urban cohort having rates of 24.4% using the same criteria[7]. Within this cohort, the diabetes rates may be due to the cohort being formed of participants over the age of 25 years, an age group where prevalence of type 2 diabetes is expected to be substantially higher. In addition, collection of fasting blood glucose levels at enrolment identified those who met the IDF criteria for diabetes without previous diagnosis. Alternatively, it may be postulated that it is these high rates of diabetes that predispose this cohort to such high rates of NAFLD.

Pathological findings

Given the close link between NAFLD and the metabolic syndrome, it is unsurprising that the lipid components of the metabolic syndrome were higher in the NAFLD group compared to the control group, along with levels of fasting glucose, insulin and HOMA-IR. Liver enzymes were also higher in the NAFLD group. Liver enzymes have long been utilised as a biomarker for NAFLD, with data suggesting that as many as 88.4% of people with a persistently elevated ALT have NAFLD[202]. This result was therefore expected. Although raised ALT is linked to NAFLD, analysis of the impact of increased ALT on NAFLD risk within this cohort is not useful, as there is also significant evidence that NAFLD can exist in the context of a normal ALT[203-205]. The increased GGT may reflect the increased numbers of participants drinking alcohol within the NAFLD group. Platelets would be expected to drop in NAFLD only with advanced cirrhosis and portal hypertension and as such were not useful when examining their relationship to NAFLD risk.

Rates of significant liver disease as assessed by TE were 15.5% in the NAFLD group with 4.5% having evidence of advanced chronic liver disease. These rates are roughly comparable to UK data (Chalmers et al. 19.1% significant, 4.3% advanced[206]), but there is a paucity of Indian data with which to compare. Das et al. used FibroScan® in their initial cohort but only assessed patients with "significant NAFLD", as defined by presence of steatosis with an accompanying ALT rise[142], contrary to the fact that significant liver fibrosis exists in the context of a normal ALT[203, 204]. Two years later, the same group did some work looking at median liver stiffness within a small NAFLD cohort and a selection of healthy controls, but without a detailed breakdown of this work it is impossible to compare with this cohort[195]. Many other Indian studies have looked at degree of fibrosis in NAFLD using biopsy, but these are biased toward urban, tertiary centres and as such are likely to overestimate prevalence of advanced disease.

2.5.2 Strengths and weaknesses of the study

To the author's knowledge, this is the largest population-based study of NAFLD ®prevalence in an Indian population in the literature. The biggest strength of this work is the random, population-level sampling framework that was utilised. The spread across the different demographics of the district gives an accurate estimate of the population prevalence of NAFLD in Trivandrum, filling an important gap in the knowledge base. The cohort has also been deeply phenotyped, which enables accurate analysis of a wide range of factors and their impact on NAFLD risk. This population is relatively static, enabling future longitudinal and interventional studies.

The main limitation of this work from a clinical perspective was the use of ultrasound as the diagnostic tool for NAFLD. Ultrasound has been shown to have a poor yield in the diagnosis of microvesicular fat within the liver, with an overall sensitivity of 43% and specificity of 73%[37]. This means that the NAFLD prevalence estimated may even be an underestimate. Whilst liver biopsy has been used to define NAFLD and is hence generally considered the gold standard for NAFLD diagnosis, its invasive nature and inherent risk of sampling error makes it unsuitable for population-based studies. Alternative diagnostic modalities, including MRS, have now be proven to give accurate estimates of NAFLD[138], but its cost and availability meant its use was not possible for this study.

In addition to this, 8% of TE readings were not entirely reliable given their wide IQR. Evidence suggests that most discordance between TE reading and true fibrosis stage (as measured via histology) are attributable to an underestimation of degree of fibrosis[207]. Therefore, the amount of significant liver disease within this cohort is likely to be an underestimate, rather than an overestimate.

Finally, although representative of a large Indian population, the data collected from this cohort are not necessarily representative of Indian as a whole, whose diverse population includes thousands of distinct and unique cultures.

2.5.3 Study implications and future work

The finding that NAFLD prevalence within this cohort is significantly higher than both national and global estimates is an important one. Obesity plays a significant role in the NAFLD risk profile of this study population, as well as in those of other studies across India. As such, those limited resources available should focus on obesity prevention strategies and NAFLD screening in those areas that have the highest rates of obesity, which have been shown to include the southern states and Kerala[201]. FibroScan®, with controlled attenuated parameter scoring (as a measure of hepatic steatosis), may be a useful screening tool to use on a population-level. Its portable nature and ability to give instant results may prove invaluable for the identification and stratification of disease in the future. The validity of this tool would need to be confirmed in an Indian population.

Whilst examining the impact of demographic, anthropometric and pathological factors, further work needs to be done to examine the interaction between lifestyle

and genetics factors within this population and their influence on NAFLD risk. The following chapter will examine the impact of diet and physical activity on NAFLD within this cohort, and future work will include a genome-wide association study using the DNA samples collected from this cohort.

2.6 Conclusion

NAFLD prevalence in this South Indian population was 49.8%, which is higher than other national and global estimates. This is likely to be due to high rates of diabetes and obesity within the population (higher than rates documented in academic literature and government reports), both of which are shown to be significant risk factors for NAFLD within this group. This confirms both hypotheses outlined at the beginning of this chapter.

To obtain a better understanding of the NAFLD risk profile and potential reasons for this higher disease prevalence, it is necessary to understand the impact of lifestyle factors on NAFLD risk. The following chapter will analyse food frequency and physical activity data for this cohort.

CHAPTER 3: Lifestyle factors and NAFLD risk in India

3.1 Introduction

3.1.1. Lifestyle and NAFLD

Changes in dietary composition and exercise habits over the last century may be contributing to the rising burden of obesity and diabetes. Evidence shows that the adverse cardiovascular risk profile connected to the metabolic syndrome is associated with a diet high in carbohydrate, low in unsaturated fat and with low levels of physical activity[174]. As outlined in Chapter 1, a diet rich in carbohydrate in persons with excess adipose tissue increases hepatic fat deposition and insulin resistance and therefore has been shown to also increase risk of NAFLD[72]. Further studies have demonstrated the link between dietary habits and NAFLD (Table 1-2). For example, Musso et al found that the diet of patients with NASH was higher in saturated fat and cholesterol [59]. Zelber-Sagi et al. confirmed these findings in their study of nutrition and the risk of NAFLD in an Israeli group [77]. They found that a diet with a higher intake of meat and soft drinks was associated with NAFLD, independent of total calorie intake. Although, a meta-analysis of Fructose consumption and risk of NAFLD concluded that the association seemed to be confounded by total calorie intake [208]).

Decreasing levels of physical activity and increasing sedentary behaviour have also been shown to increase risk of NAFLD, even when adjusted for insulin resistance, diet and BMI[89, 91]. In one study, it was found that engaging in resistance exercise was inversely associated with NAFLD even after adjusting for insulin resistance, nutritional factors and BMI. However, the statistical significance of this disappeared when waist circumference was added to the model, suggesting this association may be mediated by reduced rate of abdominal obesity [89]. More recent prospective studies that examined changing lifestyle habits over 20 years have shown that the excess NAFLD risk conferred by obesity is attenuated by increased physical activity; at any BMI, increased physical activity reduced risk of liver-related mortality[209].

3.1.2 Indian lifestyle and NAFLD

Whilst there is a wealth of evolving data on the effects of diet and exercise in relation to NAFLD in general, little is known about the specific impact of lifestyle on NAFLD risk in people of Indian origin. This is of particular relevance given the increased rates of disease in this population compared to other ethnic groups, as presented in Chapters 1 and 2. Presented here are the only available data related to lifestyle and NAFLD within Indian populations from the two dedicated studies published in the literature.

Sathiaraj et al examined risk factors between NAFLD cases and controls from 200 sequential attendees to routine health check appointments in Hyderabad, India

(urban)[176]. Interestingly, within this apparently healthy cohort, the NAFLD prevalence was 49%, similar to the Trivandrum NAFLD cohort (Chapter 2.). They looked at differences in nutritional intake between cases and controls and showed that total calorie, percent carbohydrate and percent fat were higher in the NAFLD group, with only percent fat remaining an independent risk factor for NAFLD on logistic regression. The average total calorie intake was 8100±1205kJ for NAFLD and 6890±1255kJ for controls respectively, which is likely to represent a degree of under-reporting, as current recommended calorie intake is 10,500kJ per day for an average male[210]. They also collected data on physical activity including amount and type of exercise undertaken, although the precise method of this data collection was not clear. Only 46% of the study population were involved in physical activity and there were no differences in amount or type between groups.

Singh et al published the only other study that included lifestyle data in relation to NAFLD within an Indian population[177]. They examined NAFLD risk factors comparing 464 subsequent patients diagnosed with NAFLD to 181 controls. This included assessment of dietary intake which was split into particular dietary habits (e.g. vegetarian diet vs. non-vegetarian diet, rice-predominant vs. grain-predominant), intake of particular food groups including fruits, fast food and aerated drinks, and methods of cooking, including consumption of fried food and level of spice. They found that there were no differences in rice/grain predominance between groups, but that those with NAFLD were more likely to consume fried or spicy food or follow a non-vegetarian diet – the only factor that remained a risk factor after adjustment for confounders. Again, low levels of physical activity were reported, with only 33% of NAFLD and 32% of controls undertaking any regular exercise, with no difference between groups.

3.2 Hypotheses, aims and objectives

This chapter will address the following hypothesis using the subsequent aims and objectives:

- Dietary composition influences NAFLD risk within India – with increased consumption of fat and carbohydrate increasing NAFLD risk

AIM 2: To identify the impact of different NAFLD risk factors within this population

Objective E) Clean and analyse dietary data from the Trivandrum NAFLD cohort, creating a database which outlines the average macronutrient intake as grams per day and percentage of total calorie intake

Objective F) Summarise the dietary intake of the Trivandrum NAFLD cohort with regard to 12 basic food groups

Objective G) Summarise the physical activity of the Trivandrum NAFLD cohort as metabolic equivalent hours per week

Objective H) Perform case-control analysis of NAFLD risk with regard to dietary intake and physical activity within the Trivandrum NAFLD cohort, identifying which lifestyle factors are independently associated with NAFLD risk

3.3 Methods

3.3.1 Study design

The Trivandrum NAFLD cohort is outlined in detail in the previous chapter. Briefly, the cohort was created in 2013 through un-weighted, multi-stage, randomised cluster sampling of all inhabitants of the Trivandrum district over the age of 25 years, totalling 2222 participants. Those who did not undergo a liver ultrasound or who had an alternative cause than NAFLD for an echo-bright liver were excluded. This resulted in a nested NAFLD cohort of 2089 participants: n=1,040 NAFLD; n=1,049 controls (disease prevalence 49.8%).

3.3.2 Ethical approval

Ethical approval for the study was granted by the Sree Gokulam Medical College and Research Foundation, Venjaramoodu, Trivandrum ethics committee. This study and all relevant documentation received approval from the University of Nottingham Faculty of Medicine and Health Sciences Research Ethics Committee (REC: 26/299/05/2017, 14/06/17) and the Nottingham University Hospitals NHS Trust Research and Development department. As participants were not NHS patients, NHS ethical approval was not needed and was thus not sought.

3.3.3 Consent

All participants provided written informed consent at the time of recruitment, having had the opportunity to discuss the study in full with those taking consent (particularly if the participant was illiterate) and to receive written information about the study. Participants had the opportunity to ask questions prior to signing consent and before any data was collected.

3.3.4 Data collection

The lifestyle data that form the basis of this chapter were collected during the development of the Trivandrum NAFLD cohort in 2013. Participants were approached and consented to the cohort through house-to-house survey by social workers within their wards. These data were entered onto paper CRFs and were then transcribed onto an electronic access database (Microsoft®, Redmond, USA) using their unique identifier number as outlined in Chapter 2 (2.3.5). This meant that lifestyle data were collected prior to identifying whether individuals were NAFLD or controls, reducing the risk of intake bias.

Dietary data collection

Dietary data were collected using a region-specific FFQ which was created locally through collection of detailed recipes of frequently consumed foods/drinks and has been validated against repeated food diaries[211]. The FFQ was completed through direct questioning from a social worker. It consists of a list of 389 items (see Appendix 1), some of which are complex recipes. Participants recorded how often they consumed each item over the previous year based on the following categories:

- Daily
- Weekly
- Monthly
- Occasional/seasonal
- Never

Each item had a standard portion size attached to it and participants were also asked to report the size of the portion eaten, as well as the frequency as above.

Dietary data analyses

The nutrient composition of each item was calculated by entering the weight and volume of one portion of its raw ingredients into a nutrition programme (Dietsoft®, Noida, India) which utilises the Nutritive Value of Indian Foods database[212].

To estimate the dietary and macronutrient intake of each individual (carbohydrate, protein, fat), the reported frequency of each item was multiplied by the reported portion size and its respective nutrient composition. From this, the mean daily macronutrient intakes (grams per day, g/day) were calculated.

Example:

a) Participant "A" consumed dosai (item 22) once a week, with a portion size of 3 (3x60g).
   Portion size per day is calculated = portion size x frequency per day
   0.42 = 3 x 0.14

b) One portion of dosai contains:
   Fat 0.72g
   Carbohydrate 25.70g
   Protein 4.10g
   Energy 125.70kcal

c) For participant "A", their consumption of dosai (item 22) contributed; 1.72g per day of protein (4.10x0.42), 0.30g per day of fat (0.72x0.42) and 10.79g of carbohydrate (25.70x0.14)

d) The same calculations were done for all 389 items, and were combined to give the average macronutrient intake per day for each participant

Macronutrient data were also presented as proportion (%) of total kilocalories (kcal) eaten as each macronutrient. This percentage kcal intake for each macronutrient was calculated through energy adjustment i.e. 1g fat = 9kcal, 1g protein = 4kcal, 1g carbohydrate = 4kcal. This enabled detailed analysis of the dietary composition in line with government recommendations – utilising percentage macronutrient intake rather than grams per day, which requires adjustment for body weight[210].

Example continued:

e) The average total calorie intake per day (across all 389 FFQ items) of participant "A" is 1941.42kcal made up of:

Fat 54.58g/day x 9	=	491.22kcal/day	= 25.30% of total kcal
Carbohydrate 315.53g/day x 4	=	1262.12kcal/day	= 65.01% of total kcal
Protein 47.02g/day x 4	=	188.08kcal/day	= 9.69% of total kcal

In addition to dietary data being presented in terms of macronutrient intake (g/day and %total calories), data were also presented in terms of the basic food groups, which include:

1. Milk and milk products
2. Fruits
3. Vegetables
4. Sugars
5. Nuts and oil seeds
6. Condiments and spices
7. Meat, fish and poultry
8. Cereals and millets
9. Fats and edible oils
10. Pulses and legumes
11. Beverages
12. Bakery and snacks

Each item/recipe was broken down into its constituent ingredients, which were allocated to one of the above food groups. This enabled calculation of average consumption in grams per day of each food group using the same calculation outlined in the example above.

<u>Physical activity data collection</u>

Physical activity data were collected using a modified global physical activity questionnaire (GPAQ – Appendix 2). The GPAQ was completed through direct questioning from a social worker. It consists of three main domains: occupational, non-occupational and transportation physical activity. For the occupational and non-occupational domains, participants were asked what proportion (giving the number of whole hours) were spent doing little (sedentary), moderate or vigorous activity for every 8 hours in a day (8am-4pm). For the transportation domain, they were asked how much time was spent travelling by vehicle, by walking or cycling over 24 hours.

Physical activity data analyses

These data were used to calculate the average MET for each participant across the three domains using the following calculation: sedentary 1 MET; moderate exercise 4 METs; vigorous exercise 8 METs; self-transportation 4 METS to give total MET hours per week per person (METhr/week). These data were presented as hours per week, as the GPAQ was designed to collect data as proportion of 8 hours within a day, across an average week.

3.3.5 Data cleaning

As described in Chapter 2 (2.3.5) data were converted into STATA format (version 15, Statacorp, Texas, USA) utilising individual unique identifiers to link this lifestyle data to clinical data for analysis. Data were examined for outliers and those data that were implausible were deleted and treated as missing. Repeated reporting of items (frequency and portion size of each item should only be included once) by an individual participant were deleted, leaving the highest frequency response. Those participants that reported consumption of less than ten different items from the FFQ were excluded.

3.3.6 Data analysis

Baseline data from the cohort were presented as follows: Categorical data were presented as numbers (percentage) and continuous data were presented as mean (SD). The reported variables within this work would usually be Normally distributed within the populations sampled and central limit theorem suggests that when testing for differences between groups (NAFLD cases and controls), these differences are also expected to be Normally distributed[198]. In addition, it can be assumed that the data had homogeneity of variance and that data points were independent and were derived from an equal-interval scale. For these reasons, all continuous data were considered parametric. Comparison was made between the lifestyle cohort (n=2,047) and the NAFLD cohort (n=2,089) to ensure the residual sample (due to missing lifestyle data) were representative of the cohort overall.

When comparing continuous data between two groups (i.e. between lifestyle cohort and NAFLD cohort and between NAFLD and control groups), unpaired, two-tailed Students t test of equal variance was used for continuous data, and Chi squared and two-sample test of proportion test for categorical data. Statistical significance was set at $p<0.05$.

Unadjusted odds ratios were calculated through univariate logistic regression (including 95% confidence interval and p value) to identify which lifestyle variables were associated with NAFLD risk.

To understand the confounding effect of different variables on NAFLD risk, an adjusted logistic regression model was created. An initial model was created through the forward, stepwise addition of variables. This included variables identified a priori from the previous Chapter (e.g. gender, BMI, central obesity, diabetes and dyslipidaemia), as well as variables that showed a univariate difference between

NAFLD and control groups (e.g. proportion of macronutrient intake as fat, saturated fat and fibre intake (g/day) and METhr/week). For each sequential variable added, likelihood testing was applied to compare explanatory power(Pseudo-R^2) of the subsequent iteration to its previous iteration. The p value for the likelihood test demonstrated whether the subsequent iteration resulted in a significant change in likelihood ratio from the previous iteration (i.e. the variable should be kept in). Those variables that did not improve the model were removed.

3.4 Results

3.4.1 Baseline characteristics

Clinical data

Dietary and physical activity data were available for 2,047 (lifestyle cohort) of the original NAFLD cohort (n=2,089) representing 2% data-loss (27 participants gave <10 responses, 15 participants FFQ data were missing). When comparing the characteristics of this lifestyle sub-cohort to the original NAFLD cohort, there were no differences in age (p=0.956), gender (p=0.839), domicile (p=0.992), socioeconomic class (p=0.957) or body composition (BMI p=0.881).

These data were therefore representative of the cohort overall. The full baseline demographic characteristics of the lifestyle cohort are presented in Table 3-1.

	NAFLD n=1021	Control n=1026
Age (years, mean SD)	48.31 (10.77)	46.03 (12.37)
Gender (n, %)		
- Male	468 (45.84)	334 (32.55)
- Female	553 (54.16)	692 (67.45)
Domicile (n, %)		
- Rural	412 (40.35)	529 (51.56)
- Urban	609 (59.65)	497 (48.44)
Higher education (n, %)	199 (19.49)	162 (15.79)
BMI (kg/m^2, mean SD)	26.85 (4.13)	24.30 (4.07)
BMI category (n, %)		
- Underweight (<18kg/m^2)	4 (0.39)	66 (6.43)
- Normal (18.5-22.9kg/m^2)	162 (15.87)	329 (32.07)
- Overweight (23-24.9kg/m^2)	193 (18.90)	218 (21.25)
- Obese (>25kg/m^2)	661 (64.74)	413 (40.25)
Central obesity (n, %)	553 (54.16)	410 (39.96)
Metabolic syndrome (n, %)	530 (51.91)	295 (28.78)
- Diabetes	687 (67.29)	480 (46.78)
- Hypertension	412 (40.35)	305 (29.73)
- Dyslipidaemia	713 (69.83)	551 (53.70)

Table 3-1. Lifestyle cohort baseline characteristics

Abbreviations: BMI = body mass index
Higher education = graduate or above; central obesity = ≥80cm in women and ≥90cm in men; metabolic syndrome = central obesity and more than one of: triglycerides ≥1.7, HDL ≤1.07, lipid lowering therapy, fasting glucose ≤5.6, diabetes, hypertension; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy; hypertension = history of hypertension, antihypertensives, systolic blood pressure >130, diastolic blood pressure >85.

As presented in Chapter 2, those with NAFLD were older, were more likely to be male, from an urban domicile and of higher socioeconomic class. Again, NAFLD cases had a higher BMI and were more likely to have components of the metabolic syndrome.

Lifestyle data

The average portion per day of each item (as listed in the FFQ) across the whole cohort is presented in Appendix 1. The 10 items that were consumed most commonly are presented in Table 3-2 and included tea, coffee, rice, dosai (rice batter "pancake"), iddli (fermented rice and black lentil cake) and dried fried fish (often mackerel with chilli and turmeric).

FFQ item	Average portion per day
Boiled rice	4.6
Tea with milk and sugar	2.6
Tea with milk	2.3
Raw rice	2.0
Tea with sugar	1.8
Dried fried fish	1.6
Black tea	1.5
Black coffee	1.4
Dosai	0.9
Iddli	0.9

Table 3-2. Ten most commonly consumed FFQ items across the lifestyle cohort

Abbreviations: FFQ = food frequency questionnaire

The average daily intake (g/day) of each of the 12 food groups across the cohort is outlined in Figure 3-1. This demonstrates that a large proportion of diet in Trivandrum was consumed as cereals and millets (including rice), which concurs with the portion size of the items in Table 3-2

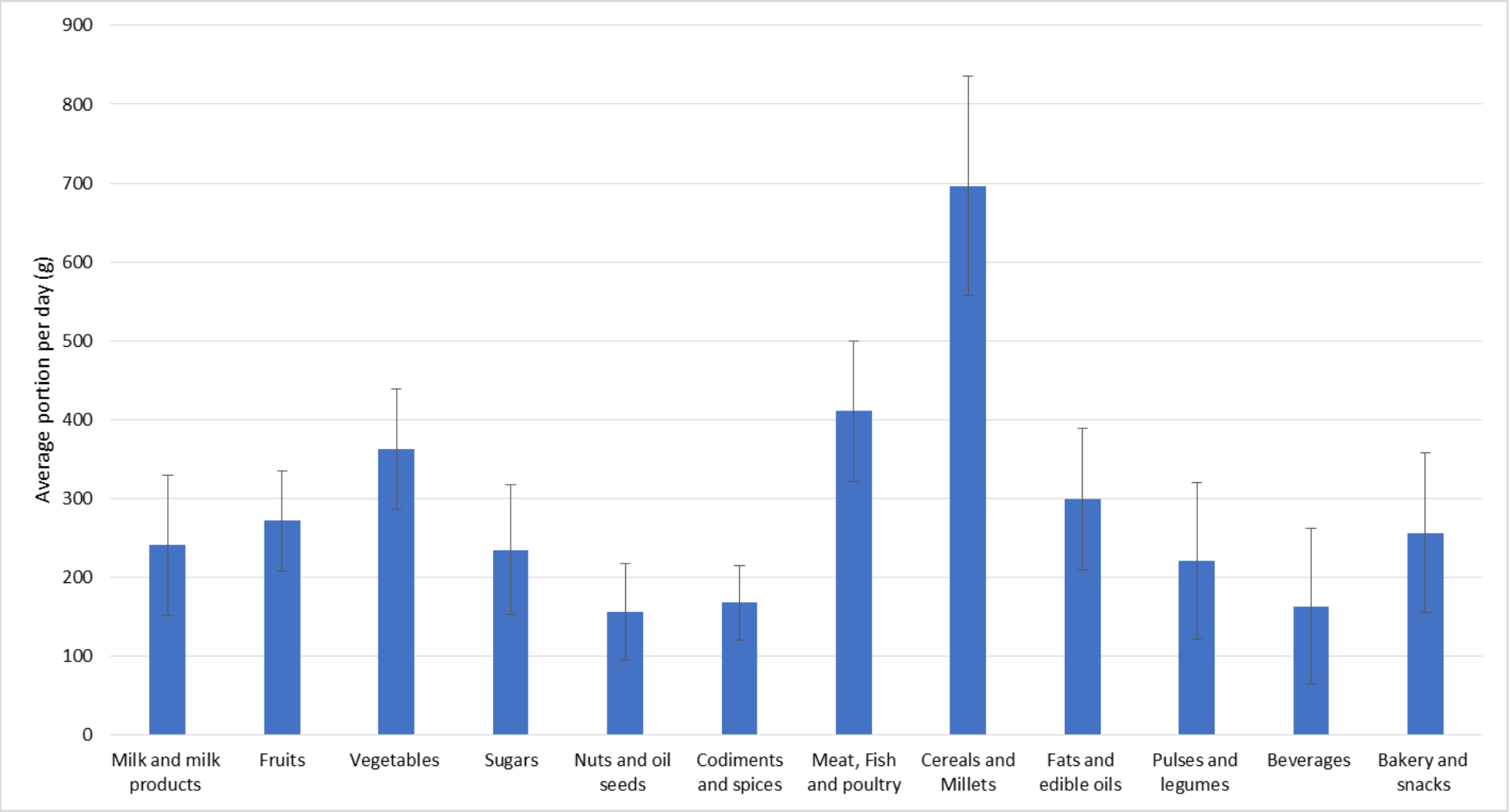


Figure 3-1. Mean intake per day of each food group across the lifestyle cohort (error bars indicating standard deviation)

To enable comparison of the intake of the 12 food groups presented in Figure 3-1 between NAFLD cases and controls, their average daily consumption (g/day) are presented in Table 3-3.

Food group (g/day)	NAFLD n=1021 mean (SD)	Control n=1026 mean (SD)	P value
Milk and milk products	121.23 (81.34)	112.43 (75.34)	0.284
Fruits	128.92 (82.41)	141.58 (89.16)	0.081
Vegetables	165.23 (94.82)	182.65 (91.78)	0.056
Sugars	119.23 (73.65)	109.89 (72.21)	0.344
Nuts and seeds	75.23 (41.23)	73.23 (43.82)	0.645
Condiments and spices	79.43 (48.14)	81.23 (46.76)	0.387
Meat, fish and poultry	222.23 (97.37)	162.23 (82.86)	0.022
Cereals and millets	480.43 (196.64)	210.23 (98.34)	0.003
Fats and edible oils	167.27 (78.32)	63.23 (31.98)	0.001
Pulses and legumes	121.23 (87.56)	90.23 (90.23)	0.082
Beverages	89.23 (41.91)	70.23 (37.81)	0.324
Bakery and snacks	134.78 (91.89)	108.23 (48.76)	0.062

Table 3-3. Average food group intake (g/day) in NAFLD and control groups (rice included within “cereals and millets”)

Those with NAFLD consumed on average more animal protein, more cereals and millets and more fats and edible oils than controls. Fruit was the only food group that NAFLD cases appeared to consume less of than controls, although not statistically so.

The average dietary intake of each macronutrient as grams per day, proportion of total energy intake (kcal/day) and proportion fibre consumed as insoluble fibre is compared between NAFLD and controls in Table 3-4.

	NAFLD n=1021 mean (SD)	**Control n=1026 mean (SD)**	**P value**
Total calories (kcal/day)	2775.35 (850.46)	2769.63 (789.80)	0.875
Carbohydrate (g/day)	373.70 (120.55)	379.83 (109.53)	0.228
% total kcal as Carbohydrate	53.59 (7.75)	54.65 (7.91)	0.002
Protein (g/day)	91.07 (37.55)	90.97 (37.18)	0.951
% total kcal as Protein	12.78 (2.49)	12.76 (2.49)	0.840
Fat (g/day)	101.81 (37.82)	98.49 (36.38)	0.043
- Saturated fat	58.20 (23.10)	55.58 (21.59)	0.008
- Monounsaturated fat	10.71 (4.44)	10.51 (4.28)	0.299
- Polyunsaturated fat	5.61 (2.47)	5.49 (2.36)	0.239
- Cholesterol	130.01 (83.12)	131.09 (81.55)	0.766
% total kcal as Fat	32.53 (6.81)	31.51 (6.96)	<0.001
Fibre (g/day)	48.86 (17.74)	47.09 (15.39)	0.016
- Soluble (g/day)	9.42 (3.55)	9.05 (2.95)	0.010
- Insoluble (g/day)	39.44 (14.30)	38.05 (12.51)	0.019
% total fibre as insoluble	80.67 (1.24)	80.71 (1.21)	0.433

Table 3-4. Average daily macronutrient intake in NAFLD and control groups

There was no difference in total calorie intake between groups. There was also no difference in intake of protein. However, those with NAFLD did consumed more fat per day (both in terms of grams per day and percentage of total calorie intake). Through further, more detailed analysis of the components of dietary fat, it appears that this difference is largely accounted for by an increased intake of saturated fat within the NAFLD group. Those with NAFLD ate more fibre of both soluble and insoluble components with no difference in proportion of fibre sub-types consumed. There was no difference in total carbohydrate intake per day, but those with NAFLD appeared to eat a smaller proportion of calories as carbohydrate, with the increased proportion of calories consumed as fat.

The physical activity data is compared between NAFLD and control groups in Table 3-5 (note n=2,045 as physical activity data were missing for two participants).

	NAFLD n=1020 mean (SD)	Control n=1025 mean (SD)	P value
Total MET hr/week	56.28 (13.73)	54.99 (13.05)	0.028
Sedentary (hr/week)	8.08 (2.4)	8.32 (2.35)	0.023
- % of total METhr/week	50.51 (15.26)	52.00 (14.71)	
Moderate (hr/week)	18.36 (6.30)	18.15 (6.00)	0.447
- % of total METhr/week	28.69 (9.84)	28.37 (9.35)	
Strenuous (hr/week)	26.64 (15.11)	25.14 (14.15)	0.020
- % of total METhr/week	20.81 (11.80)	19.64 (11.06)	

Table 3-5. Average physical activity in NAFLD and control groups

Abbreviations: MET = Metabolic Equivalent

Interestingly, those with NAFLD appeared to do more physical activity than controls, with an increased amount of strenuous exercise and less sedentary time.

3.4.2 Univariate analysis

In Chapter 2, multivariate analysis was performed to identify the baseline characteristics that independently influenced NAFLD risk within this cohort (section 2.4.1, Table 2-3). This showed that male sex, BMI, central obesity, diabetes and hyperlipidaemia were independently associated with increased NAFLD risk. In addition, the data presented above (Table 3-5, Table 3-6) identified lifestyle factors that appeared to influence NAFLD risk, including percentage of intake as fat and carbohydrate, saturated fat and fibre intake (g/day), amount of sedentary and strenuous activity (METhr/week) and total METhr/week. As such, all of these variables were be analysed in isolation for their impact on NAFLD risk, through univariate analysis (Table 3-6). This enabled identification of those variables that should be included in a multivariate model.

	Unadjusted OR (95% confidence interval)	P value
Male gender	1.75 (1.46, 2.09)	<0.001
BMI (kg/m^2)	1.17 (1.14, 1.20)	<0.001
Central obesity	1.78 (1.49, 2.12)	<0.001
Diabetes	2.27 (1.85, 2.78)	<0.001
Dyslipidaemia	1.99 (1.66, 2.38)	<0.001
Fat intake (% total kcal)	1.02 (1.01, 1.03)	0.001
Saturated fat intake (g/day)	1.01 (1.00, 1.01)	0.009
Carbohydrate intake (% total kcal)	0.98 (0.97, 0.99)	0.003
Fibre intake (g/day)	1.01 (1.00, 1.01)	0.017
MET hr/week (total)	1.01(1.00, 1.01)	0.028
Strenuous (MET hr/wk)	1.01 (1.00, 1.01)	0.021
Sedentary (MET hr/wk)	0.96 (0.92, 0.99)	0.024

Table 3-6. Unadjusted NAFLD odds ratios within the lifestyle cohort

Abbreviations: BMI = Body mass index, MET = Metabolic equivalent
Central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy

Considering lifestyle factors, a 1% increase in proportion of total calorie intake as fat is associated with an increased NAFLD risk of 2%. The reverse is seen with proportion of carbohydrate intake; the proportion of carbohydrate decreases as fat intake increases. For this reason, both variables cannot be included in a multivariate model due to collinearity, so fat intake (% of total) was selected. In addition, a 1g/day increase in fibre intake appears to increase NAFLD risk by 1%.

An increase of 1 MET hour per week is associated with an increased NAFLD risk of 1%, which appears to be related to increased strenuous activity and decreased sedentary time in those with NAFLD within the cohort. Again, collinearity of increased time undertaking strenuous activity and decreased time being sedentary means that total METhr/week alone was included in the multivariate model.

3.4.3 Logistic regression

In order to understand the confounding effect of both demographic and lifestyle variables on NAFLD risk, an adjusted logistic regression model was created. This included the clinical variables as outlined above, as well as fat intake (% of total), saturated fat intake (g/day), fibre intake (g/day) and total METhr/week.

The development of the final model is outlined through the tables below showing the sequential addition of variables and the resultant likelihood ratios, R^2 values and a p value. This p value demonstrates whether the subsequent iteration results in a significant change in likelihood ratio from the previous iteration:

	Adjusted OR (95% CI) n=2045
Constant	1.00 (0.91, 1.09)

Table 3-7. Iteration 0 (R^2=0.00, log likelihood=-1417.48)

	Adjusted OR (95% CI) n=2045
Male sex	1.75 (1.46, 2.09)

Table 3-8. Iteration 1, addition of male sex to the model (R^2=0.01, log likelihood=-1398.72, p<0001)

	Adjusted OR (95% CI) n=2045
Male sex	2.20 (1.82, 2.67)
BMI (kg/m^2)	1.19 (1.16, 1.22)

Table 3-9. Iteration 2, addition of BMI to the model (R^2=0.09, log likelihood=-1287.64, p<0.001)

Abbreviations: BMI = Body mass index

	Adjusted OR (95% CI) n=2045
Male sex	3.09 (2.41, 3.95)
BMI (kg/m^2)	1.15 (1.12, 1.18)
Central obesity	1.88 (1.43, 2.46)

Table 3-10. Iteration 3, addition of central obesity to the model (R^2=0.10, log likelihood=-1277.10, p<0.001)

Abbreviations: BMI = Body mass index
Central obesity = ≥80cm in women and ≥90cm in men

	Adjusted OR (95% CI) n=2045
Male sex	2.88 (2.25, 3.71)
BMI (kg/m^2)	1.26 (1.11, 1.43)
Central obesity	1.75 (1.33, 2.31)
Diabetes	1.98 (0.68, 5.65)

Table 3-11. Iteration 4, addition of diabetes to the model (R^2=0.11, log likelihood=-1257.32, p<0.001)

Abbreviations: BMI = Body mass index
Central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl

	Adjusted OR (95% CI) n=2045
Male sex	2.92 (2.27, 3.76)
BMI (kg/m^2)	1.14 (1.11, 1.18)
Central obesity	1.73 (1.31, 2.28)
Diabetes	1.91 (1.54, 2.38)
Dyslipidaemia	1.69 (1.39, 2.06)

Table 3-12. Iteration 5, addition of dyslipidaemia to the model (R^2=0.12, log likelihood=-1243.48, p<0.001)

Abbreviations: BMI = Body mass index
Central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy

	Adjusted OR (95% CI) n=2045
Male sex	2.92 (2.27, 3.77)
BMI (kg/m^2)	1.14 (1.11, 1.18)
Central obesity	1.72 (1.30, 2.27)
Diabetes	1.90 (1.53, 2.36)
Dyslipidaemia	1.69 (1.38, 2.05)
Fat intake (% calories)	1.02 (1.00, 1.03)

Table 3-13. Iteration 6, addition of fat intake to the model (R^2=0.12, log likelihood=-1241.01, p<0.001)

Abbreviations: BMI = Body mass index
Central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy

	Adjusted OR (95% CI) n=2045
Male sex	2.98 (2.31, 3.84)
BMI (kg/m^2)	1.15 (1.11, 1.18)
Central obesity	1.72 (1.30, 2.27)
Diabetes	1.87 (1.50, 2.33)
Dyslipidaemia	1.70 (1.39, 2.07)
Fat intake (% calories)	1.02 (1.01, 1.04)
Saturated fat intake (g/day)	1.00 (0.99, 1.00)

Table 3-14. Iteration 7, addition of saturated fat intake to the model (R^2=0.13, log likelihood=-1239.94, p=0.144)

Abbreviations: BMI = Body mass index
Central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy

	Adjusted OR (95% CI) n=2045
Male sex	2.96 (2.30, 3.83)
BMI (kg/m^2)	1.15 (1.11, 1.18)
Central obesity	1.72 (1.30, 2.27)
Diabetes	1.89 (1.52, 2.35)
Dyslipidaemia	1.69 (1.39, 2.06)
Fat intake (% calories)	1.02 (1.00, 1.03)
Fibre intake (g/day)	1.00 (0.99, 1.00)

Table 3-15. Iteration 8, removal of saturated fat intake and addition of fibre intake to the model (R^2=0.12, log likelihood=-1240.58, p<0.001)

Abbreviations: BMI = Body mass index
Central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy

	Adjusted OR (95% CI) n=2045	P value
Male sex	2.93 (2.27, 3.78)	<0.001
BMI (kg/m^2)	1.14 (1.11, 1.18)	<0.001
Central obesity	1.74 (1.32, 2.30)	<0.001
Diabetes	1.89 (1.52, 2.35)	<0.001
Dyslipidaemia	1.70 (1.39, 2.06)	<0.001
Fat intake (% calories)	1.02 (1.00, 1.03)	0.019
Fibre intake (g/day)	1.00 (0.99, 1.00)	0.479
MET hours/week	1.01 (1.00, 1.01	0.157

Table 3-16. Iteration 9, addition of MET hr/week to the model (R^2=0.13, log likelihood=-1239.48, p<0.001)

Abbreviations: BMI = Body mass index, MET = Metabolic equivalent
Central obesity = ≥80cm in women and ≥90cm in men; diabetes = history of diabetes &/or fasting glucose>126mg/dl; dyslipidaemia = triglycerides>150mg/dl &/or high-density lipoproteins<50mg/dl female, high-density lipoproteins<40mg/dl male &/or lipid lowering therapy

This final model demonstrated that increased proportion of macronutrient intake as fat was the only lifestyle factor that appeared to independently influence NAFLD risk

within this cohort. In addition to those risk factors identified a priori (Chapter 2), a 1% increase in proportion of dietary calories consumed as fat (e.g. 25kcal/2.8g of fat per day from an average intake of 2,500kcal for a man) appeared to increase NAFLD risk by 2%. Although addition of fibre intake and METhr/week improved the fit of the model (and were therefore included), they were not independently associated with NAFLD risk (p=0.479, p=0.157).

3.5 Discussion

3.5.1 Principal findings

Within this large, population based cohort in Trivandrum, India, rice was the most commonly consumed food (average 4.6 portions per day), with cereals and millets being the most commonly consumed food group (average 697g/day). Those with NAFLD consumed more animal protein, cereals and millets and fats and edible oils than controls (p=0.022, p=0.003, p=0.001). When comparing the macronutrient intake between NAFLD cases and controls, there was no difference in total calorie intake, but the proportion of energy intake as fat was higher in the NAFLD group (32.53% vs 31.51%, p<0.001), with a corresponding lower proportion of carbohydrate intake (53.59% vs 54.65%, p=0.002). There was an increased consumption of fat overall in the NAFLD group (g/day) with an increased intake of saturated fats in particular (58.2g/day vs. 55.6g/day, p=0.008). Saturated fats are consumed most commonly from animal sources (both meat and dairy), and within this cohort, it is likely therefore that the increased saturated fat intake in the NAFLD group comes from an increased consumption of animal protein and/or edible oils. The NAFLD group also consumed more fibre of both subtypes (48.86g/day vs 47.09g/day, p=0.016). Interestingly, the NAFLD group were more active than the control group (56.28 vs 54.99 METhr/week p=0.028), spending less time sedentary (8.08 vs 8.32 hours/week, p=0.023).

Through multivariate modelling, it was shown that an average daily increase of 1% in the proportion of diet consumed as fat is associated with an increased risk of NAFLD of 2%, independently of other variables (adjusted OR 1.02, 1.00-1.03, p=0.019). None of the other lifestyle factors appeared to influence NAFLD risk when adjusted for confounders. Of note, are the low log likelihood and pseudo-R^2 within the multivariate models. The adjusted R^2 (pseudo-R^2) denotes the percentage of variation explained by the variables within the model i.e. a perfect model would predict 100% of the variation (log likelihood = zero, R^2 = 1.0). In the final multivariate model presented within this chapter, the pseudo-R^2 was 0.13, despite including variables that are known to strongly influence NAFLD risk (BMI, diabetes). This highlights the inherent randomness within real-world data, particularly given the large sample size and variation within this population in question. Cox and Wermuth noted in their 1992 paper that “with binary responses, R^2 tends to be low, even for

an underlying perfect regression relationship"[213], and whilst small, an R^2 value of 0.13 can still be considered significantly different to zero.

As demonstrated through the work of Chapters 1 and 2, people of Indian ethnicity are at increased risk of NAFLD due to their propensity to insulin resistance and predisposition to central adiposity, which, at a lower BMI compared to other ethnicities, has been shown to further compound metabolic risk. The data presented in this chapter has also demonstrated that lifestyle factors may play a role in the Indian NAFLD phenotype.

This work supports the findings of Sathiaraj et al, who examined nutritional risk factors for NAFLD within a population with a similar disease prevalence[176]. They showed that increased percentage of fat intake resulted in a much more dramatic rise in NAFLD risk than this study found (OR 2.51, 1.99-3.31). Despite there being much larger differences in nutrient intake between their cases and controls, much like this study, fat intake remained the only nutritional variable that independently influenced disease risk.

Parallels can also be drawn between this work and that of Singh et al, who demonstrated that those with NAFLD were more likely to consume fried food and follow a non-vegetarian diet[177]. Whilst they did not collect detailed macronutrient data, the increased consumption of fried food and animal protein are mirrored in the findings within this work.

3.5.2 Strengths and weaknesses of the study

Study cohort

The Trivandrum NAFLD cohort, from which this data is collected, is, to the author's knowledge, the largest population-based study of NAFLD prevalence and associated risk factors within an Indian population. The biggest strength of this work is the use of random, population-level sampling across the breadth of geographic and social regions of this district, resulting in the data collected forming an accurate representation of the population as a whole. The combination of demographic, anthropometric and lifestyle data mean the cohort is deeply characterised. This enables accurate assessment of overall NAFLD phenotype, allowing adjustment for multiple possible confounders when examining specific variables – in this case, lifestyle. It is also a large study. The two other studies of nutrition and NAFLD risk within India are significantly smaller. The larger study comprised 642 participants who were sampled from consecutive patients attending hospital appointments within a large East Indian city[177] - resulting in a bias towards the urban, higher socioeconomic class and resulting in a sex disparity (M:F 3:1). The other study was smaller still (n=200), recruiting its participants from those attending routine employment check-ups[176]. Again, this resulted in significant selection bias towards those of a socioeconomic class who have employment that included health checks, and towards men (76%) who hold these types of jobs. The results of both of these

studies are therefore limited in their accuracy and generalisability even within the populations in which they were conducted.

Dietary data

Another strength of this study comes from the utilisation of a region-specific FFQ. This FFQ was designed and validated within the population in question using repeated food diaries, making it the ideal tool to capture detailed information about their precise dietary habits. In this piece of work, it has been utilised to look at the macronutrient composition of diet in relation to NAFLD risk, but it could also be used to examine the interaction between different dietary patterns and risk of disease.

Sathiaraj et al reported using an FFQ within their study, although they do not specify which they used, and appeared to only include foods that were eaten daily, weekly or monthly[176]. They do however use the same reference for the calculation of the macronutrient composition of food as used in this work[212]. Singh et al used a different approach, combining categorical questions (e.g. predominant grain type consumed) with questions about frequency of intake of certain foods, which was limited to daily/alternate days or weekly/twice weekly[177]. The method of dietary data collection within their study is very vague, and although the results may be more useful for interventional strategies (i.e. reduction in fried food consumption), the lack of detail and breadth of data collected may have resulted in important findings being missed.

The use of FFQs have been well validated in large, epidemiological studies[214], but they do have their limitations. FFQ data are vulnerable to systemic reporting/recall bias, comprised of both intake-related and person-specific biases[215]. Intake bias occurs when study groups feel pressure to follow a certain dietary pattern. In this case, participants were aware that they are taking part in a study related to fatty liver disease, so may have tended to under-report unhealthy dietary practices. Person-specific bias refers to the difference between a persons reported average intake, and the true average intake. This may be affected by socio-cultural influence or certain personality characteristics, and as such are difficult to quantify. If one were to compare this work to other studies that utilise FFQs, it is reasonable to think that all data would be biased to a similar degree. Despite the predisposition of FFQs to attenuate results between cases and controls through systemic bias, this study has conclusively demonstrated a difference in fat intake between cases and controls, an effect which therefore may even be under-reported. Alternatives to FFQ for dietary data collection include 24-hour dietary recall and dietary records. 24-hour dietary recall is time consuming, requiring interviewers to document the recalled dietary intake of the participant over the past 24 hours, over a number of days. Whilst it may be more detailed, it is subject to significant recall and interviewer bias, and was not appropriate for use in this study. Dietary records require participants themselves to document their dietary intake over a number of days. Whilst this generates detailed data, it has a large respondent burden, requiring both literacy and motivation from participants. Again, this was not appropriate in this setting.

Physical activity data

Throughout the literature, there is a lot of conflicting evidence about the impact of physical activity on NAFLD risk. Within Indian populations, there is a particular lack of large, detailed data related to exercise in relation to disease in general. The PURE study compared physical activity data from high-income countries (where physical activity is often recreational) to low-income countries including India (where activity is more likely to be occupational)[216]. Similar to this work, an international physical activity questionnaire (IPAQ) was used and data presented in terms of MET as the measure of activity. It was found that, regardless of setting, exercise improves cardiovascular morbidity and mortality across different socioeconomic strata. The same could be assumed for NAFLD. Although GPAQ is broadly considered a suitable tool to use for large, epidemiological studies, its limitation, in this work in particular, relates to the way in which data on the different forms of exercise was collected. Participants were asked to report how many hours were spent active within an 8hr working day, and how many hours were spent active in an 8hr period outside the working day; the total number of hours spent active within 16 hours of a day (which is why this data is presented as MET hours, rather than minutes). This is likely to have resulted in overestimation of the amount of exercise done, as culturally women are less likely to work than men but did not report zero hours of occupation-related activity. The wording of the questions implied an assumption that every day contained both occupational and non-occupational activity. Again, this is likely to result in inclusion of more occupational activity than is actually done for any given week. This general overestimation is highlighted when comparing these results with those from the PURE study, where people from low-income countries did on average 2520 METmin/week, whereas the cohort in this work did on average 3338 METmin/week (56 METhr/week). Regardless of the inability to compare these findings with other works, reporting errors in this study are the same across the population, meaning comparison between cases and controls was still relevant.

Finally, whilst this large quantity, high quality data has given insights into the impact of diet and exercise on NAFLD risk within Trivandrum, this may not be entirely representative of the whole of India, which is made up of many distinct and unique cultures.

3.5.3 Study implications and future work

The work of this chapter demonstrates the impact of lifestyle on risk of NAFLD within a population already shown to have high rates of disease compared to current global estimates. In particular, consumption of a diet composed of higher proportion of fat is independently associated with increased risk of NAFLD. This is an important finding and has public health implications within India, which could inform local education strategies and future government policy on a larger scale.

Whilst this work adds to the evidence base that dietary fat influences NAFLD risk (particularly in an Indian population), attempts at interventional strategies that involve alterations in dietary macronutrient composition are not pragmatic at a

population level. Future work that utilises these data are likely to involve analysis of dietary patterns within the population. This will enable further understanding of the major dietary constituents in terms of types of fats and edible oils, and sources of protein consumed and their association with NAFLD risk. Findings from this and planned further work are likely to inform future dietary intervention studies within this cohort (ClinicalTrials.gov NCT03844165).

3.6 Conclusion

As a result of economic changes over the last 30 years, there have been significant alterations in dietary patterns across India. Where previously childhood malnutrition and unsafe water were a major health burden, diet-related non-communicable diseases including diabetes are now the leading cause of death within India[154]. This is due to a shift away from consumption of coarse grains towards consumption of rice and wheat[217], and more pertinently, there has also been an increase in fat intake across all socioeconomic groups, often in the form of highly saturated ghee[218]. The work of this chapter has demonstrated that this increase in dietary fat may be associated with an increased risk of NAFLD within India. This confirms, in part, the hypothesis outlined at the beginning of this chapter. This finding may help inform future observational and interventional studies focussing on the components of the Indian diet and their relationship with metabolic risk, further strengthening the need for education and policy change in the future.

To understand whether a change in lifestyle results in altered NAFLD risk profile, the following chapter will examine whether acculturation through migration to the UK results in a change in NAFLD risk profile in people of South Asian ethnicity.

CHAPTER 4: The impact of socio-cultural factors on NAFLD risk in UK Indians

4.1 Introduction

Through industrialisation and urbanisation of Asia, rates of obesity, diabetes and the metabolic syndrome have dramatically risen[219] – matching and overtaking those of Western countries[7, 220]. This highlights the important role of changing lifestyle habits on NAFLD risk in a population who already possess an innate risk of disease (as outlined in Chapter 1). Westernisation of Indian culture has resulted in increasing consumption of fat across the country[218], which, as demonstrated in the previous chapter, is likely to be a contributory factor in the increased NAFLD risk in India. However, less is known about the impact of migration into Western countries and the resultant change in socio-cultural habits on the NAFLD risk profile in people of Indian ethnicity.

4.1.1 Acculturation

Acculturation is a sociological process by which a person assimilates to a different culture, typically the dominant culture[221]. As Indian migrants in the UK become more accultured, exposure to new lifestyles, diets, beliefs and environments may change their risk of developing chronic disease, including NAFLD. Understanding this accultural effect may provide further insight into the impact of different factors that influence NAFLD risk and may help inform the development of culturally appropriate interventions for disease prevention and management. In 2005 Patel et al undertook a study looking at the impact of migration on coronary heart disease risk, comparing risk profiles of Gujarati Indians living in Britain, who have higher rates of hypertension and hyperlipidaemia, to their contemporaries in villages of origin in India[222]. They demonstrated that British Gujaratis had a BMI that was 6kg/m^2 (4.5-7.4kg/m^2) higher than their native counterparts. They had a greater total calorie intake and proportion of intake as fat despite an increased energy expenditure (almost 40% more Kcal per day) and with little difference in glucose tolerance. This implied that the adverse effect of migration on cardiovascular risk is nutritional in nature. Studies looking at dietary acculturation suggest that it is the “accessory foods”, such as sugary drinks and fat-rich snacks, which are the first to be exchanged[223], and that the adoption of a Western diet as opposed to maintaining an Eastern or mixed diet strongly correlates with presence of the metabolic syndrome[224].

Acculturation may not however be a linear process, and speed of acculturation may be influenced by numerous factors related both to place of origin (urban vs rural; socioeconomic class) and place of migration as well as duration of residence and generation within the UK[225]. In 2003, Harding examined the mortality of South Asian migrants by duration of residence in the UK and found that a yearly increase in

residence was associated with an all-cause mortality HR of 1.07 (1.02-1.13) and cardiovascular disease mortality HR of 1.09 (1.03-1.16)[226]. Other studies did not however confirm this finding, showing no association between language preference, generation or duration of residence and risk of NAFLD[227].

Based on this, people of Indian ethnicity who migrate to the UK are likely to adopt British lifestyle habits. This may include increased physical activity, increased total calorie intake and increased intake of fat and sugar. Although innately more insulin resistant, this diet may potentiate insulin resistance and result in altered body composition, (increased BMI and decreased waist-hip ratio) compared to their native counterparts, thus increasing the risk of metabolic disease and NAFLD through dietary acculturation.

4.1.2 NAFLD risk in UK-South Asians

There are only a few studies published in the literature relating to NAFLD risk in migrant South Asian groups.

Alazawi et al. performed a large, retrospective, population-based study looking at the rates of abnormal LFTs and coded NAFLD diagnoses from an ethnically diverse region of London[228]. They analysed primary care data from over 690,000 adults and found that rates of NAFLD diagnoses were highest in people of Bangladeshi origin (1.8%), and that Bangladeshi origin was itself an independent risk factor for NAFLD. They compared risk factors between cases and controls and demonstrated that increasing BMI was the strongest risk factor for NAFLD across the whole study population (adjusted OR 1.63 overweight, OR 2.44 obese). However, this study does not include data from native Bangladeshi populations, making assessment of the impact of acculturation difficult.

Neukam et al also conducted a study looking at rates of NAFLD within a UK South Asian cohort[229]. They analysed data from a cardiovascular screening programme from members of two London Hindu temples, largely formed of UK Gujarati migrants. This retrospective, cross-sectional study calculated NAFLD prevalence within this UK South Asian population using the fatty liver index (based on BMI, GGT, triglycerides and waist circumference). Within a sample size of 597 UK Gujarati Indians, the NAFLD prevalence was 31%, demonstrating higher rates of NAFLD compared to the global pooled average[1]. This study did not however include any comparator groups, making ethnic/migrant comparisons impossible.

The only other work in the literature that directly compared a migrant South Asian population with their native counterparts was undertaken by Patel et al[222]. This group compared the cardiovascular risk profile of UK Gujarati Indians with their siblings living in India. Whilst not looking specifically at people with NAFLD, many of the parameters measured within this study form part of the general NAFLD risk profile, and as such is a useful comparator paper to the work of this chapter. Their study of 513 participants across both countries contained a large quantity of high

quality demographic, anthropometric and lifestyle data including use of repeated food diaries and accelerometers to enable analysis of energy intake vs expenditure.

It was shown that the UK-migrant participants had a higher BMI and waist-hip ratio, as well as increased rates of diabetes and obesity. Following migration to the UK, South Asians had an increased calorie intake and increased intake as fat. Again, lack of a control group made evaluation of NAFLD risk within this study impossible.

The aim of this chapter is therefore to add to the above evidence by comparing a UK-migrant South Asian cohort with a native Indian cohort, using a UK-Caucasian cohort as a comparator, examining whether acculturation may result in differences in NAFLD risk profile.

4.2 Hypotheses, aims and objectives

This chapter will address the following hypothesis, using the subsequent aims and objectives:

- Migration of South Asians to the UK results in alteration of NAFLD risk profile through dietary acculturation

AIM 3: To compare NAFLD risk profile of native Indians with their UK migrant counterparts

Objective I) Design and create a multi-ethnic study comprised of native Indian, UK-migrant Indians and Caucasians with and without NAFLD, collecting data regarding NAFLD risk profile

Objective J) Compare NAFLD phenotype between native Indian, UK-migrant Indian and Caucasian cohorts

Objective K) Perform case-control analysis of NAFLD risk across the multi-ethnic study

Objective L) Analyse the impact of ethnicity on NAFLD risk within the multi-ethnic study

4.3 Methods

4.3.1 Study design

This observational, cross-sectional study was designed to enable collection and comparison of data on both clinical and behavioural risk factors associated with NAFLD from three distinct self-identified ethnic cohorts, each comprised of NAFLD and control arms.

a) UK South Asian
b) Native Indian (a sub-group of the Trivandrum NAFLD cohort – Chapter 2)
c) UK Caucasian

For this chapter, it should be noted that the modern definitions of South Asia include Afghanistan, India, Pakistan, Bangladesh, Sri Lanka, Nepal, Bhutan and Maldives.

4.3.2 Recruitment

Indian cohort

Participants of the Indian cohort of this study were recruited from the Trivandrum NAFLD cohort (Chapter 2). As recruitment and data collection for the original cohort was completed in 2016, risk profiles and lifestyle habits of its participants may have changed. For this reason, those participants randomly selected for this study underwent the same study regime as the UK cohorts (including re-collection of anthropometric, pathological and lifestyle data), enabling accurate comparison at the same time point. As the Trivandrum NAFLD cohort participants were managed locally, their site team recruited participants for this study.

UK cohort

Originally, the intention was to recruit all participants of the UK cohorts of this study (UK South Asian and UK-Caucasian) from a previous study which identified patients with NAFLD within the community[206]. Participants of this previous study who had consented to be contacted about future research were allocated a number, and through use of a random number generator were invited to participate. The aim was to recruit a random sample of the local population, whilst ensuring an adequate number of participants with NAFLD were identified. During recruitment, the NAFLD sub-groups became over-subscribed (as NAFLD was only identified on the study day), and there was difficulty in recruiting participants to the South Asian cohort. Therefore, a more pragmatic approach was adopted. Additional participants were recruited through distribution of promotional material and Participant Information Sheets within the community, and by identifying potential participants more likely to be controls based on age, medical history and/or FibroScan® result.

4.3.3 NAFLD risk factor data

To compare NAFLD risk profiles between ethnic cohorts, data was collected from the following domains:

Body composition

There are strong links between BMI, the metabolic syndrome and NAFLD. This is particularly evident in South Asian populations where every increase in BMI of 1kg/m^2 results in an increased risk of diabetes and hypertension of 15%, compared to 11% in Caucasians[158]. Evidence also shows that rates of NAFLD are higher with increasing BMI even within the “normal” range (18.5-22.9kg/m^2) in Indian populations[142]. In addition to BMI, increased visceral adiposity has also been linked with hepatic steatosis and increased inflammation/fibrosis[56] and evidence

shows that Indian populations tend to central adiposity with increased visceral fat[167, 168].

Therefore, in this study, BMI, waist-to-hip ratio, presence of obesity and total body fat percentage (via bioimpedance) were selected as appropriate measures of body composition.

Insulin resistance

The concept of lipid-induced insulin resistance is key to the pathogenesis of NAFLD, where excess energy intake saturates cutaneous and visceral adipose tissue lipid stores, leading to spill-over of free fatty acids which are deposited in skeletal muscle (worsening whole-body insulin resistance) and the liver – an early step in the development of NAFLD[50]. Whilst glycaemic clamping remains the gold standard for measuring insulin resistance, its cost and complexity limits its use in observational studies like this. An alternative technique that is often used in both clinical practice and epidemiological studies is HOMA-IR[230]. HOMA-IR shows significant correlation with glycaemic clamps in non-diabetic patients and has been used in a number of studies relating to NAFLD[231, 232].

For the purposes of this study, HOMA-IR was calculated from fasting glucose (mmol/l) and fasting insulin (mU/l) levels in non-diabetic patients ((fasting insulin*fasting glucose)/405), whilst presence of diabetes (formal diagnosis, treatment or fasting blood glucose >7.0mmol/l) was dichotomous.

Dietary data

There are numerous methods of dietary data collection including single or multiple 24-hour dietary recalls, weighed diet records, self-reports of diet history and food frequency questionnaires (FFQ). In terms of habitual diet, all of these methods are subject to measurement errors from random misreporting or systemic reporting bias, thus limiting their accuracy[233]. FFQs have however long been used over other dietary methods in epidemiological studies. They are easy to deliver and have a lower responder burden when capturing long-term habitual intakes, rather than collecting diary data, which, whilst accurate, only include data on foods eaten over the last few days.

The EPIC FFQ, which has been validated against biomarkers and repeated 24-hour dietary recalls[234], was utilised for collection of dietary data for the UK groups in this study (i.e. for participants who are either UK South Asian or UK-Caucasian). As outlined in Chapter 3, a locally devised and validated FFQ was used to collect dietary data from the native Indian cohort[211].

Physical activity data

Physical activity or inactivity can be measured in a number of different ways. Direct methods, such as use of accelerometers or multi-sensor array, offer optimal accuracy for a given time frame but may not give a representative result when looking at physical activity over longer periods or habitual levels, plus they require use of instruments which may reduce their practicality in clinical settings[235]. Indirect

methods include logbooks and questionnaires. Logbooks may provide a detailed assessment of activity over a relatively short period of time, but have a high administrative burden. Self-reported questionnaires are used more commonly, are easy and quick to deliver and give a global view of an individual's activity but are subject to recall bias[236]. Different physical activity questionnaires report different outputs, and are therefore chosen depending on the activity domain of interest (work/leisure), time within each of these domains, or physical activity as continuous data expressed in metabolic equivalents (MET) of energy expenditure. This method assigns MET values to different levels of exercise e.g. moderate (4 METs) or vigorous (8 METs).

This study utilised the Global Physical Activity questionnaire[237] for the collection of physical activity data within the UK cohorts (UK South Asian and Caucasian), which has been well validated across different populations[238]. It provides MET data that could be compared to the physical activity data collected using the local questionnaire for the Indian cohort (Chapter 3).

4.3.4 Sample size calculation

When designing this study, a sample size calculation was undertaken to ensure ability to identify statistically significant differences for these variables between groups. As there is a paucity of data on NAFLD risk in relation to acculturation, sample size was calculated using a number of different studies that compared risk profiles between cases and controls within Indian populations[176], compared baseline profiles between Caucasians and Indians in America[173] and compared cardiovascular risk between UK migrant Indians and their native counterparts[222].

To be able to detect a difference of 250kcal in daily total calorie intake between groups (10% of the recommended daily total calorie intake of 2500kcal for men[239]), using standard assumptions of 80% power and a significance level of 5% and taking data on diet from a case-control study on risk factors for NAFLD in India[176], a difference between means calculation gave a sample size of 21 per group. Similarly, to detect a 6g difference in daily total fat (6% of the daily recommended intake of 97g of fat), a sample size of 18 participants per group was required. Finally, to detect a difference of 50g in daily total carbohydrate intake (15% of the daily recommended intake of 333g, a sample size of 22 participants per group was needed.

Under the same conditions (80% power and significance level of 5%), a sample size of 14 participants per group would enable detection of a 1.35 mass units difference in HOMA-IR, a difference that was statistically significant when comparing insulin resistance between Caucasians and Indians[173], and in a case control study of native Indians[177]

Powering the study to look for differences in physical activity was more difficult, as there is conflicting evidence in the literature about whether reduced levels of activity are independently associated with NAFLD. Sathiaraj et al [176] did not find a

difference between NAFLD/control groups, neither did a comparable study done by Singh et al[177]. A study done looking at levels of physical activity in UK South Asians and Caucasians did however show a difference between ethnicities. Using this data, a sample size of 22 would enable detection with 80% power a 3METhr/day difference between ethnicities – the difference detected when comparing Asian and Caucasian groups of similar sample size in the United States[137].

Based on these calculations, each ethnic cohort was comprised of 20 NAFLD and 20 controls (Figure 4-1).

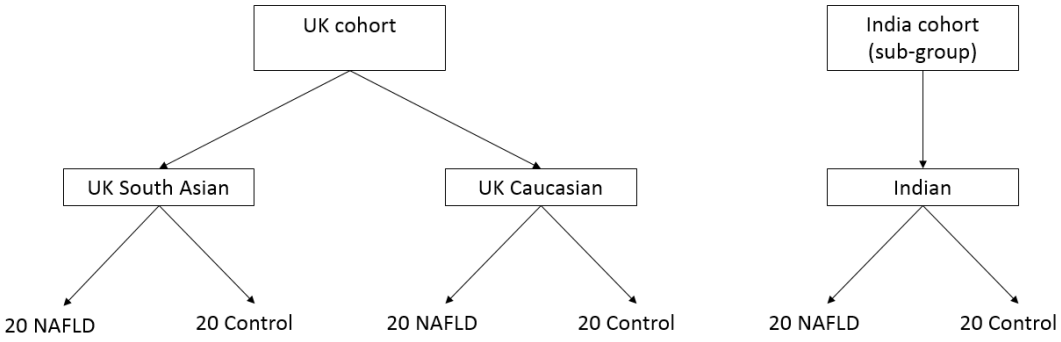


Figure 4-1. Ethnic cohort study design

4.3.5 Eligibility criteria

Inclusion:

- Adult participants aged 18 years or over
- Male
- Able to give informed consent
- UK cohorts: Self-identify as either “White” or “South Asian”
- For NAFLD groups: Fatty/echo-bright liver on ultrasound

This study also forms the basis of work in Chapter 5, which examines BAT activity in relation to ethnicity and NAFLD risk. Due to a number of gender-specific confounders (menstrual cycle, age) which were likely to affect results from small study groups, recruitment was limited to men.

Due to the conflicting evidence on the impact of duration of residence of migrant populations on acculturation, no restriction was imposed on length of time in the UK.

Exclusion:

- Female
- Known or suspected cirrhosis on clinical/histological/radiological grounds
- Current or recent significant history of self-reported alcohol consumption (>14units per week)
- Other documented causes of chronic liver disease including:
    - Hepatitis B or C infection

    - Drug-induced liver disease
    - Alcohol-related liver disease
    - Autoimmune hepatitis
    - Wilson's disease or Haemochromatosis
    - Primary biliary cholangitis or primary sclerosing cholangitis
- Currently taking medication that can induce steatosis (corticosteroids, Amiodarone, Tamoxifen, Methotrexate)
- Evidence of any other unstable or untreated clinically significant immunological, endocrine, haematological, gastrointestinal, neurological, neoplastic or psychiatric disease
- Presence of hair within the region of interest (neck) that cannot be removed prior to the study visit (this relates to the work in Chapter 5)

4.3.6 Ethical approval

Indian cohort

This study and all relevant local documentation received approval from the Sree Gokulam Medical College Institutional Ethics Committee (04/36/01/2013, 18/05/17), Trivandrum, India.

UK cohorts

This study and all relevant documentation received approval from the University of Nottingham, Faculty of Medicine and Health Sciences Research Ethics Committee (REC: 26/299/05/2017, 14/06/17) and the Nottingham University Hospitals NHS Trust Research and Development department. It also received Health Research Authority approval for the site (28/06/17: 17/HRA/3356). Amendments to the protocol and participant information sheet to include recruitment from the community were subsequently approved through the local REC (18/04/18). As participants were not NHS patients, NHS ethical approval was not needed and was thus not sought.

4.3.7 Consent

All participants provided written informed consent on the day of the study regime, having received a detailed Participant Information Sheet and having had sufficient time to consider participating or not. Participants had the opportunity to ask questions prior to signing consent and before undergoing any investigations related to the study. If needed, translator services were available to assist with discussion of the study as well as during the study visit.

Indian participants received the information sheet and consent form in Malayalam, which had undergone forward and backward translation for verification of information. Local staff were available to answer questions.

4.3.8 Study regime

Participants attended a single morning study visit lasting approximately 3 hours, having fasted from midnight and avoided caffeine and strenuous exercise in the

preceding 24 hours. On arrival, eligibility was checked against the inclusion/exclusion criteria which included a brief medical, alcohol and drug history, after which informed consent was given. They then underwent the following investigations and were offered refreshments on regime completion:

To determine body composition:

i. Height: measured to the nearest 0.5cm using a stadiometer (Marsden, Rotherham, UK) placed against a wall
ii. Weight: measured to the nearest 0.1kg using electronic standing scales (Seca, Hamburg, Germany)
iii. Body mass index: kg/m^2
(BMI cut-offs for: Caucasians; <18.5 underweight, 18.5-25 normal, 25-29.9 overweight, ≥30 obese, Asians; <18.5 underweight, 18.5-23 normal, 23-27.5 overweight, ≥27.5 obese)[240]
iv. Waist circumference: measured to the nearest cm using a non-expanding tape at the midpoint between the lower rib margin and the iliac crest

Central obesity defined as South Asians ≥90cm, Caucasians ≥102cm as per Indian consensus statement and IDF metabolic syndrome definition[194, 241].

v. Hip circumference: measured to the nearest cm using a non-expanding tape at the largest point around the buttocks
vi. Waist-hip ratio: waist(cm)/hip(cm)
vii. Body fat percentage (%) measured using bi-frequency terapolar bioimpedance analyser (Bioelectric Impedance Analyser – Analisi Composizione Corporea; Biotekna S.R.I®, Italy)

To determine baseline liver function (including markers of non-specific liver injury) and metabolic profile, 8ml of whole blood was taken by a healthcare professional trained in venepuncture for:

i. Full blood count
ii. Urea, creatinine and electrolytes
iii. Liver function tests (LFTs - ALT, AST, ALP, Albumin, Bilirubin)
iv. Lipid profile (Cholesterol, HDL, LDL, Triglycerides)
v. HOMA-IR (Fasting Glucose and Insulin)

To collect lifestyle data:

i. Average dietary intake of macronutrients over the past year: Assessed through completion of an EPIC FFQ [234], or using a locally validated Indian FFQ[211] presented as intake of macronutrients as grams per day (g/day)
ii. Average physical activity as MET-minutes per week as measured by a GPAQ[237], or using the local Indian physical activity questionnaire.

To identify presence of NAFLD through ultrasound of the abdomen, imaging was undertaken by a qualified radiographer/radiologist within the radiology department

of the Queens Medical Centre, Nottingham, or within the Sree Gokulam Medical College, Trivandrum. Presence or absence of NAFLD was binary and was assessed using the hepatic/renal echo-intensity ratio (>0.5) [242]. Due to logistical constraints, a specific ultrasonographer was not used at each site, and the ultrasound scans were not crosschecked.

4.3.9 Data collection and management

Study participants were assigned a participant study code (UK = SA-XXX, India = IN-XXX) for use on case report forms, other study documents and electronic database. A separate confidential record of participants' name, date of birth, local hospital number, address and participant study number was created to enable identification of all participants enrolled into the study. This confidential record was password protected and held on an NHS server for a year after the study ended.

Participants were informed of the local General Data Protection Regulation policies within the Participant Information Sheet and agreed to these by signing the consent form. Data collected during the study regime were documented on paper CRFs using individual participant study code as the identifier. These paper records were stored in secured filing cabinets within a locked office on NHS premises to which only authorised members of staff have access – all of whom have a duty of confidentiality to the participants. Data from paper CRFs were directly inputted into a master spreadsheet and were held and backed up on a University server requiring both electronic permission to access the storage drive and password access to the database. Under UK Data Protection laws, the University was the Data Controller (was legally responsible for the data security) and the Chief Investigator of this study was the Data Custodian (managed access to the data). Laboratory data were entered manually using local NHS clinical software. All research data has been and will be kept securely for seven years after which time it will be disposed of securely.

4.3.10 Data analysis

Baseline data from each group (NAFLD and control) within each ethnic cohort were initially presented as follows (Table 4-1): Categorical data were presented as numbers (percentage) and continuous data were presented as mean (SD) for Normally distributed data and median (IQR) for non-normal data. The reported variables within this work would usually be Normally distributed within the populations sampled (except for HOMA-IR which is only measured in non-diabetics, making it positively skewed) and study numbers within groups were too small to make statistical testing for Normality (such as the Shapiro-Wilks test) meaningful. Central limit theorem suggests that when testing for differences between groups (NAFLD cases and controls), these differences are also expected to be Normally distributed[198]. In addition, it can be assumed that the data had homogeneity of variance and that data points were independent and were derived from an equal-interval scale. For these reasons, the majority of continuous data (except HOMA-IR) were considered parametric. The variables BMI and body fat% were checked for

collinearity using Pearson's correlation coefficient, using a correlation coefficient of $r \geq 0.4$ as evidence of a moderate correlation and $r \geq 0.7$ as strong correlation[243].

Comparison was then made between the UK Indian cohort and the Trivandrum NAFLD cohort (from which they were recruited) to ensure adequate sampling, using pertinent variables from anthropometric, medical and lifestyle domains (age, BMI, presence of diabetes, physical activity and calorie intake). As the UK Indian cohort was comprised solely of male participants, the comparison was made between the UK cohort and male members of the Trivandrum NAFLD cohort. Inter-cohort comparisons were then made (comparing ethnic cohorts as a whole). Statistical significance was set at $p<0.05$.

When comparing continuous data between two groups (i.e. between UK-Indian and Trivandrum cohorts and between NAFLD and Control groups), unpaired, two-tailed Students t test of equal variance was used for parametric data, and Mann-Whitney U test for non-parametric data (HOMA-IR). Chi squared and two-sample test of proportion were used for categorical data. For comparison between more than two groups (i.e. between the three ethnic cohorts), analyses are done using ANOVA (with Bonferroni correction where differences were identified) for parametric data or Kruskal-Wallis test for non-parametric data.

Intra-cohort comparisons (comparing NAFLD and control groups within each ethnic cohort) were initially presented as the difference in mean/median for continuous data and difference in proportion for categorical data, along with 95% confidence intervals (95% CI) (Table 4-3). This helped identify which variables were significantly different (both statistically and clinically) between NAFLD and control groups across the ethnic cohorts. Further to this, unadjusted odds ratios for NAFLD (95% CI) were calculated for the selected variables within each ethnic cohort (Table 4-4). NAFLD odds ratio for waist-hip ratio was calculated for an increase in ratio of 0.01 as a log reduction for ease of data handling.

In order to understand the confounding effect of different variables on NAFLD risk, an adjusted logistic regression model was created. To begin with, this was done using the entire dataset (n=130) to give the most power to detect differences between the NAFLD and control groups overall. An initial model was created through the forward, stepwise addition of variables. This included variables identified a priori from this work (Chapter 4) and the literature (e.g. age and diabetes). It also included variables that showed the strongest univariate difference between NAFLD and control groups across the ethnic cohorts of this study (e.g. presence of obesity, diabetes and central obesity) as well as variables that appeared to have a modest impact on NAFLD risk (p value <0.20, e.g. dyslipidaemia). Finally, variables that were of specific interest to this body of work were included to see whether they improved the model (METmin/wk and total calorie intake). For each sequential variable added, likelihood testing was applied to compare the explanatory power (Pseudo-R^2) of the subsequent iteration to its previous iteration. The p value for the likelihood test demonstrated whether the subsequent iteration resulted in a significant change in

likelihood ratio from the previous iteration (i.e. the variable should be kept in). Those variables that did not improve the model were removed.

To understand the impact of ethnicity on the final model, further analyses were undertaken using likelihood ratio testing. Ethnicity could not be included as an independent variable within the final model, as the numbers within each ethnic group were fixed. Therefore, comparator models were created to examine whether there was an additional effect of ethnicity on each of the variables included within the null model (except age, which is not influenced by ethnicity). South Asian and Indian ethnicity were compared to Caucasian ethnicity as the base category.

Statistical analyses were performed using STATA v. 15 (Statacorp, Texas, USA).

4.4 Results

4.4.1 Participant characteristics

A total of 130 participants were recruited to the study (UK Caucasian 44, UK South Asian 45, Indian 41). As presence of NAFLD was identified on the day of the study visit, and because recruitment of participants from both countries were from cohorts of people at risk of NAFLD, the NAFLD groups were oversubscribed. The baseline participant characteristics are presented in Table 4-1.

	UK South Asian		UK Caucasian		Indian	
	NAFLD n=25	**Control n=20**	**NAFLD n=24**	**Control n=20**	**NAFLD n=20**	**Control n=21**
Anthropometry						
Age (years) Mean (SD)	49 (10)	45 (12)	60 (13)	49 (16)	49 (12)	48 (12)
BMI (kg/m^2) Mean (SD)	28.32 (3.51)	25.46 (2.67)	31.57 (4.98)	25.49 (3.72)	26.46 (3.43)	22.20 (3.69)
- Underweight n (%)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	4 (19.05)
- Normal n (%)	1 (4)	5 (25)	9 (37.5)	17 (85)	3 (15)	10 (47.62)
- Overweight n (%)	2 (8)	5 (25)	0 (0)	1 (5)	4 (20)	3 (14.28)
- Obese n (%)	22 (88)	10 (50)	15 (62.5)	2 (10)	13 (65)	4 (19.05)
Waist-hip ratio Mean (SD)	0.98 (0.07)	0.93 (0.10)	1.01 (0.07)	0.89 (0.08)	1.01 (0.03)	0.96 (0.07)
- Central obesity n (%)	21 (84)	15 (75)	24 (100)	11 (55)	18 (90)	11 (52)
Body fat (%) Mean (SD)	39.28 (4.58)	36.25 (6.17)	42.79 (4.21)	34.2 (6.60)	37.4 (5.0)	2.48 (7.51)
Medical history						
Diabetic n (%)	11 (44)	3 (15)	20 (83.33)	6 (30)	6 (30)	5 (23.81)
Hypertension n (%)	7 (28)	4 (20)	16 (66.7)	4 (20)	6 (30)	4 (19.1)
Dyslipidaemia n (%)	19 (76)	12 (57.1)	22 (91.7)	9 (45)	10 (50)	12 (57.1)

(continued)

	UK South Asian		UK Caucasian		Indian	
	NAFLD n=25	**Control n=20**	**NAFLD n=24**	**Control n=20**	**NAFLD n=20**	**Control n=21**
Metabolic syndrome n (%)	12 (48)	6 (30)	21 (87.5)	5 (25)	7 (35)	5 (23.8)
Lifestyle						
Total calories kcal/day Mean (SD)	2015 (904)	1809 (718)	1849 (871)	1786 (550)	2658 (424)	2754 (548)
Protein g/day Mean (SD)	81.9 (37.4)	70.4 (27.4)	88.9 (36.2)	82.6 (24.4)	88.1 (27.2)	81.3 (19.4)
Carbohydrate g/day Mean (SD)	259.1 (119.7)	233.9 (101.3)	209.6 (116.8)	198.3 (68.3)	385.5 (68.5)	411.6 (103.9)
Fat g/day Mean (SD)	78.4 (39.4)	70.5 (31.5)	76.0 (39.4)	71.9 (25.7)	85.3 (24.5)	86.9 (19.6)
MET min/week Mean (SD)	2952 (4114)	3101 (3323)	1840 (2007)	3962 (3744)	2886 (1148)	2693 (815)
LFTs						
ALT U/l Mean (SD)	33.8 (23.7)	23.9 (10.9)	32.8 (20.7)	26.3 (11.6)	52.2 (27.4)	47.0 (34.3)
ALP U/l Mean (SD)	83.2 (22.2)	76.4 (21.6)	88.7 (30.3)	66.3 (22.7)	84 (32.7)	76.8 (30.0)
Platelets						
Platelets 10^9/l Mean (SD)	257.8 (47.0)	276.3 (65.2)	233.2 (55.0)	235.3 (48.5)	191.6 (38.8)	246.5 (48.5)

(continued)

	UK South Asian		UK Caucasian		Indian	
	NAFLD n=25	**Control n=20**	**NAFLD n=24**	**Control n=20**	**NAFLD n=20**	**Control n=21**
Lipids						
Cholesterol mmol/l Mean (SD)	4.6 (1.0)	4.7 (0.7)	4.2 (1.5)	4.7 (1.2)	5.6 (1.2)	5.9 (1.3)
Triglycerides mmol/l Mean (SD)	0.7 (0.5)	0.4 (0.5)	0.9 (0.3)	0.3 (0.5)	0.4 (0.5)	0.5 (0.5)
HDL mmol/l Mean (SD)	0.5 (0.5)	0.3 (0.5)	0.9 (0.3)	0.3 (0.5)	0.4 (0.5)	0.2 (0.4)
LDL mmol/l Mean (SD)	2.6 (1.0)	2.7 (0.6)	2.1 (1.1)	2.6 (1.1)	3.7 (1.0)	3.9 (1.0)
LDL-C mmol/l Mean (SD)	3.0 (0.9)	3.1 (0.7)	2.6 (1.3)	2.9 (1.1)	4.0 (1.1)	4.2 (1.2)
Metabolic						
Glucose mmol/l Mean (SD)	7.7 (4.3)	5.7 (1.6)	9.4 (6.1)	5.9 (1.9)	5.3 (2.1)	5.1 (2.5)
Insulin mU/l Mean (SD)	11.5 (7.8)	7.7 (5.2)	24.7 (39.9)	5.9 (5.0)	25.8 (50.6)	14.1 (15.4)
HOMA-IR Median (IQR)	0.13 (0.07)	0.06 (0.04)	0.16 (0.2)	0.02 (0.07	0.11 (0.09)	0.06 (0.09)

Table 4-1. Participant characteristics across the ethnic cohorts

Abbreviations: BMI = body mass index, MET = Metabolic equivalents, ALT = alanine transferase, ALP = alkaline phosphatase, HDL = high-density lipoproteins, LDL = low-density lipoproteins, LDL-C = low-density lipoprotein cholesterol, HOMA-IR = homeostatic model assessment of insulin resistance

*BMI cut-offs for: Caucasians; <18.5 underweight, 18.5-25 normal, 25-29.9 overweight, ≥30 obese, Asians; <18.5 underweight, 18.5-23 normal, 23-27.5 overweight, ≥27.5 obese, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/dl. HOMA-IR calculated if not on diabetic treatment. Hypertension = diagnosis or treatment. Dyslipidaemia = TG>1.7mg/dl, HDL<1.03mg/dl or on treatment. Metabolic syndrome = central obesity + two of (raised triglycerides, reduced HDL, presence of hypertension, presence of diabetes or glucose ≥5.6mg/dl), HOMA-IR ((fasting insulin*fasting glucose)/405) calculated for those not on treatment for diabetes. LDL-C = (total cholesterol-HDL cholesterol-(triglycerides/5))*

Correlation between BMI and body fat % across the whole study population was strongly positive ($r=0.91$, $p<0.001$) (Figure 4-2) suggesting collinearity of variables, and as such BMI alone was used for further analyses. BMI had a weak but significant positive correlation with waist-hip ratio ($r=0.42$, $p<0.001$) and as such waist-hip ratio was included as a variable in further analyses.

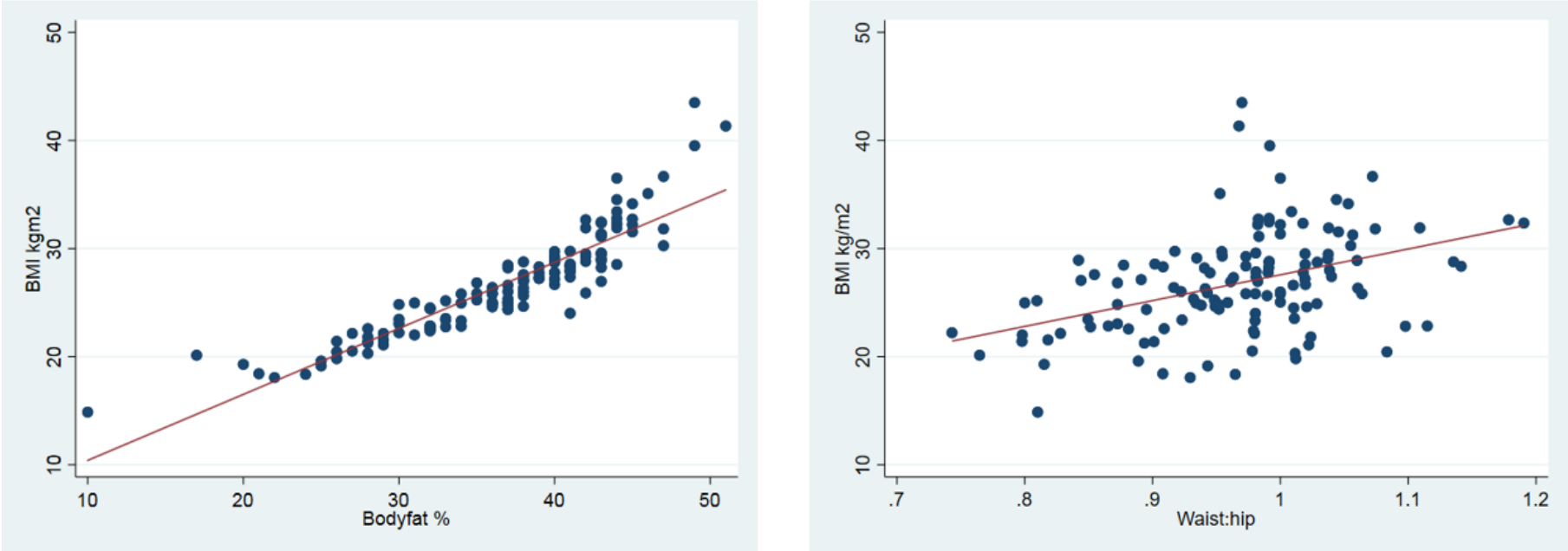


Figure 4-2. Correlation between BMI and a) body fat % (r=0.91, p<0.001) b) waist-hip ratio (r=0.42, p<0.001) across all ethnic cohorts

Abbreviations: BMI = Body mass index

Indian cohort comparison

Prior to analysing differences between the ethnic cohorts, comparison was made between the Indian cohort and male participants of the Trivandrum NAFLD cohort to ensure adequate sampling. These data are presented in Table 4-2.

	Indian cohort **n=41**	**Trivandrum NAFLD cohort** **n=802**	**P value**
Age (years) Mean (SD)	48.3 (11.7)	47.91 (11.97)	0.839
BMI (kgm^2) Mean (SD)	24.28 (4.12)	24.95 (3.83)	0.277
Diabetes n (%)	10 (24.39)	486 (60.60)	<0.001
MET (min/week) Mean (SD)	2787.32 (983.50)	2991.02 (1009.83)	0.208
Total calories (kcal/day) Mean (SD)	2707.21 (487.52)	2901.08 (863.49)	0.154

Table 4-2 Comparison of UK-Indian and Trivandrum NAFLD cohort

Abbreviations: BMI = Body mass index, MET = Metabolic equivalent
Diabetic = diagnosis or fasting glucose>7.0mg/dl

The UK-Indian cohort is representative of the Trivandrum NAFLD cohort in terms of age, BMI, amount of physical activity undertaken and total calories consumed per day. However, there were significantly fewer participants with diabetes in this study cohort, although rates of diabetes within the Trivandrum NAFLD cohort overall (both men and women) were comparable – 26.33%.

Ethnic comparison

Table 4-1 shows that the UK Caucasian cohort was significantly older than both the UK South Asian (p=0.005) and Indian cohorts (p=0.024), who were of similar age (p=0.585). The Indian cohort had a significantly lower BMI and body fat % than both UK Caucasians (p<0.001, p=0.001) and UK South Asians (p=0.001, p=0.002), who had similar body composition. Rates of being overweight or obese (as per ethnic cut-offs) were highest in the UK South Asian cohort (n= 39, 86.7%) compared to Indian (n=24, 58.5%) and UK Caucasian (n=18, 40.9%) cohorts. There was no difference in waist-hip ratio between cohorts (p=0.30).

Rates of diabetes were highest in the UK Caucasian cohort (59.1%), with lower rates in the UK South Asian (31.1%) and Indian (26.8%) cohorts. There was no difference in HOMA-IR between ethnic groups (p=0.15). There were no differences in rates of hypertension or dyslipidaemia between groups (p=0.051, p=0.601 respectively), but

rates of metabolic syndrome overall were highest in the UK Caucasian cohort (59.1%) compared to UK South Asian (40%) and Indian cohorts (29%).

The Indian cohort consumed significantly more total calories per day compared to the two UK cohorts ($p<0.001$), who consumed similar amounts ($p=1.00$). There was no difference in protein or fat intake between cohorts ($p=0.071$, $p=0.150$), but the Indian cohort consumed significantly more carbohydrate per day compared to the two UK cohorts ($p<0.001$), who consumed similar amounts ($p=0.125$). There was no difference in METmin/wk between groups ($p=0.920$).

The Indian cohort had higher levels of total cholesterol, LDL-C and ALT than the UK cohorts ($p<0.001$). The low levels of glucose in the Indian cohort compared to the UK Caucasians is likely to be a result of the low rates of diabetes. There were no other differences in blood parameters that were of clinical note.

Case-control comparison

Next, the differences between NAFLD cases and controls within each cohort were examined and are presented in Table 4-3.

	UK South Asian NAFLD vs Control Difference in mean/median/% (95% CI) n=45		UK Caucasian NAFLD vs Control Difference in mean/median/% (95% CI) n=44		Indian NAFLD vs Control Difference in mean/median/% (95% CI) n=41	
Anthropometry						
Age (years) (Mean)	3.6 (-3.03, 10.23)	p=0.280	10.71 (2.01, 9.41)	p=0.017	0.99 (-6.49, 8.48)	p=0.800
BMI (kg/m^2) (Mean)	2.86 (0.95, 4.78)	p=0.004	6.09 (3.36, 8.80)	p<0.001	4.25 (2.00, 6.50)	p<0.001
Obesity (%)	0.38 (0.13, 0.63)	p=0.005	0.5 (0.29, 0.76)	p<0.001	0.46 (0.19, 0.73)	p=0.003
Waist-hip ratio (Mean)	0.044 (-0.007, 0.095)	p=0.091	0.123 (0.076, 0.170)	p<0.001	0.044 (0.011, 0.078)	p=0.011
Central obesity (%)	0.09 (-0.15, 0.33)	p=0.453	0.45 (0.23, 0.67)	p<0.001	0.38 (0.13, 0.63)	p=0.008
Medical history						
Diabetes (%)	0.29 (0.04, 0.54)	p=0.041	0.53 (0.28, 0.78)	p<0.001	0.06 (-0.21, 0.33)	p=0.650
Hypertension (%)	0.08 (-0.17, 0.33)	p=0.532	0.47 (0.21, 0.72)	p=0.002	0.11 (-0.15, 0.37)	p=0.411
Dyslipidaemia (%)	0.31 (0.04, 0.58)	p=0.033	0.62 (0.39, 0.85)	p<0.001	-0.07 (-0.38, 0.23)	p=0.653
Metabolic syndrome (%)	0.18 (-0.10, 0.46)	p=0.220	0.63 (0.39, 0.86)	p<0.001	0.11 (-0.17, 0.39)	p=0. 434

(continued)

	UK South Asian NAFLD vs Control Difference in mean/median/% (95% CI) n=45		UK Caucasian NAFLD vs Control Difference in mean/median/% (95% CI) n=44		Indian NAFLD vs Control Difference in mean/median/% (95% CI) n=41	
Lifestyle						
Total Calories (kcal) (Mean)	205.15 (-295.16, 05.45)	p=0.413	63.13 (-390.73, 16.99)	p=0.780	-95.44 (-405.93, 215.04)	p=0.538
Protein (g/day) (Mean)	11.50 (8.7, 31.7)	p=0.257	6.28 (-12.94, 25.49)	p=0.513	6.75 (-8.10, 21.59)	p=0.363
Carbohydrate (g/day) (Mean)	25.19 (-42.54, 92.92)	p=0.457	11.25 (-48.54, 71.05)	p=0.706	-26.10 (-81.98, 29.77)	p=0.351
Fat (g/day) (Mean)	7.93 (-13.93, 29.80)	p=0.468	4.14 (-16.55, 24.83)	p=0.688	-1.63 (-15.63, 12.37)	p=0.815
METmin/wk (Mean)	-149 (-2438.8, 2140.8)	p=0.896	-2122 (-335.8, -3908.2)	p=0.021	193 (-433.7, 819.0)	p=0.538
LFTs						
ALT (U/l) (Mean)	9.9 (-1.7, 21.5)	p=0.091	6.5 (-3.9, 17.0)	p=0.215	5.2 (-14.5, 24.9)	p=0.060
ALP (U/l) (Mean)	6.8 (-6.5, 20.1)	p=0.308	22.5 (5.9, 39.0)	p=0.009	7.2 (-12.6, 27.0)	p=0.468
Platelets						
Platelets (10^9/l) (Mean)	-18.5 (-52.2, 15.3)	p=0.280	-2.1 (-34.5, 30.3)	p=0.890	-54.9 (-85.6, 24.2)	p<0.001

(continued)

	UK South Asian NAFLD vs Control Difference in mean/median/% (95% CI) n=45		UK Caucasian NAFLD vs Control Difference in mean/median/% (95% CI) n=44		Indian NAFLD vs Control Difference in mean/median/% (95% CI) n=41	
Lipids						
Cholesterol (mmol/l) (Mean)	-0.1 (-0.6, 0.4)	p=0.719	-0.6 (-1.4, 0.3)	p=0.181	-0.3 (-1.1, 0.5)	p=0.408
Triglycerides (mmol/l) (Mean)	0.3 (-0.01, 0.6)	p=0.063	0.6 (0.3, 0.8)	p=0.001	-0.1 (-0.4, 0.2)	p=0.857
HDL (mmol/l) (Mean)	0.2 (-0.1, 0.5)	p=0.144	0.6 (0.3, 0.8)	p=0.001	0.2 (-0.1, 0.4)	p=0.143
LDL (mmol/l) (Mean)	0.1 (-0.4, 0.6)	p=0.700	-0.6 (-1.2, 0.1)	p=0.140	0.2 (-0.5, 0.9)	p=0.530
LDL-C (mmol/l) (Mean)	0.02 (-0.5, 0.5)	p=0.690	-0.3 (-1.0, 0.4)	p=0.420	-0.2 (-1.0, 0.5)	p=0.580
Metabolic						
Glucose (mmol/l) (Mean)	2.0 (-0.05, 4.0)	p=0.055	3.5 (0.6, 6.4)	p=0.019	0.2 (-1.3, 1.6)	p=0.797
Insulin (mU/l) (Mean)	3.8 (-0.3, 7.9)	p=0.071	18.8 (0.6, 37.0)	p=0.041	11.6 (-11.7, 35.0)	p=0.321
HOMA-IR (Median)	0.07	p=0.073	0.09	p=0.014	0.05	p=0.791

Table 4-3. Intra-cohort differences between NAFLD cases and controls

Abbreviations: BMI = body mass index, MET = Metabolic equivalents, ALT = alanine transferase, ALP = alkaline phosphatase, HDL = high-density lipoproteins, LDL = low-density lipoproteins, LDL-C = low-density lipoprotein cholesterol, HOMA-IR = homeostatic model assessment of insulin resistance

*Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/dl. HOMA-IR calculated if not on diabetic treatment. Hypertension = diagnosis or treatment. Dyslipidaemia = TG>1.7mg/dl, HDL<1.03mg/dl or on treatment. Metabolic syndrome = central obesity + two of (raised triglycerides, reduced HDL, presence of hypertension, presence of diabetes or glucose ≥5.6mg/dl), HOMA-IR ((fasting insulin*fasting glucose)/405) calculated for those not on treatment for diabetes. LDL-C = (total cholesterol-HDL cholesterol-(triglycerides/5))*

NAFLD and control groups were matched for age in the South Asian cohorts (UK South Asian and Indian), but NAFLD cases were significantly older than controls in the Caucasian cohort. Across all ethnic cohorts, BMI and proportion of those who were obese were higher in the NAFLD group, as was central adiposity, except for the UK South Asian group. Of note, the differences between cases and controls were much larger within the Caucasian cohort compared to the South Asian and Indian cohorts. For example, the difference in BMI between cases and controls within the Caucasian cohort was 6kg/m^2 compared to 4kg/m^2 in Indian and 3kg/m^2 in UK South Asian cohorts.

Diabetes rates were higher in the NAFLD groups of both UK-based cohorts (UK South Asian and UK Caucasian), but were similar between groups in the Indian cohort. All components of the metabolic syndrome were more prevalent in the Caucasian NAFLD group compared to controls, with only diabetes and dyslipidaemia being more prevalent in the UK South Asian NAFLD group (vs. controls).

There was no difference in total calorie intake, or intake of individual macronutrients between NAFLD cases and controls in any of the ethnic cohorts. UK Caucasians with NAFLD were significantly less physically active than their control counterparts (p=0.021), although the confidence interval was wide (-335.8,-3908.2). This difference was not seen in either of the South Asian cohorts.

There was minimal difference in blood-based parameters between cases and controls within cohorts, except for marginally increased Triglyceride, HDL and ALP levels (a difference that is statistically but not clinically significant) and raised measures of glucose and insulin sensitivity within the Caucasian NAFLD group. This difference in insulin sensitivity is accounted for by differences in rates of diabetes.

As planned, variables known a priori to impact NAFLD risk, along with those that were significantly different between cases and controls across these ethnic cohorts, as well as those of particular interest to this body of work (e.g. lifestyle factors) were included for calculation of univariate NAFLD odds ratios. This included; age, BMI, presence of obesity (including central obesity as a separate variable), components of the metabolic syndrome, METmin/wk and total calories. Unadjusted odds ratios of NAFLD are presented in Table 4-4.

	UK South Asian n=45 unadjusted OR (95% CI)		UK Caucasian n=44 unadjusted OR (95% CI)		Indian n=41 unadjusted OR (95% CI)	
Age (years)	1.03 (0.98, 1.09)	p=0.274	1.05 (1.01, 1.10)	p=0.023	1.01 (0.96, 1.06)	p=0.784
BMI (kg/m^2)	1.39 (1.08, 1.79)	p=0.012	1.42 (1.15, 1.76)	p=0.001	1.41 (1.12, 1.77)	p=0.003
Obesity	11.00 (0.90, 134.06)	p=0.021	14.17 (1.91, 105.10)	p<0.001	10.83 (1.39, 84.17)	p=0.004
Central obesity	1.75 (0.40, 7.63)	p=0.456	*	*	8.18 (1.50, 44.49)	p=0.015
Diabetes	4.45 (0.94, 21.04)	p=0.045	11.67 (2.07, 65.68)	p<0.001	1.37 (0.34, 5.60)	p=0.655
Hypertension	1.56 (0.37, 6.45)	p=0.536	8.00 (1.64, 39.04)	p=0.002	1.82 (0.41, 8.00)	p=0.417
Dyslipidaemia	3.87 (1.00, 15.03)	p=0.037	25.67 (2.63, 250.90)	p<0.001	0.75 (0.21, 2.61)	p=0.647
Metabolic syndrome	2.15 (0.60, 7.69)	p=0.224	21.00 (2.67, 164.91)	p<0.001	1.72 (0.43, 6.91)	p=0.433
Total calories (kcal/day)	1.00 (0.99, 1.00)	p=0.407	1.00 (0.99, 1.00)	p=0.775	1.00 (0.99, 1.00)	p=0.528
METmin/wk	1.00 (0.99, 1.00)	p=0.893	0.99 (0.98, 0.99)	p=0.012	1.00 (0.99, 1.00)	p=0.529

Table 4-4. Unadjusted odds ratio for NAFLD risk factors within each ethnic cohort

Abbreviations: BMI = Body mass index, MET = Metabolic equivalent Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/dl. Hypertension = diagnosis or treatment. Dyslipidaemia = TG>1.7mg/dl, HDL<1.03mg/dl or on treatment, Metabolic syndrome = central obesity + 2 of (raised triglycerides, reduced HDL, presence of hypertension, presence of diabetes or glucose ≥5.6mg/dl)

*Central obesity predicted NAFLD perfectly in the Caucasian group and as such, these data were omitted.

Age was associated with increased odds of NAFLD within the Caucasian group only. As previously noted, the Indian and UK South Asian cohorts were comprised of participants of similar age across both NAFLD and control groups, which may be why this effect was not seen within these cohorts.

Measures of body composition were associated with increased NAFLD risk across the whole study population, with the largest effect from presence of obesity (unadjusted OR 10.83-11.00). The confidence intervals for these variables were however extremely large. Presence of diabetes also appeared to increase NAFLD risk although not significantly so in the Indian cohort – which may be related to the low rates of diabetes. There was variation in impact of the individual components of the metabolic syndrome on NAFLD risk between cohorts. Again, within the UK Caucasian cohort, presence of the metabolic syndrome was strongly associated with NAFLD (unadjusted OR 21.00, 2.67-164.91, p<0.001). This variable also appeared to be strongly related to NAFLD within the other cohorts, but not significantly so. Total calorie intake did not influence NAFLD risk across the cohorts, and increased physical activity appeared to reduce risk of NAFLD marginally in the Caucasian cohort only.

Overall, the impact of individual variables on NAFLD risk showed larger odds ratios for the UK Caucasian cohort compared to the other cohorts (e.g. presence of the diabetes appeared to increase NAFLD risk 11-fold in Caucasians, but only 4-fold in UK South Asians and not at all in native Indians).

4.4.2 Logistic regression

A logistic regression model was created using the whole dataset initially (n=130), including the variables identified as having an impact on NAFLD risk a priori, from the literature (i.e. age, diabetes) or through calculation of unadjusted odds ratios (obesity, central obesity, dyslipidaemia). Variables of specific interest to this body of work (METmin/wk, total calorie intake) were added subsequently to see whether they improved the final model. The development of the final model is outlined through the tables below which demonstrate the sequential addition (and removal if no improvement) of variables and the resultant likelihood ratios, R^2 values and a p value. This p value demonstrates whether the subsequent iteration results in a significant change in likelihood ratio from the previous iteration:

	Adjusted OR (95% CI) n=130
Constant	1.13 (0.80, 1.60)

Table 4-5. Iteration 0 (R^2=0.00, log likelihood=-89.86)

	Adjusted OR (95% CI) n=130
Age (years)	1.03 (1.00, 1.06)

Table 4-6. Iteration 1, addition of age to the model (R^2=0.03, log likelihood=-87.22, p=0.022)

	Adjusted OR (95% CI) n=130
Age (years)	1.01 (0.98, 1.05)
Obesity	6.91 (3.15, 15.16)

Table 4-7. Iteration 2, addition of obesity to the model (R^2=0.17, log likelihood=-74.29, p<0.001)

Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5

	Adjusted OR (95% CI) n=130
Age (years)	1.02 (0.99, 1.05)
Obesity	5.10 (2.21, 11.76)
Central obesity	2.89 (0.96, 8.68)

Table 4-8. Iteration 3, addition of central obesity to the model (R^2=0.19, log likelihood=-72.42, p=0.053)

Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort.

	Adjusted OR (95% CI) n=130	P value
Age (years)	0.99 (0.96, 1.04)	0.811
Obesity	4.50 (1.92, 10.56)	0.001
Central obesity	2.76 (0.91, 8.39)	0.074
Diabetes	2.83 (0.98, 8.23)	0.055

Table 4-9. Iteration 4, addition of diabetes to the model (R^2=0.22, log likelihood=-70.51, p=0.050)

Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/d,

	Adjusted OR (95% CI) n=130
Age (years)	1.00 (0.96, 1.04)
Obesity	6.00 (2.69, 13.39)
Diabetes	2.895 (1.03, 8.48)

Table 4-10. Iteration 5, removal of central obesity from the model (R2=0.20, log likelihood=-72.18, p=0.067)

Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, Diabetic = diagnosis or fasting glucose>7.0mg/d,

	Adjusted OR (95% CI) n=130
Age (years)	0.99 (0.96, 1.03)
Obesity	4.31 (1.81, 10.23)
Central obesity	2.54 (0.81, 7.97)
Diabetes	2.56 (0.83, 7.82)
Dyslipidaemia	1.35 (0.52, 3.48)

Table 4-11. Iteration 5, addition of dyslipidaemia to the model (R^2=0.22, log likelihood=-70.32, p=0.541)

Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/d, d*yslipidaemia = TG>1.7mg/dl, HDL<1.03mg/dl or on treatment*

	Adjusted OR (95% CI) n=130
Age (years)	0.99 (0.96, 1.04)
Obesity	4.55 (1.93, 10.72)
Central obesity	2.93 (0.96, 8.95)
Diabetes	3.09 (1.03, 9.30)
Total calories (kcal/day)	1.00 (0.99, 1.00)

Table 4-12. Iteration 6, addition of total calories to the final model (R^2=0.22, log likelihood=-69.81, p=0.240)

Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/d

	Adjusted OR (95% CI) n=130
Age (years)	0.99 (0.95, 1.03)
Obesity	4.57 (1.92, 10.86)
Central obesity	2.76 (0.90, 8.50)
Diabetes	3.09 (1.05, 9.10)
METmin/wk	1.00 (0.99, 1.00)

Table 4-13. Iteration 7, addition of METmin/wk to the final model (R^2=0.23, log likelihood=-69.51, p=0.163)

Abbreviations: MET = Metabolic equivalent
Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/d

Whilst the inclusion of age and diabetes did not increase the proportion of variation explained, they are well documented as strong risk factors for NAFLD and were therefore included. Presence of obesity had a strong influence on NAFLD risk (adjusted OR 4.50, 1.92-10.56). Presence of central adiposity also appeared to increase NAFLD risk (adjusted OR 2.76, 0.91-8.39), but not statistically so (p=0.074). Subsequent removal of the central obesity variable did not improve the model (Table 4-10), so it was kept in. Diabetes, as a risk factor alone, improved the strength of the model compared to components of, or presence of, the metabolic syndrome. The inclusion of this variable alone also improved the power of the model. Subsequent addition of dyslipidaemia, calorie intake and METmin/wk did not improve the model so age, obesity, central obesity and diabetes form the null model (Table 4-9).

4.4.3. Impact of ethnicity

So far the data have been used in their entirety (n=130) to give the strongest model for NAFLD risk prediction (Table 4-9). To understand the impact of ethnicity on this model, further analyses were undertaken through likelihood ratio testing, comparing the null model (Table 4-9) with models including ethnicity as an interaction effect for each variable (except age – which is not influenced by ethnicity). Again, each iteration includes the resultant likelihood ratios, R^2 values and a p value. This p value demonstrates whether the subsequent iteration results in a significant change in likelihood ratio from the null model (Table 4-9):

	Adjusted OR (95% CI) n=130
Age	0.99 (0.95, 1.03)
Obesity	6.00 (0.89, 40.36)
Obesity - South Asian - Indian	 0.50 (0.09, 2.90) 0.63 (0.09, 4.35)
Central obesity	2.78 (0.91, 8.48)
Diabetes	2.74 (0.90, 8.31)

Table 4-14. Comparator model examining the interaction effect of ethnicity on obesity (R^2=0.22, log likelihood -69.97, p=0.900)

Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/d

	Adjusted OR (95% CI) n=121**
Age	0.99 (0.95, 1.03)
Obesity	5.53 (2.05, 14.92)
Central obesity	3.05 (0.50 18.78)
Central obesity - South Asian - Indian	 0.50 (0.16, 1.61) ¥Omitted
Diabetes	2.49 (0.82, 7.53)

Table 4-15. Comparator model examining the interaction effect of ethnicity on central obesity (R^2=0.19, log likelihood -67.01)

Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/d

**When examining the impact of ethnicity on NAFLD risk in relation to central obesity, nine observations were dropped from the Caucasian group as they predicted NAFLD perfectly. This resulted in unequal numbers between the null model (Table 4-9) and this model (Table 4-15) so they could not be directly compared (i.e. no p value presented).

	Adjusted OR (95% CI) n=130
Age	0.99 (0.95, 1.03)
Obesity	6.03 (2.18, 16.68)
Central obesity	2.64 (0.83, 8.28)
Diabetes	1.54 (0.20, 11.66)
Diabetes - South Asian - Indian	 2.73 (0.37, 20.23) ¥Omitted

Table 4-16. Comparator model examining the interaction effect of ethnicity on presence of diabetes (R^2=0.27, log likelihood -66.80, p=0.290)

Obesity = Caucasians; BMI≥30, Asians; BMI≥27.5, central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetic = diagnosis or fasting glucose>7.0mg/d¥ When examining the impact of ethnicity on NAFLD risk in relation to central obesity and diabetes, data from the Indian group were omitted due to collinearity between these risk factors and Indian ethnicity i.e. the characteristics of the Indian NAFLD participants with central obesity or diabetes perfectly predicted NAFLD.

Through likelihood ratio testing, it was shown that none of these comparator models improved the null model - which did not include ethnicity (p=0.901, p=0.290). This means that the results of the original model stand even when testing the individual ethnic cohorts (i.e. there is no need for a stratified model).

Therefore, regardless of ethnicity and despite apparent differences in baseline data between groups, obesity is associated with a 4-fold increase in NAFLD risk across the study population as a whole. Presence of diabetes and central obesity appear to also have an impact on NAFLD risk, although the results are not statistically significant based on the predefined criteria on multivariate analysis. Increasing age does not appear to increase NAFLD risk within this study population.

4.5 Discussion

4.5.1 Principal findings

The baseline characteristics of the ethnic cohorts within this study were different. The Indian cohort had a lower BMI, whilst the UK Caucasian group were older, with a higher prevalence of diabetes. The Indian cohort appeared to consume more carbohydrate and more calories overall compared to the UK cohorts. When comparing NAFLD cases and controls within each ethnic cohort, BMI and rates of obesity were higher in all NAFLD groups, whilst rates of diabetes were only higher in the UK-based cohorts. Of note, all differences seen between cases and controls were much more dramatic in the UK-Caucasian cohort, with the difference in BMI between cases and controls within the Caucasian cohort being 6kg/m^2 compared to 4kg/m^2 in Indian and 3kg/m^2 in UK South Asian cohorts. There was no difference in difference in dietary intake or physical activity between any case-control groups, except that UK-Caucasians with NAFLD did significantly less exercise than controls. Through univariate analysis, increasing BMI and presence of obesity had a strong association with NAFLD across all cohorts (obesity unadjusted OR 10.83-14.17). Through multivariate logistic regression modelling, presence of obesity was strongly and independently associated with increased NAFLD risk across the whole study population (adjusted OR 4.50, 1.92-10.56, p=0.001). Central adiposity and presence of diabetes also appeared to be associated with NAFLD, but not statistically so (central obesity adjusted OR 2.76 (p=0.074), diabetes adjusted OR 2.83 (p=0.055)). Through likelihood ratio testing, it was demonstrated that the inclusion of ethnicity as an interaction effect did not alter the final regression model (Table 4-9); although it did highlight collinearity of factors within certain ethnic groups (e.g. rates of central obesity and diabetes in the Indian cohort). This may reflect the limitations of the study sample size and presence of sampling error.

The principal finding of this chapter is, therefore, that there is commonality of risk factors for NAFLD across the ethnic cohorts within this study – namely presence of obesity and diabetes. Despite this commonality of risk, conclusions can be drawn about the difference in NAFLD risk profile between the native Indian and UK-migrant South Asian cohorts. South Asians that live in the UK had a higher BMI than their native Indian counterparts (27.1kg/m^2 vs 24.3kg/m^2) and had the highest rates of obesity compared to the other ethnic cohorts (87% vs 59% vs 41%). Rates of diabetes were higher in the UK South Asian group (31.1%) than the Indian group (26.8%), whose rates of diabetes appeared significantly lower than the male participants of the large NAFLD cohort from which they were sampled (60.60%), although rates of diabetes within that cohort overall were comparable (26.33%)[244]. There does appear to be an element of dietary acculturation in the UK South Asian group, who have total calorie and carbohydrate intakes similar to the UK Caucasian group, which is less than their Indian counterparts.

This work adds to the work of Alazawi et al, who noted that increasing BMI was the strongest risk factor for NAFLD within an ethnically diverse population in London (OR 1.63 overweight, OR 2.44 obese)[228]. Whilst their study sample size was significantly larger, the baseline characteristics of their South Asian contingent were similar to the cohort used in this work (rates of overweight or obese 73% vs. 71%), suggesting the results are broadly comparable. Alazawi et al. demonstrated an increased risk of NAFLD diagnosis in people of Bangladeshi origin living in London compared to their Caucasian neighbours. Their study however was unable to analyse the cause for this increased risk or assess whether this risk has changed through migration to the UK.

Similarly, this work draws parallels with the findings of Patel et al, who demonstrated an increase in BMI and rates of obesity in UK-migrant Gujarati Indians[222]. Contrary to the findings of the work of this chapter, they found an increased intake of calories and fat in their UK-migrant group. Whilst a larger sample size, a more robust sampling technique and lifestyle measures with better accuracy mean that the data from Patel et al. is probably more reliable, the lack of a control group make evaluation of NAFLD risk impossible. Whilst it could be thought that the heterogeneity in ethnic origin within the UK South Asian cohort used in this chapter may have "diluted" the results, the use of such a specific community in a Gujarati town in India in their study mean their results may not be generalisable to the South Asian or Indian community as a whole.

4.5.2 Strengths and weaknesses of the study

This is the first prospective study to examine the differences in NAFLD phenotype between native Indians and UK-migrant groups, utilising UK Caucasians as comparators. It also included data from both NAFLD and controls, enabling case-control analysis of risk within each ethnic group. This study incorporated a large breadth of data, including demographic, anthropometric, physiological and lifestyle components, resulting in deeply phenotyped groups. Other studies that have looked at NAFLD risk in South Asians living in the UK have been limited by the diagnostic tools used to identify NAFLD, a lack of comparator groups, or by the breadth of data collected (as above).

Whilst being the only study of its kind, confirming findings that are consistent with the available literature, and demonstrating that although the NAFLD phenotype is different in migrant populations, overall risk factors for NAFLD are similar to those in India, this study does have weaknesses – most notably related to its power. The sample size calculations performed prior to commencing the study used the available evidence from NAFLD case control studies in people of South Asian ethnicity[176, 177]. Whilst powering the study adequately for each individual variable, the combination of multiple variables is likely to have reduced the power of the study overall. This means that where there appeared to be important differences between groups, the confidence intervals were too large to give any certainty to their impact. This is of particular importance in relation to the lifestyle data, where

differences in diet may have influenced differences in NAFLD risk profile, but were not conclusively demonstrated. Another limitation is the use of different FFQs for different cohorts. Whilst this was a considered decision to enable accurate, culture-specific collection of macronutrient data, this does mean that comparison of dietary data between groups may not be entirely accurate. In addition, whilst the original sampling framework for the UK cohorts of this study utilised random sampling techniques, the oversubscription of the NAFLD groups meant that the sampling technique had to become more pragmatic, which may have led to sampling bias. Whilst the Indian cohort within this work was broadly representative of the large NAFLD cohort from which they were sampled (Table 4-2), the resultant NAFLD group had certain characteristics (presence of central obesity or diabetes) that perfectly predicted NAFLD, making detailed logistic regression analysis difficult. This is demonstrated through the low log likelihood and pseudo R^2 of the final model (Table 4-13). Whilst a pseudo-R^2 value of 0.23 suggests that only 23% of the variation is explained by the variables within the model, as discussed in the last chapter (section 3.5.1), models with a binary outcome tend to have a low R^2 value, even in the presence of a perfect underlying regression relationship[213].

4.5.3 Study implications and future work

Although those factors that influence NAFLD risk are the same in both UK-migrant South Asian and native Indian groups, this work does also demonstrate some differences in body composition and dietary intake that could be attributable to a degree of acculturation. Further work is needed to analyse the impact of changes in macronutrient intake and adaptation of dietary habits (potentially through longitudinal data collection), including types of protein consumed, the ways in which food are cooked and choice of available snacks. More detailed analysis could also include data on the micronutrient composition of diet and include data on caffeine intake, which has also been linked to NAFLD risk[245].

True rates of NAFLD amongst people of South Asian ethnicity living here in the UK are unknown, and therefore there is a need for larger epidemiological studies looking at rates of disease across ethnicities. It is already known that people of South Asian heritage have higher rates of diabetes[5], and data presented within this thesis and within the literature support the suggestion that they have higher risk of NAFLD compared to Caucasians[228]. This is particularly pertinent as people of South Asian ethnicity form approximately 5.3% of the UK population[246], meaning there are a large number of people at risk of NAFLD, or with undiagnosed disease within the community. It may therefore be important to target NAFLD screening programmes within areas with a high prevalence of people of South Asian heritage.

Little is known about the interaction of genes and lifestyle on NAFLD risk. Whilst multiple genome-wide association studies have shown which variants may predispose to NAFLD risk within European cohorts, to date there has not been a study looking at people of South Asian ethnicity. Future work may include a GWAS of the Trivandrum NAFLD cohort, which will enable analysis of the interaction of dietary

habits and NAFLD genotype, and promote a deeper understanding of the change in NAFLD risk in relation to dietary acculturation.

4.6 Conclusion

Despite a commonality of risk for NAFLD across ethnic groups, South Asians who have migrated to the UK may have a different NAFLD risk phenotype compared to their native counterparts, particularly in relation to measures of body composition. These differences could be related to acculturation to a Western lifestyle, although this study was not designed to demonstrate causality. Whilst the work of Chapter 2 (Trivandrum NAFLD cohort) showed that people of Indian ethnicity were at increased risk of NAFLD compared to global estimates, further work is required to see whether differences in phenotype of UK-based South Asian groups results in increased NAFLD risk (compared to their native counterparts), and whether interaction between diet and genetics impacts this further.

Finally, in addition to changes in lifestyle factors in migrant South Asians, there may be other factors that alter NAFLD risk through migration. One such hypothesis is that change in BAT activity following relative cold adaptation of people whose heritage are from hotter climates may alter NAFLD risk. Further work looking at this hypothesis is presented in the following chapter.

CHAPTER 5: The association of brown adipose tissue activity with risk of NAFLD in Indians

5.1 Introduction

5.1.1 Brown adipose tissue

As described in Chapter 1, brown adipose tissue was first identified in the 16th century[96]. It is a highly metabolically active organ, which is able to release energy as heat through uncoupling of mitochondrial respiration and is a major effector of adaptive thermogenesis. Having previously been thought to disappear with age, it has now been shown to persist into adulthood[98], and as such has recently become a focus for research as a possible target for anti-obesity treatment.

Numerous studies have demonstrated decreased BAT activity with increasing BMI[113, 247, 248], in the presence of diabetes[114, 249] and as such may influence NAFLD risk and ethnic variation in disease prevalence.

5.1.2 BAT activity and non-alcoholic fatty liver disease (NAFLD)

The links between obesity, diabetes and NAFLD have been described in detail in Chapter 1. Briefly, excess lipid is deposited in skeletal muscle and the liver, resulting in organ-specific insulin resistance, which in turn drives further lipid accumulation within the liver through de novo lipogenesis from carbohydrate which is not utilised by insulin resistant muscle[50]. Reduced BAT activity has also been shown to play an important role in insulin resistance[114], whole body energy expenditure[249], obesity[113] and lipid metabolism[115], suggesting a possible role in the development of NAFLD.

When activated, BAT utilises fat and glucose as substrates for thermogenesis. This includes the use of intracellular triglyceride droplets, and the uptake of lipids and glucose from the circulation[250]. This was initially demonstrated in mouse models, where cold-induction or transplantation of BAT resulted in improved insulin sensitivity, lipid profiles and decreased body weight[115, 251]. Further studies in man showed that increasing BAT activity through cold activation could improve measures of insulin sensitivity[116, 117] and, on retrospective review of PET-CT scans done for clinical purposes, that those with NAFLD were less likely to have active BAT[121, 122].

5.1.3 BAT activity and ethnicity

Through Chapters 1, 2 and 4 it has been consistently shown that people of South Asian ethnicity have a more unfavourable metabolic profile compared to Caucasians, and that this results in increased risk of NAFLD. BAT activity has also been shown to be closely related to obesity and insulin sensitivity. It is therefore possible that

reduced BAT activity may underpin the disadvantageous metabolic phenotype within this group.

At present, there are limited data in the literature about the influence of ethnicity on BAT activity. The large, early retrospective studies of clinical PET-CT scans did not include data on patient ethnicity[252, 253]. Cronin et al. reported a large retrospective cohort of 6867 patients who underwent PET-CT, demonstrating again that sex, age and BMI are associated with presence of BAT, but that there were no differences between ethnic groups[254]. However, this cohort was mainly comprised of Caucasians, with small numbers of Black, Hispanic and Asian patients. More recently, a Dutch group prospectively examined the differences in BAT activity between healthy, lean Caucasians and South-Asians. Hindustani-Surinamese participants – people of South Asian ancestry who migrated to the Netherlands from Surinam (a former Dutch colony) – were recruited for both studies undertaken by this group. These South Asian participants were born in the Netherlands but had Hindustani-Surinamese parents. Their initial study of 10 Caucasian, 10 South Asian men showed no difference in cold-induced BAT activity as measured by PET-CT or BAT volume between the groups[178]. A further study undertaken a year later repeated this work, comparing BAT activity between 12 Caucasians and 12 South Asians, including measures of energy expenditure[179]. At baseline, the South Asian group had a lower energy expenditure and a lower BAT volume. Glucose uptake on cold activation did not differ between groups, but energy expenditure did seem to increase to a lesser degree in the South Asian group (20% vs 13% p=0.09) – a difference that may have been more pronounced if a larger sample size had been used. Of note, the reduced BAT volume in the South Asian group may influence the detection of small changes in glucose uptake via PET-CT. If the foci of tracer uptake are small, the radioactivity concentration can be underestimated due to resolution losses[255].

These are the only two studies to directly compare BAT activity between South Asians and Caucasians, and are limited by their small sample size, and by the fact that, although their South Asian groups are genetically South Asian, they are likely to be culturally and environmentally European, which may impact their metabolic profile and BAT activity.

5.1.4 BAT activity and genetics

BAT activity is influenced by environment, with evidence demonstrating changes in activity according to season and outdoor temperature, despite imaging being undertaken in a temperature-controlled environment[249, 252, 256]. Early studies showed that prolonged cold exposure reduced presence of shivering but did not change metabolic rate, indicating an increase in non-shivering thermogenesis from BAT[257]. More recent studies confirm this finding, and that cold acclimation can accentuate acute cold-induced BAT activity, but whether long-term cold acclimation can result in long-term changes to basal metabolic rate remain to be seen[258]. In addition to environment, genetics may also play a role in potential differences in BAT

activity between ethnicities. It has been theorised that the ancestral exposure to cold necessitated the enhancement of thermogenesis through the activity of BAT, and that the reverse may be true for the descendants of heat-adapted populations[259]. Therefore, it may be that evolutionary climate exposure plays a part in the ethnic variation in obesity and diabetes. In addition to these proposed "thrifty" genotypes, presence of genetic variants that directly influence BAT activity through UCP1 or beta-adrenergic receptors have also been shown to impact rates of obesity and diabetes and may influence ethnic differences in BAT activity[260-262]. Variants within the FTO haplotype (the strongest genetic association with obesity[263]) have recently been mechanistically linked to BAT activity, so may be of interest when examining differences between populations[264].

Another SNP that has been linked to perception and adaptation to cold is the Transient Receptor Potential (TRP) cation channel subfamily member 8 gene (TRPM8), whose allele frequency varies significantly with latitude[265]. It has been postulated that positive selection raised its frequency in the Eurasian population during the last 25,000 years making it an interesting focus for future studies of its role in determining BAT activity.

5.2 Hypothesis, aims and objectives

This chapter will address the following hypotheses, using the subsequent aims and objectives:

- Indians have lower brown adipose tissue activity than Caucasians
- Reduced brown adipose tissue activity increases NAFLD risk

AIM 4: To investigate the ethnic differences in BAT activity and its impact on NAFLD risk

Objective M) Assess the BAT activity across the multi-ethnic study

Objective N) Compare BAT activity between native Indian, UK-migrant Indian and Caucasian cohorts

Objective O) Perform case-control analysis of NAFLD risk with regard to BAT activity across the multi-ethnic study

Objective P) Analyse the impact of ethnicity on BAT activity and NAFLD risk within the multi-ethnic study

Objective Q) Analyse the impact of BAT activity on diabetes risk within the multi-ethnic study

5.3 Methods

5.3.1 Background

Methods of assessing BAT function

BAT activity can be measured in several ways, which can be broadly divided into direct and indirect measures.

Direct measures of activity require tissue samples through biopsy – usually taken from the supraclavicular depot under imaging guidance. Activity can be assessed by measuring gene expression markers (UCP1 mRNA[266], lipoprotein lipase mRNA[267]), gene products (UCP1 or BAT lipid content), or from measures of mitochondrial respiration[268, 269]. These measurements are usually limited to animal models or small human studies due to their invasive nature.

Indirect measures of BAT activity are more commonly used in research. The principal output from activated BAT is heat, so calorimetry may be the most suitable tool. Indirect calorimetry measures oxygen consumption and carbon dioxide production, enabling calculation of energy expenditure using the Weir equation (EE(*J*) = 15.818VO_2(l/min) + 5.176VCO_2(l/min)). It can therefore be used to measure the increase in energy expenditure in response to cold-activation of BAT[249, 270]. The limitation of this method, however, is that a rise in metabolic rate in response to cold may not be entirely attributable to increased BAT activity. Studies have shown that UCP1 knockout mice maintain the ability to increase their metabolic rate in response to cold[271, 272]. This means that the use of calorimetry alone may not be a sufficient method to measure BAT activity.

PET-CT

As previously described, it was the symmetrical glucose uptake in the supraclavicular fossae on PET-CT that led to the resurgence of interest in BAT physiology[273]. This uptake was subsequently proven to be due to BAT, through direct measures of activity via biopsy[274-278]. Since then, PET-CT (specifically ^{18}F-FDG PET-CT) has been considered the gold standard for measuring BAT activity in humans, where PET measures the uptake of glucose (the substrate utilised by active BAT) and CT measures the density of the tissue of interest (distinguishing it from WAT). Often the detection of active BAT on PET-CT is reported in a binary fashion as “BAT positive” or “BAT negative” rather than as a continuous measure. This is because multiple factors may influence the reporting of small changes in glucose uptake (BAT activity) between participants or studies, including variability in scanner model, system calibration, resolution, ^{18}F-FDG dose, scan duration and the region of interest (ROI) selected.

PET-CT has other limitations. It requires exposure to relatively high levels of radiation (8mSv), costs around $450 per scan and requires around 1.5hours of time from administration of ^{18}F-FDG to completion, not including time required for BAT

activation. This means that studies of healthy individuals are limited to small numbers, particularly if repeated imaging is required, reducing confidence in results. Measurement of BAT volume and activity can also vary according to PET resolution, the ROI and PET threshold criteria chosen. This means that studies may not be comparable. Larger studies can be undertaken by retrospectively looking at imaging done for clinical reasons, but conditions and protocols are optimised to minimise BAT activity, resulting in much lower prevalence of BAT activity detected [277, 279] than in those studies that activate BAT [276, 278]. Data from these studies may not necessarily be extrapolated to healthy individuals.

In addition to the practical constraints, there are also methodological limitations. PET measures the uptake of glucose (^{18}F-FDG) within tissues, and, because BAT utilises glucose as a substrate for oxidation, the amount of glucose uptake within BAT could be used as a measure of the tissue's activity. However, there are studies that show that glucose (^{18}F-FDG) uptake is unaffected in UCP1 knockout mice where thermogenesis from BAT activity is diminished[280]. It is also noted that BAT uses fatty acids as its main substrate[249]. It is the process of lipolysis of triglycerides to produce fatty acids for BAT substrate that requires glucose, rather than the utilisation of glucose itself as a substrate[281]. This means that PET-CT may not always detect metabolically active BAT.

Due to the multiple limitations to the current "gold-standard" for measurement of BAT activity, a number of other indirect measures of activity are being used in clinical research.

Magnetic resonance imaging

MRI has properties that make it an alternative to PET-CT for the measurement of BAT activity. MRI (whether MRI or MRS) can differentiate between BAT and WAT through differences in lipid/water fractions. White adipocytes contain a single lipid vacuole[282] and minimal water, whereas brown adipocytes contain multiple vacuoles[250] and significantly higher amounts of water. Studies using either MRI or MRS have demonstrated the ability to detect BAT under both cold-stimulated and thermoneutral conditions, even in the supposed absence of BAT on PET-CT[283, 284]. The triglyceride content of BAT has also been linked to insulin sensitivity in healthy volunteers[285] and has led to further studies comparing BAT fat fraction in patients with different levels of insulin resistance[114]. This clinical utility, along with absence of ionising radiation make MRI an attractive modality for the assessment of BAT activity, particularly in larger, longitudinal studies. The main restrictions of this method are its cost and limited availability.

Infrared thermography (IRT)

IRT is a process by which the infra-red radiation emitted from an object is measured, converted to radiometric data and finally presented as temperature data (often in pictorial form), using Planck's law to take into account the emissivity of the object being measured[286].

IRT is non-invasive and does not require ionising radiation and therefore offers an attractive alternate method for the indirect measurement of BAT activity. IRT requires very little equipment and can give good quality data through repeated or serial measurements of large numbers of individuals. Contrary to PET-CT, IRT measures heat as the key outcome of thermogenesis, rather than an upstream substrate. As the largest depot of BAT in adults is found within the supraclavicular fossae, this region is usually imaged, often comparing basal to stimulated thermogenesis, and, at times, in relation to a reference point.

IRT has been used to study BAT since the 1970's[287]. Over time, animal models confirmed that BAT activity as measured by IRT correlated with other methods such as PET-CT[288] or calorimetry[289]. More recently, IRT has been used to assess cold-induced BAT activity in humans[290, 291]. When comparing IRT to PET-CT, it was initially shown that those individuals who were "BAT-positive" on PET-CT had an increased temperature of the BAT hotspot, relative to a reference point in response to cooling, but there was no response in "BAT-negative" individuals[292, 293]. Subsequent work has confirmed that the BAT hotspot as measured by IRT closely corresponds to the area of maximal uptake on PET-CT in repeated imaging of healthy volunteers, demonstrating that IRT may be a suitable alternative to PET-CT for the measurement of BAT activity.

A major limitation of IRT is the fact that the measurement of BAT activity relies on the radiation of the heat it produces through adipose and subcutaneous tissues. Adipose tissue is highly insulating[294] and increased adiposity reduces skin temperature[295], which can make detection of BAT through thick adipose tissue of the neck difficult. It is also well documented that BMI is inversely associated with BAT activity[296-298], making detection of reduced activity through an area of decreased conduction potentially more difficult[279]. At present, there is a paucity of studies that examine the use of IRT in overweight/obese individuals. One study comparing lean and obese men showed no increase in energy expenditure or supraclavicular temperature via IRT in the obese group in response to cold. The limitations of this work are that the imaging protocol was not outlined, suggesting that IRT was done at a single time-point, thus losing data on the dynamic nature of BAT activity. In addition, the cooling protocol involved immersion of both hands into ice water, which is likely to have elicited a pain response, rather than pure cold-activation of BAT. This study was also limited by a small sample size (obese n=16). Contrary to this finding, Hartwig et al. examined the use of IRT in assessing BAT activity in response to glucose load and cold in 5 healthy and 5 overweight women[299]. In this study, a thermogenic response as measured by IRT was noted in all participants, both lean and obese. Again, the imaging protocol was not clear, highlighting the need for further detailed studies of this nature.

Methods of activating BAT

In addition to differences in methods of measuring BAT activity, it is also important to consider the different methods with which BAT is activated. BAT is activated a

number of ways: through cooling[248, 258], use of pharmacological agents[300], diet[301] or even stress[302]. As the core function of BAT is thermogenesis, cooling is the most common method of stimulation used. This can be done acutely, using short cooling protocols[303], or through cold acclimation over a period of time[258]. Cooling can be achieved through lowering of ambient air temperature[292], through immersion in cold water[290] or through use of cooling blankets[293]. Whichever modality is used, the temperature selected must not be so low as to cause shivering, thus switching off non-shivering thermogenesis, or pain, which activates the sympathetic nervous system and may influence BAT activity. Cooling protocols that utilise water or blankets can be standardised for a given study or personalised to the individual participant, which involves reducing the temperature to a point below which shivering is induced, and then bringing the temperature up until shivering stops. With such variability in methods and protocols, comparison of data between studies can be difficult, as highlighted by a Dutch group who demonstrated different outcomes when comparing BAT activity between South Asians and Caucasians when using different techniques[304].

When cooling is used as the method of activation, it is also pertinent to consider its duration. Initial response to cold stimulus is rapid; 5 minutes of cooling is enough to detect a change in temperature via IRT[290]. Longer periods of cooling have however shown further increases in temperature, with studies using durations of 30 minutes, 1 hour or even 2 hours[291, 292, 303]. The comfort of participants must also be considered, as well as how pragmatic it is to conduct long imaging protocols in studies of reasonable sample size.

5.3.2 Study design

This study used data collected from the multi-ethnic study outlined in Chapter 4. As previously described, this cohort was created to enable collection and analysis of data on risk factors associated with NAFLD, and to enable comparison of both clinical and environmental risk profiles between ethnic groups. Data from different variables known to impact NAFLD risk and/or BAT activity were collected in a cross-sectional manner, at a single time point. This included age, measures of body composition (including subclavicular skinfold thickness as a measure of neck adiposity) and presence of diabetes – variables that are known to impact BAT activity.

The eligibility criteria and recruitment method for this study are outlined in Chapter 4 (4.3). This part of the study regime was outlined within the participant information sheet, was included within the consent form and was covered by the same ethical approvals (both in India and in the UK).

Through the work of this chapter, the aim is to investigate whether ethnic differences in BAT activity may contribute to the increased metabolic/NAFLD risk in people of Indian ethnicity. BAT activity was measured using IRT in both NAFLD cases and controls, comparing native Indian, UK Indian and UK Caucasian populations using the study outlined in Chapter 4. Although evidence for the use of IRT in participants with a metabolic phenotype is limited, pragmatically it enables assessment of BAT

activity within a large sample size (n=120), using tools that are portable and practical for a study that involves assessment of people within a lower-middle income environment. In addition, the inclusion of metabolic variables (obesity, presence of diabetes) in final logistic regression analysis will reduce any impact of confounding.

A cooling protocol that utilised a water blanket wrapped around the right arm and set to 14˚C was chosen. Cooling occurred after 20 minutes of acclimation to a room temperature of 22˚C and continued for 20 minutes to enable dynamic assessment of BAT activity. An individualised cooling protocol was not suitable for this study, as the induction of shivering to identify an appropriate cooling temperature would invalidate further assessment of non-shivering thermogenesis during the single study visit. Whilst preferable, electromyography to assess for presence of shivering was unavailable, so participants were closely monitored and asked at regular intervals about presence of shivering. This protocol has been utilised in other studies of this nature within this study group and has shown to provide adequate activation of BAT[293].

5.3.3 Study regime

As outlined in Chapter 4, participants attended a single morning study visit lasting approximately 3 hours, having fasted from midnight and avoided caffeine and strenuous exercise in the preceding 24 hours. These specifics were outlined in the participant information sheet as caffeine may increase sympathetic activation of BAT and strenuous exercise can increase skin temperature, both of which may influence BAT activity as measured by IRT. On arrival, participant eligibility was checked against the inclusion/exclusion criteria, which included a brief medical, alcohol and drug history. Participants then underwent measures of body composition and venepuncture for assessment of basic blood parameters. UK-Caucasian and UK-South Asian participants had skinfold thickness measured to the nearest 0.1cm in the midline below the right clavicle using Harpenden callipers (HaB direct, Warwickshire, UK). This was used as a surrogate marker of neck adiposity and enabled analysis of its correlation with other measures of body composition. Participants then completed questionnaires on lifestyle habits and underwent an abdominal ultrasound for assessment of liver fat.

For the assessment of BAT activity, IRT was undertaken in line with the Delphi consensus on the measurement of human skin temperature[305], and through modification of existing protocols used within the local study group (Early Life Research Group, Division of Child Health, University of Nottingham). Imaging was done in a room under closely monitored conditions. The ambient temperature was set to 22°C (thermoneutral zone 22-24°C), with documentation of the external air temperature being noted from the Meteorological Office (www.metoffice.gov.uk). If air conditioning was required to maintain a thermoneutral ambient temperature, a partition was used to ensure no direct airflow across the ROI. Participants were positioned away from sources of infrared radiation (e.g. electronic devices or lighting). Room temperature and humidity were measured and monitored using a

psychrometer to the nearest 0.01°C (M0297 Extech). Thermal imaging was undertaken using a FLIR E60 60Hz infrared camera to the nearest 0.01°C (FLIR systems, Wilsonville, Oregon, USA), which was calibrated before each participant using atmospheric temperature and humidity, reflective temperature and with emissivity set at 0.98. The camera was positioned one metre from and perpendicular to the ROI. The ROI is the supraclavicular fossa, as outlined by the outline of the neck laterally, using the point where the sternocleidomastoids reach the lateral part of the neck superiorly, the clavicle inferiorly, with the sternal notch as the central point (Figure 5-1.). Foil tabs were attached to each of these anatomical landmarks, enabling analysis of subsequent thermal data.

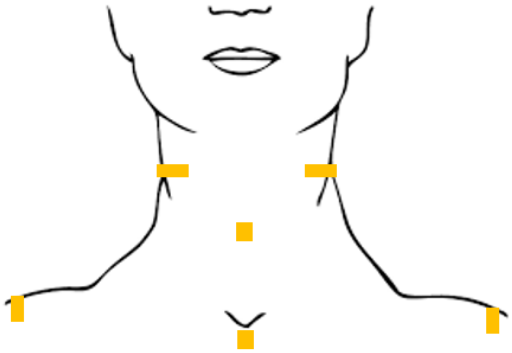


Figure 5-1. Anatomical landmarks of the region of interest (ROI)

Participants wore standardised clothing provided by the investigator – including shorts and T-shirt with an open neck away from the ROI. Participants were asked to sit upright, with their head in a neutral position for the entirety of the protocol. A cooling blanket (Balnketrol II, Cincinnati subzero) was attached to the participants' right arm, and an oxygen saturation probe (V100, Dinamap technology) was attached to a finger on the left hand. Participants' temperature was taken using an ear thermometer (Thermoscan pro 6000, Braun) and blood pressure was measured (V100, Dinamap technology) from the left arm. Ibuttons (Maxim, San Jose, USA) were secured at seven anatomical positions (Figure 5-2.), having been programmed to take measurements every minute, enabling calculation of mean skin temperature throughout the imaging protocol using the Hardy-Du Bois formula[306]: (0.07xhead)+(0.14xarm)+(0.05xhand)+(0.07xfoot)+(0.13xleg)+(0.19xthigh)+(0.35xtrunk).

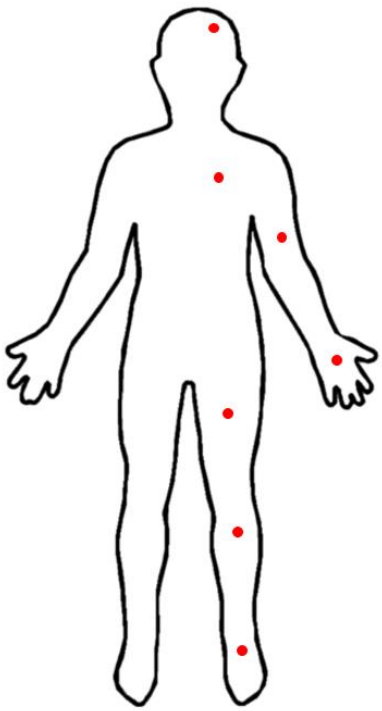


Figure 5-2. Anatomical locations of ibuttons

Imaging was recorded through thermal video of the participant's ROI using FLIR+ software (FLIR systems, Wilsonville, Oregon, USA). The first 20 minutes of imaging was completed at the thermoneutral temperature of the room (22°C). Pulse rate was documented every minute (V100, Dinamap technology) to show if sympathetic activation had occurred during cooling, and at regular intervals, the participant was asked if there was presence of shivering. After 20 minutes of acclimatisation, the cooling blanket was activated at a temperature of 14°C. Imaging and monitoring continued for a further 20 minutes of cooling. On completion of the imaging protocol, measurement of blood pressure was repeated, as were measurements of ambient temperature and humidity. Participants were then given a packed lunch and allowed to leave.

5.3.4 Data collection and management

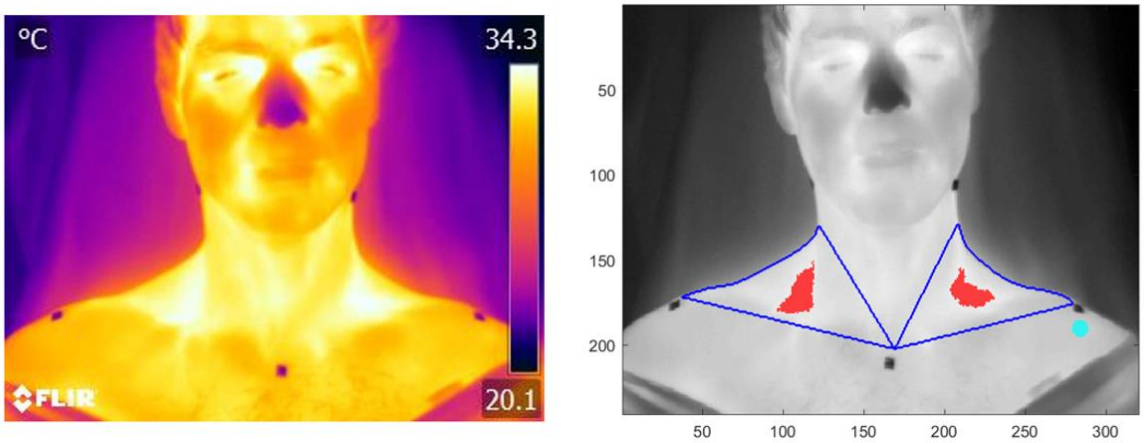
Data collection and management strategies for the study overall were outlined in Chapter 4 (4.3). Below is information pertaining to data used specifically in this chapter.

Data management

Thermal imaging data files were saved in duplicate using the same study number format as outlined previously (SA-XXX) – one on an external hard drive kept within a locked drawer within a locked office on NHS premises, and one on a University of Nottingham internet-based storage platform.

Data types

Thermal imaging data were configured utilising code written using MATLAB software (MathWorks, Massachusetts, USA). This code was designed and written by Dr James Law. Still images were identified for every 5 seconds of video, with re-selection of an image within a 2.5 second range if the ROI was not clearly presented (i.e. if participant had turned their head). For each image, the ROI as demarcated by the foil markers was outlined, where the medial and inferior borders were defined by straight lines between the appropriate apices (Figure 5-3.).



a) b)

Figure 5-3. a) Direct thermal image captured using FLIR E60 60Hz infrared camera b) MATLAB output outlining region of interest bilaterally (blue lines), reference point (turquoise circle) and hottest 10% of pixels (red)

The contour of the neck was defined programmatically by identifying the temperature gradient between the participant and the background. The hottest 10% of pixels within the ROI were identified and the median of these points were calculated (equivalent to the 95th percentile). Corresponding graphical outputs were created to visualise this "hotspot". A reference region consisting of a circle of 10-pixel diameter was selected just below the left acromioclavicular apex for comparison. This area was identified as the closest comparator to the mean skin temperature from the ibutton data as opposed to a reference point just below the sternal notch apex as used in other studies (Figure 5-4.)[303]. This thermal imaging reference was used in analysis of change of ROI temperature in response to cold, as opposed to using mean skin temperature, as it enabled calculation of change in relative temperature using a single modality (IRT) rather than using different modalities (IRT vs. ibutton). All data conversion and analysis was completed by the author.

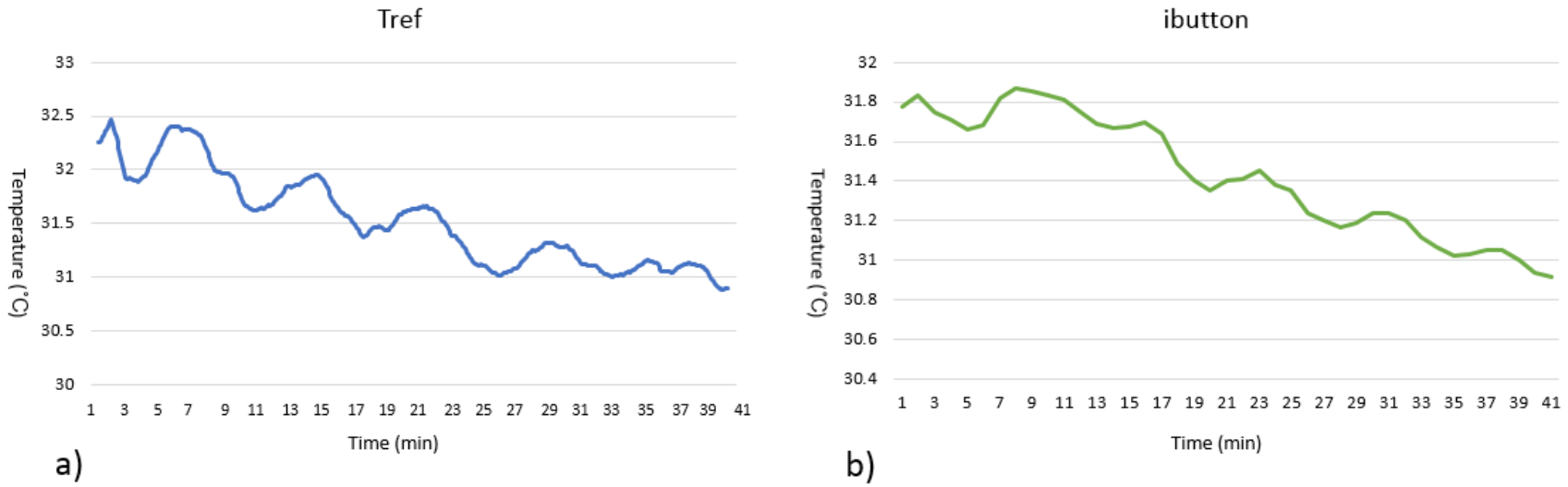


Figure 5-4. a) Thermal data from infrared thermography reference point (rolling average over 1 minute) b) Average skin temperature from ibuttons (mean skin temperature)

As with other imaging modalities, radiometric data are subject to an error term, through physiological rhythms, inherent "noise" within the measurement, or a combination of the two. "Noise" can be limited through averaging out data over a time period. Therefore, a rolling average (10 time-points, 50 seconds) was applied to the thermal data created. Thermal data was collected and presented as per example data in Figure 5-5.

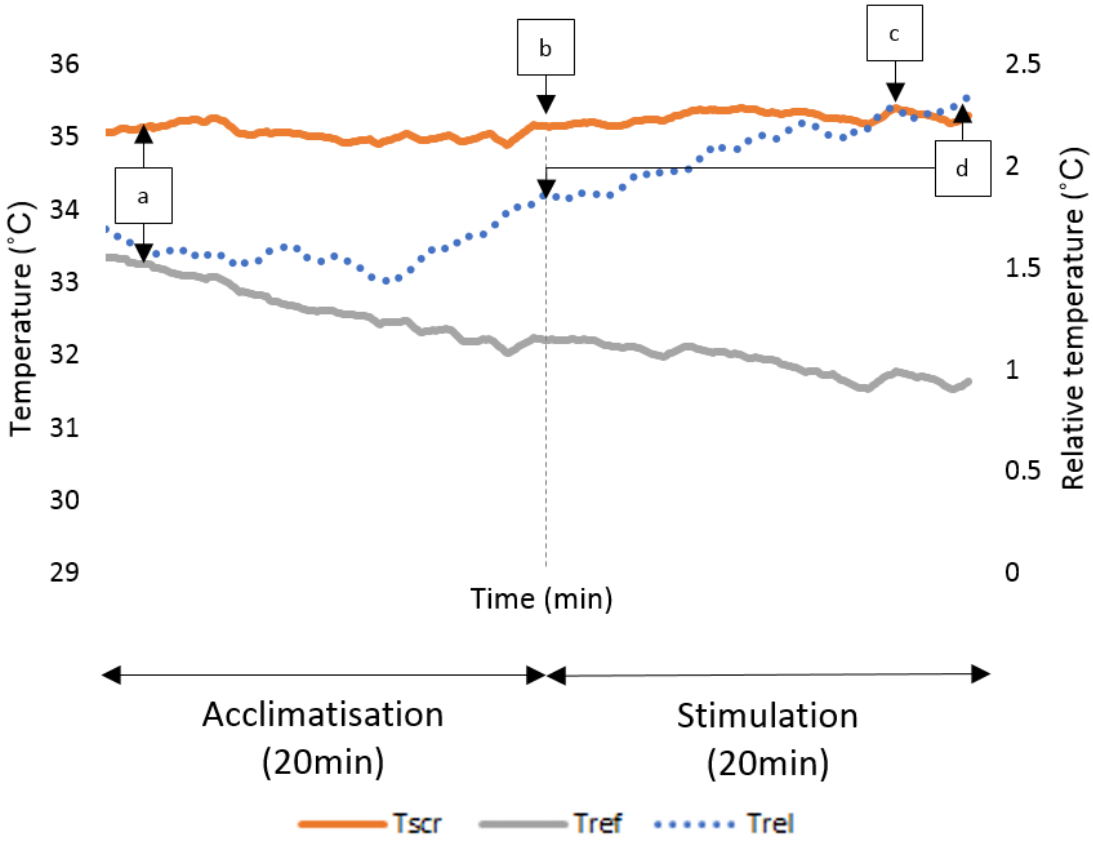


Figure 5-5. Example thermal imaging data output a) Trel = Tscv-Tref b) baseTscv c) peakTscv d) ΔTrel = peakTrel-baseTrel

baseTscv = baseline temperature of the supraclavicular ROI in the minute prior to cooling, baseTref = baseline temperature of the reference point in the minute prior to cooling, baseTrel = baseline relative temperature (baseTscv-baseTref), peakTscv = maximal temperature of the supraclavicular ROI during cooling, peakTref = temperature of the reference point at peak Tscv, peakTrel = maximal relative temperature, ΔTrel = Change in relative temperature (peak relative temperature – base relative temperature (Tscv-Tref)

Whilst there were numerous measures of BAT activity reported in response to cold, ΔTrel has been shown to correlate closely to BAT activity as measured via PET-CT[303], and as such was the primary endpoint for analysis.

5.3.5 Data analysis

Baseline data from each group (NAFLD and control), within each ethnic cohort were initially presented as follows (Table 5-1): Categorical data were presented as numbers (percentage) and continuous data as mean (SD). As described previously, the reported variables within this work would usually be considered Normally distributed within the populations sampled and study numbers within groups were too small to make statistical testing for Normality (such as the Shapiro-Wilks test) meaningful. Central limit theorem suggests that, when testing for differences between groups (NAFLD cases and controls), these differences are also expected to be Normally distributed[198]. In addition, it can be assumed that the data had homogeneity of variance and that data points were independent and were derived from an equal-interval scale. For these reasons, all continuous data were considered parametric. Pearson's correlation coefficient was used to examine for collinearity between BMI and subclavicular skinfold thickness, with a correlation coefficient of $r \geq 0.4$ considered moderate correlation and $r \geq 0.7$ strong correlation[243]. Correlation between subclavicular skinfold thickness and BAT activity (ΔTrel) was also assessed.

Differences were analysed between baseline characteristics of the ethnic cohorts as a whole and between NAFLD cases and controls (data as presented in Chapter 4). Comparisons were then made in environment (temperature/humidity) and measures of sympathetic activity pre- and post-cooling, in BAT activity between ethnic cohorts, between the control groups of each ethnic cohort and in BAT activity between NAFLD cases and controls overall. These were done using unpaired, two-tailed Students t-test of equal variance for continuous data (parametric), and chi squared and two-sample test of proportion for categorical data. For comparisons between more than two groups (i.e. between the three ethnic cohorts), analyses were done using ANOVA with Bonferroni correction where differences were identified for continuous data. Statistical significance was set at $p<0.05$.

To understand the confounding effect of different variables on BAT activity and subsequent NAFLD risk, an adjusted logistic regression model was created. This was initially done using the entire dataset (n=126) to give the most power to detect differences between NAFLD and control groups overall. A regression model was created through the addition of sequential variables known to impact BAT activity and NAFLD risk, including those identified a priori from the literature and the work of the preceding chapters i.e. age, BMI and presence of diabetes.

To understand the impact of incorporating the effect of ethnicity within this model, further analysis was undertaken through likelihood ratio testing, using ethnicity as an interaction term for each variable (as demonstrated previously in Chapter 4 – 4.4.3). Ethnicity could not be included as an independent variable within the final model, as the number within each ethnic group was fixed. Comparator models were therefore

created to examine whether there was an additional effect of ethnicity on BAT activity and BMI (not age - which is not influenced by ethnicity). South Asian and Indian ethnicity were compared to Caucasian ethnicity as the base category.

Statistical analyses were performed using STATA v. 15 (Statacorp, Texas, USA).

5.4 Results

A total of 130 participants were recruited to the study (UK Caucasian 44, UK South Asian 45, Indian 41). Thermal imaging data were corrupted in three participants and one did not undergo IRT. These participants were therefore excluded from analysis leaving a study size of n=126. The baseline demographics of the study group are presented in Table 5-1. More detailed participant characteristics were presented in Chapter 4. Data presented here are limited to those that are pertinent to this body of work.

	UK South Asian		UK Caucasian		Indian	
	NAFLD n=22	**Control n=20**	**NAFLD n=24**	**Control n=19**	**NAFLD n=20**	**Control n=21**
Age (years) Mean (SD)	50 (10)	45 (12)	60 (13)	50 (16)	49 (12)	48 (12)
BMI (kg/m^2) Mean (SD)	27.64 (2.84)	25.46 (2.67)	31.57 (4.98)	25.31 (3.73)	26.46 (3.43)	22.20 (3.69)
Central obesity n (%)	18 (81.82)	15 (75.00)	24 (100.00)	10 (52.63)	18 (90.00)	11 (52.38)
Subclavicular skinfold thickness (cm) Mean (SD)	7.54 (1.95)	5.74 (1.84)	8.69 (3.00)	5.31 (1.98)	N/A	N/A
Diabetes n (%)	10 (45.5)	3 (15)	20 (83.3)	6 (30)	6 (30)	5 (23.8)

Table 5-1. Brown adipose tissue study participant characteristics

Abbreviations: BMI = Body mass index

Central obesity = ≥90cm for Asian cohorts, ≥102cm for Caucasian cohort. Diabetes = formal diagnosis or fasting glucose>7.0mg/dl

5.4.1 Participant characteristics

Ethnic comparison

UK Caucasian participants were older than both UK South Asian (p=0.013) and Indian participants (p=0.033) – who were of similar age (p=1.000). The Indian group had a lower BMI than both the UK Caucasian (p<0.001) and UK South Asian participants (p=0.045) – whose BMI was lower, but not statistically so than their Caucasian counterparts (p=0.060). There were no differences in rates of central obesity between ethnic cohorts. Rates of diabetes were higher in the UK Caucasian cohort compared to the UK South Asian (p=0.006) and Indian cohorts (p=0.002). There was no difference in diabetes rates between UK South Asians and Indians (p=0.679).

Body composition comparison

There was no difference in subclavicular skinfold thickness between UK Caucasians and UK South Asians (p=0.398). There was a strong positive correlation between subclavicular skinfold thickness and BMI (r=0.73 p<0.001) meaning BMI as a measure could be used to adjust for differences in neck adiposity. There was a moderate, negative correlation between subclavicular skinfold thickness and BAT activity (ΔTrel) (r=-0.49 p<0.001), but, because there were no data for the Indian cohort, this variable was not included in the final logistic regression model (Figure 5-6).

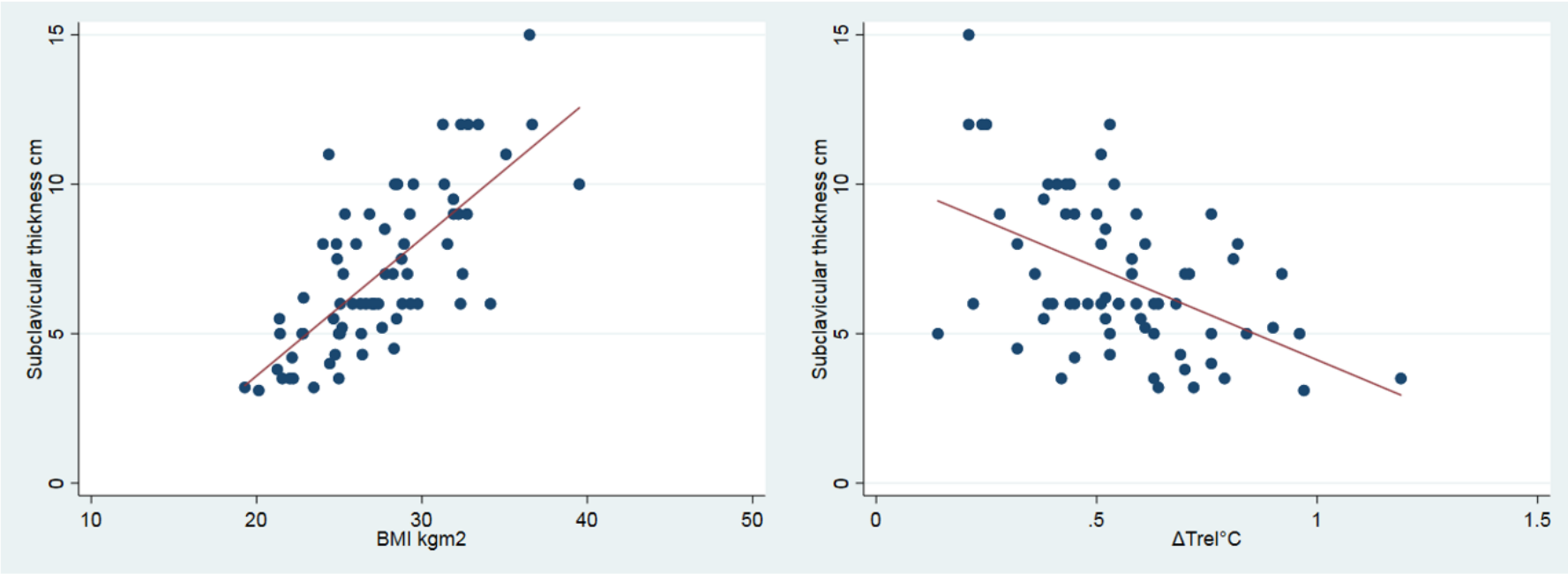


Figure 5-6. Scatterplot demonstrating correlation between subclavicular skinfold thickness (cm) and a) BMI (kg/m^2) (r=0.73, p<0.001) b) ΔTrel (°C) (r=-0.49, p<0.001)

Abbreviations: BMI = Body mass index

ΔTrel = *Change in relative temperature (peak relative temperature – base relative temperature*

Case-control comparison

Comparing NAFLD cases to controls overall (n=126), those with NAFLD were older (p=0.016) and had a higher BMI (p<0.001). Rates of central obesity and diabetes were significantly higher in the NAFLD group (p<0.001).

5.4.2 Environment

During study visits, measures were taken to ensure that the environment remained controlled for both ambient temperature and humidity, and that BAT activation was the result of cooling alone and not through activation of the sympathetic nervous system (either through anxiety or pain).

Room temperature and humidity were measured at the start and end of the imaging protocol. Mean start temperature overall was 22.21°C (0.98) and end temperature was 22.51°C (1.31), showing a statistically significant difference of 0.30°C (p=0.004). Although statistically different, this is unlikely to make a significant clinical difference, given that cold activation requires a much larger change in temperature (22˚C to 14°C). Mean overall start humidity was 49.85% (14.06) and end humidity was 49.83% (13.52), showing no significant difference through the imaging protocol (p=0.742).

Mean pulse rates for the acclimatisation period and cooling period were calculated for each participant, and the differences between these means across the whole study population were not significant (p=0.261). Blood pressure was taken before and after the imaging protocol. Systolic BP decreased overall following cooling (-3.4mmHg p=0.002) which is clinically appropriate given that participants were sitting still and were being actively cooled. It is also unlikely therefore, that sympathetic drive had increased. There was no difference in diastolic BP following cooling (p=0.277).

IRT for the Indian cohort was undertaken in Trivandrum, India between 20/05/17 and 30/05/17. The average external temperature during this time was 30˚C. IRT for the UK cohorts was undertaken in Nottingham, UK between October 2017 and March 2019, across different seasons and external temperatures. The annual average temperature in Nottingham is significantly lower at 10˚C (p<0.001).

5.4.3 BAT activity and ethnicity

Raw measures of BAT activity for the different ethnic cohorts are presented in Table 5-2.

		UK South Asian n=42	UK Caucasian n=43	Indian n=41
Absolute	Base Tscv (°C) Mean (SD)	34.09 (0.48)	34.12 (0.64)	33.61 (0.70)
	Peak Tscv (°C) Mean (SD)	34.28 (0.50)	34.29 (0.68)	33.87 (0.68)
	Δ Tscv (°C) Mean (SD)	0.20 (0.14)	0.22 (0.15)	0.28 (0.26)
Relative	Base Trel (°C) Mean (SD)	2.13 (0.54)	2.24 (0.56)	2.68 (0.64)
	Peak Trel (°C) Mean (SD)	2.70 (0.60)	2.74 (0.62)	3.04 (0.61)
	Δ Trel (°C) Mean (SD)	0.57 (0.21)	0.50 0.20)	0.36 (0.21)

Table 5-2. Measures of brown adipose tissue activity for each cohort

baseTscv = baseline temperature of the supraclavicular ROI in the minute prior to cooling, baseTrel = baseline relative temperature (baseTscv-baseTref), peakTscv = maximal temperature of the supraclavicular ROI during cooling, peakTrel = maximal relative temperature, ΔTscv = peakTscv – baseTscv), ΔTrel = Change in relative temperature (peak relative temperature – base relative temperature (Tscv-Tref)

At baseline, Tscv was significantly lower in the Indian group compared to UK Caucasians (p=0.001) and UK South Asians (p=0.002). The same is seen for peak Tscv (Caucasians p=0.009, South Asians p=0.011 Figure 5-7). There were no differences in ΔTscv between ethnic groups (p=0.132). There was no difference in supraclavicular temperatures between UK Caucasians and UK South Asians.

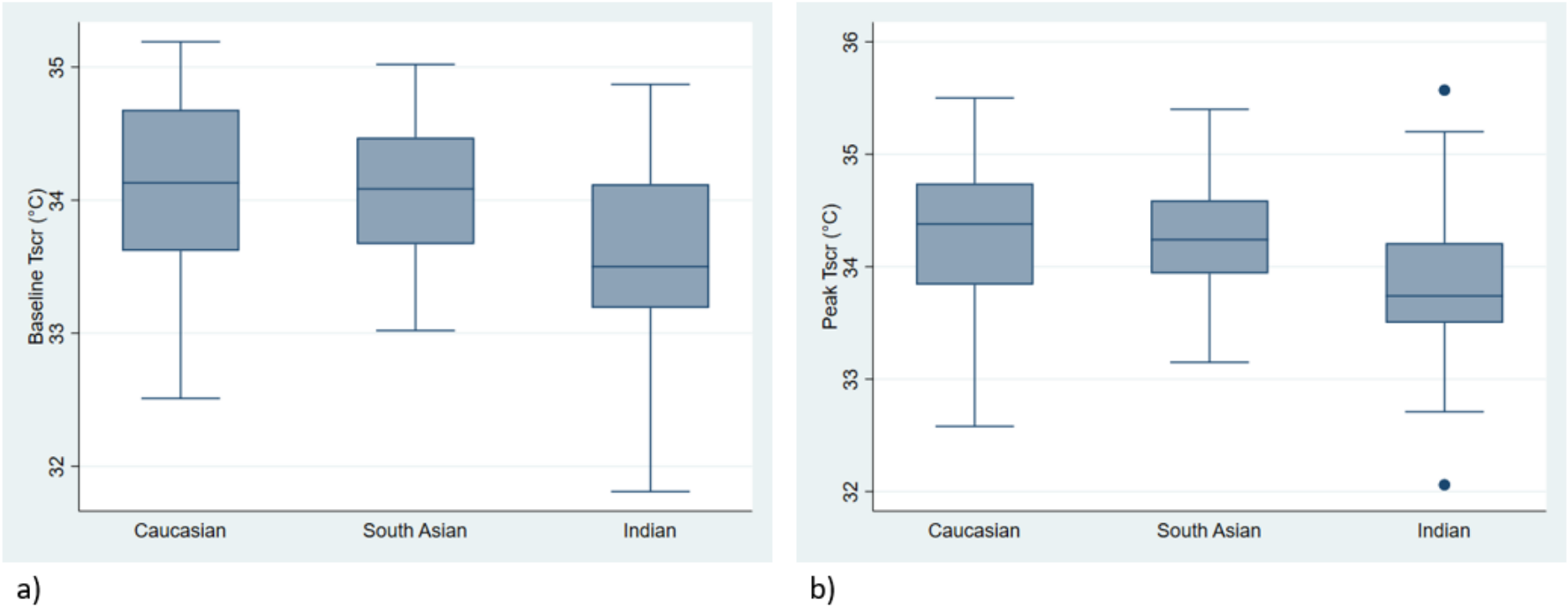


Figure 5-7. Boxplots depicting a) baseline and b) peak Tscv for each ethnicity

Tscv = temperature of the supraclavicular region of interest

Relative temperature at baseline was higher in Indians compared to UK Caucasians ($p=0.003$) and UK South Asians ($p<0.001$), with no difference between the UK cohorts ($p=1.000$). At peak, relative temperature was higher in Indians than UK South Asians ($p=0.038$) but not significantly so compared to UK Caucasians ($p=0.078$ Figure 5-8). Again, there was no difference between the UK cohorts ($p=1.000$). The change in relative temperature from baseline to peak ($\Delta Trel$) was significantly lower in Indians compared to UK Caucasians ($p=0.010$) and UK South Asians ($p<0.001$), with no difference between the two UK Cohorts ($p=0.286$).

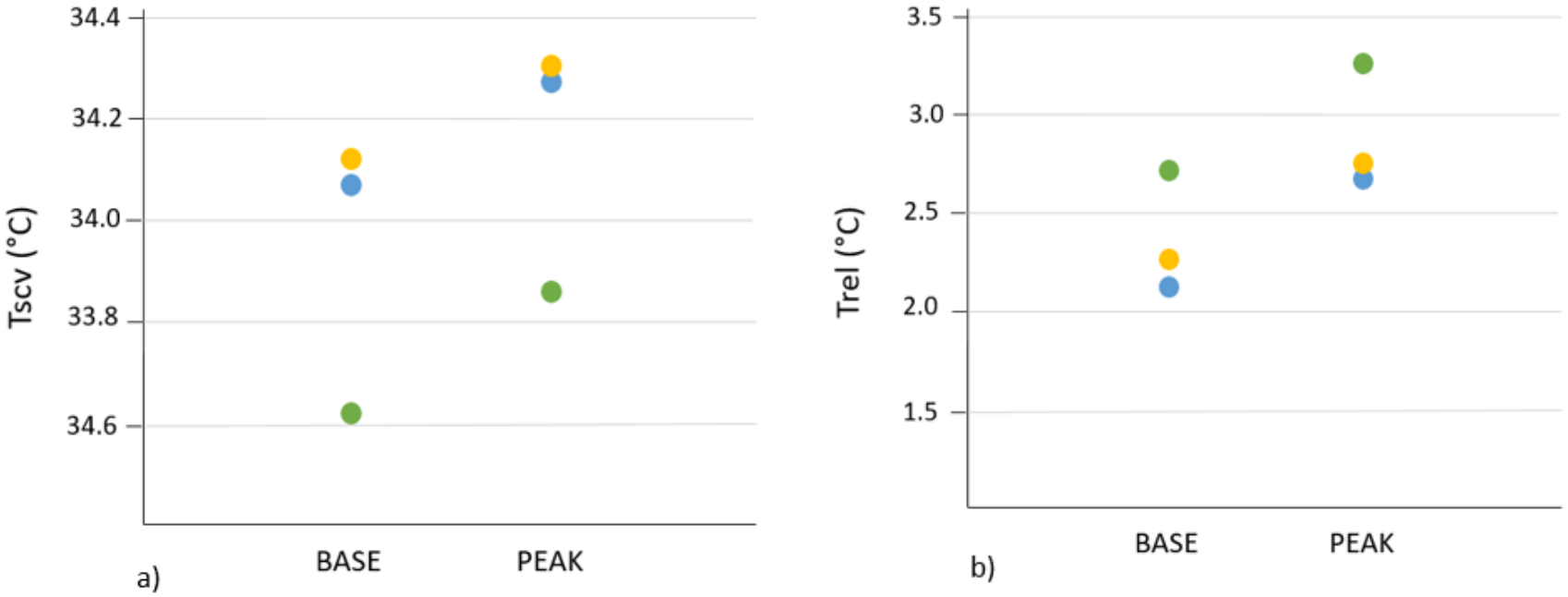


Figure 5-8. a) Mean baseline and peak Tscv b) Mean baseline and peak Trel. UK-South Asian (blue), UK-Caucasian (yellow), Indian (green)

Tscv = temperature of the supraclavicular region of interest, Trel = temperature of the supraclavicular region of interest relative to a reference point

Despite having a lower BMI (known to correlate with higher BAT activity), ΔTrel was significantly lower in the Indian cohort compared to the UK Caucasian (-0.134 p=0.010), and UK South Asian (-0.210 p<0.001) cohorts. There was no difference in BAT activity between the UK cohorts (p=0.286).

To understand the impact of external environment better, BAT activity was compared between the control groups within each cohort (Table 5-3).

	UK South Asian Controls n=20	**UK Caucasian Controls n=19**	**Indian Controls n=21**
Δ Trel (°C) Mean (SD)	0.617 (0.256)	0.567 (0.184)	0.363 (0.193)

Table 5-3. Brown adipose tissue activity in control groups

ΔTrel = Change in relative temperature within supraclavicular region of interest

Again, a significant difference was seen between the Indian and both UK Caucasian (-0.204 p=0.012) and UK South Asian control groups (-0.254 p=0.001) with no difference between the UK cohorts (0.050 p=1.000).

5.4.4 BAT activity and NAFLD

Initial, unadjusted analysis of differences in BAT activity between NAFLD cases and controls overall (n=126) showed that baseline Tscv was lower in participants with NAFLD (p=0.020), as was peak Tscv, although not significantly (p=0.058). There was no difference in ΔTscv between groups (p=0.893). Baseline Trel was again lower in the NAFLD group (p=0.028), as was peak Trel (p=0.004) but overall, although ΔTrel was lower in the NAFLD group, it was not significantly so (p=0.086). These unadjusted data are presented in Table 5-4.

		NAFLD n=66	Control n=60	P value
Absolute	Base Tscr (°C) Mean (SD)	33.81 (0.65)	34.08 (0.63)	0.020
	Peak Tscr (°C) Mean (SD)	34.04 (0.64)	34.26 (0.64)	0.058
	Δ Tscr (°C) Mean (SD)	0.23 (0.21)	0.23 (0.17)	0.893
Relative	Base Trel (°C) Mean (SD)	2.23 (0.52)	2.47 (0.71)	0.028
	Peak Trel (°C) Mean (SD)	2.67 (0.53)	2.99 (0.67)	0.004
	Δ Trel (°C) Mean (SD)	0.44 (0.20)	0.51 (0.24)	0.086

Table 5-4. Measures of brown adipose tissue activity in NAFLD and Control groups

baseTscv = baseline temperature of the supraclavicular ROI in the minute prior to cooling, baseTrel = baseline relative temperature (baseTscv-baseTref), peakTscv = maximal temperature of the supraclavicular ROI during cooling, peakTrel = maximal relative temperature, ΔTscv = peakTscv – baseTscv), ΔTrel = Change in relative temperature (peak relative temperature – base relative temperature (Tscv-Tref)

5.4.5 Logistic regression

To understand the confounding effect of different variables on BAT activity and subsequent NAFLD risk, an adjusted logistic regression model was created. This was initially done using the entire dataset (n=126) to give the most power to detect differences between NAFLD and control groups overall. This model was created through the addition of sequential variables known to impact BAT activity, including age, BMI and diabetes.

The development of the final model is outlined through the tables below, which demonstrate the sequential addition of variables and the resultant likelihood ratios, R^2 values and a p value. This p value demonstrates whether the subsequent iteration results in a significant change in likelihood ratio from the previous iteration.

	Adjusted OR (95% CI) n=126
Constant	1.10 (0.78, 1.56)

Table 5-5. Iteration 0 (R^2=0.00, log likelihood=-87.19)

	Adjusted OR (95% CI) n=126
BAT activity (ΔTrel (°C))	0.24 (0.05, 1.24)

Table 5-6. Iteration 1, addition of BAT activity to the model (R^2=0.02, log likelihood=-85.69, p=0.083)

Abbreviations: BAT = brown adipose tissue
ΔTrel = change in relative temperature within the supraclavicular region of interest

	Adjusted OR (95% CI) n=126
BAT activity (ΔTrel (°C))	0.35 (0.06, 1.89)
Age (years)	1.03 (1.00, 1.06)

Table 5-7. Iteration 2, addition of age to the model (R^2=0.04, log likelihood=-83.51, p=0.037)

Abbreviations: BAT = brown adipose tissue
ΔTrel = change in relative temperature within the supraclavicular region of interest

	Adjusted OR (95% CI) n=126
BAT activity (ΔTrel (°C))	0.42 (0.06, 1.89)
Age (years)	1.01 (0.97, 1.04)
BMI (kg/m^2)	1.33 (1.17, 1.51)

Table 5-8. Iteration 3, addition of BMI to the model (R^2=0.21, log likelihood=-69.21, p<0.001)

Abbreviations: BAT = brown adipose tissue, BMI = Body mass index
ΔTrel = change in relative temperature within the supraclavicular region of interest

	Adjusted OR (95% CI) n=126	P value
BAT activity (ΔTrel (°C))	0.47 (0.07, 3.20)	0.444
Age (years)	0.99 (0.96, 1.03)	0.769
BMI (kg/m^2)	1.31 (1.15, 1.50)	<0.001
Diabetes	1.99 (0.70, 5.66)	0.199

Table 5-9. Iteration 4, addition of diabetes to the model (R^2=0.22, log likelihood=-68.38, p=0.197)

Abbreviations: BAT = brown adipose tissue, BMI = Body mass index
ΔTrel = change in relative temperature within the supraclavicular region of interest, diabetes = *formal diagnosis or fasting glucose>7.0mg/dl*

Whilst the inclusion of age and diabetes did not increase the proportion of variation explained within the model, the evidence presented in the introduction to this chapter highlights their documented impact on BAT activity and were therefore included. Table 5-9 outlines the final logistic regression model and demonstrates that BAT activity is not associated with NAFLD risk (adjusted OR 0.47 0.07-3.20 p=0.444), and that the biggest impact on NAFLD risk appears to come from BMI (OR 1.31 1.15-1.50 p<0.001).

5.4.6 Impact of ethnicity

To understand the impact of incorporating the different ethnic groups within this model, further analysis was undertaken through likelihood ratio testing, using ethnicity as an interaction term for each variable (as demonstrated previously in Chapter 4 – 4.4.3). Ethnicity cannot be included as an independent variable within the final model, as the number within each ethnic group was fixed. Therefore, comparator models were created to examine whether there is an additional effect of ethnicity on BAT activity, BMI and presence of diabetes (not age - which is not influenced by ethnicity). South Asian and Indian ethnicity were compared to Caucasian ethnicity as the base category.

	Adjusted OR (95% CI) n=126
BAT activity (ΔTrel (°C))	0.30 (0.03, 2.73)
BAT activity (ΔTrel (°C))	
- South Asian	2.42 (0.39, 14.82)
- Indian	6.97 (0.65, 74.62)
Age (years)	0.99 (0.96, 1.04)
BMI (kg/m^2)	1.35 (1.17, 1.54)
Diabetes	2.21 (0.76, 6.46)

Table 5-10. Comparator model examining the interaction effect of ethnicity on BAT activity (R^2=0.23, log likelihood=-67.00, p=0.252)

Abbreviations: BAT = brown adipose tissue, BMI = Body mass index
ΔTrel = change in relative temperature within the supraclavicular region of interest, diabetes = *formal diagnosis or fasting glucose>7.0mg/dl*

	Adjusted OR (95% CI) n=126
BAT activity (ΔTrel (°C))	1.01 (0.12, 8.74)
Age (years)	1.00 (0.96, 1.04)
BMI (kg/m^2)	1.34 (1.17, 1.53)
BMI (kg/m^2) - South Asian - Indian	 1.02 (0.98, 1.06) 1.04 (0.99, 1.09
Diabetes	2.33 (0.78, 6.93)

Table 5-11. Comparator model examining the interaction effect of ethnicity on BMI (R^2=0.24, log likelihood=-66.62, p=0.172)

Abbreviations: BAT = brown adipose tissue, BMI = Body mass index
ΔTrel = change in relative temperature within the supraclavicular region of interest, diabetes = *formal diagnosis or fasting glucose>7.0mg/dl*

	Adjusted OR (95% CI) n=126
BAT activity (ΔTrel (°C))	1.78 (0.19, 17.00)
Age (years)	0.99 (0.95, 1.04)
BMI (kg/m^2)	1.43 (1.22, 1.67)
Diabetes	4.75 (0.64, 35.02)
Diabetes - South Asian - Indian	 2.99 (0.44, 20.35) ¥Omitted

Table 5-12. Comparator model examining the interaction effect of ethnicity on presence of diabetes (R^2=0.28, log likelihood=-63.19)

Abbreviations: BAT = brown adipose tissue, BMI = Body mass index
ΔTrel = change in relative temperature within the supraclavicular region of interest, diabetes = *formal diagnosis or fasting glucose>7.0mg/dl*

¥ As seen in Chapter 4 (4.4.3), when examining the impact of ethnicity on NAFLD risk in relation to diabetes, data from the Indian group were omitted due to collinearity between these risk factors and Indian ethnicity i.e. the characteristics of the Indian NAFLD participants with diabetes perfectly predicted NAFLD.

These models were then compared to the null model (Table 5-9) using likelihood ratio testing. This shows that none of the above comparator models improve the null model, which did not include ethnicity (p=0.252, p=0.172). This means that there was no impact of ethnicity on the null model, despite apparent differences in BAT activity between ethnic groups.

5.4.7 BAT activity and diabetes

These data show that BAT activity is not directly associated with NAFLD. However, given the links between BAT activity and insulin sensitivity, analysis was undertaken to examine whether reduced BAT activity may be associated with presence of diabetes (as a metabolic precursor to NAFLD). Using the same logistic regression model (without the inclusion of diabetes as a variable), the influence of BAT activity on presence of diabetes was assessed. These data are presented in Table 5-13.

	Adjusted Odds Ratio (95% CI)	**P value**
BAT activity (ΔTrel (°C))	0.29 (0.03 – 3.03)	0.301
Age (years)	1.12 (1.07 – 1.18)	<0.001
BMI (kg/m^2)	1.22 (1.08 – 1.38)	0.001

Table 5-13. Adjusted logistic regression for diabetes risk

Abbreviations: BAT = brown adipose tissue, BMI = Body mass index
ΔTrel = change in relative temperature within the supraclavicular region of interest

This shows that BAT activity is not associated with diabetes risk. Risk of diabetes within these cohorts is influenced mainly by BMI and age. Again, through likelihood ratio testing, the inclusion of ethnicity as an interaction term did not alter the outcome of this model (ethnicity and ΔTrel p=0.642, ethnicity and BMI p=0.645).

5.5 Discussion

5.5.1 Principal findings

BAT activity was lower in native Indians compared to UK South Asians and UK Caucasians. This difference was significant even in the presence of a lower BMI, which has been shown to correlate with higher BAT activity. This difference appeared to be environmental rather than genetic, as there was no difference in BAT activity between control groups of different ethnicities who live in the same country (UK Caucasian and UK South Asian).

BAT activity was not directly associated with NAFLD risk. Any differences in BAT activity were related to increasing BMI (even within the native Indian population who have a lower BMI), which itself is known to increase NAFLD risk. The same can be concluded for BAT activity in relation to diabetes risk. Again, this highlights the Asian-paradox, whereby small changes in BMI within this ethnic group result in significantly higher risk of disease, whether cardiovascular or in relation to components of the metabolic syndrome[158].

There is paucity of studies that examine the difference in BAT activity between ethnic groups, and its impact on NAFLD risk. This makes it difficult to compare the

findings of this work to the current literature. When examining differences in BAT activity between ethnicities, this study is similar to the work done by the Dutch group who looked at differences in BAT activity between South Asians and Caucasians[179]. Their study utilised a group of participants whose heritage was South Asian, but who were born in Europe. Although they found that BAT volume was lower in those of South Asian descent, BAT activity did not differ between groups. Their finding is echoed in this work, with no difference seen in BAT activity between Caucasians and South Asians living within the UK. Where the Dutch group had a pool of participants of a particular heritage, of the same generation, this study recruited South Asian participants from different South Asian countries, with different cultures, of different generations and of different lengths of residence in the UK. Whilst this may result in heterogeneity in diet, lifestyle and genetics within the group, it is still important to have demonstrated that living in the UK appears to result in adaptation of BAT activity to levels similar to UK Caucasians.

The only other studies that examined BAT activity in relation to NAFLD were conducted through retrospective review of PET-CTs performed for clinical purposes[121, 122]. This makes it impossible to compare the findings of this work to those. They were not designed to examine active BAT, were undertaken in patients within the diseased state, and did not include pertinent data about other potential metabolic confounders such as presence of diabetes. It is however interesting that both such studies noted increased odds of NAFLD in those without active BAT. Whilst clinical PET-CT as a modality may under-report BAT activity, the CT component may be more accurate at identifying NAFLD. This means that the use of ultrasound in this work may in fact under-report NAFLD within the study group, reducing the power of its results.

5.5.2 Strengths and weaknesses of the study

This cross-sectional, multi-cohort case control study is, to the author's knowledge, the largest study of BAT activity utilising IRT as a technique – with other studies of this kind having sample sizes of between 8-58 participants[292, 299, 303, 307]. It may also be the largest clinical, prospective study of BAT activity in the literature overall, with studies that use PET-CT as an imaging modality being limited by cost and exposure to radiation [119, 120, 178, 179]. Whilst there have been numerous retrospective studies, methodological studies and studies of BAT activity and function in healthy volunteers, there is a paucity of data on the impact of BAT activity in the disease state. This is particularly pertinent when the majority of work done in the field of BAT focusses on its physiological links to whole body energy expenditure and potential role in obesity, with only a minority of observational or interventional studies involving patients[114, 116, 120]. This study allowed comparison of real-world BAT activity between different ethnicities, in different environments, and between participants with and without metabolic components of disease.

This pragmatic approach was enabled by using IRT as an imaging modality. Thermal imaging – in this case, video imaging – provided an enormous amount of data from a large number of participants, who did not require invasive procedures, or exposure to radiation. Its portable nature meant imaging could be undertaken wherever a climate-controlled room could be created. Within this study, individual images, selected at 5-second intervals, were used to create a detailed outline of the dynamic temperature changes that occur through activation of BAT. Many previous studies that used IRT for the assessment of BAT activity took single images at different time points[279, 291, 292], or at much longer intervals[293, 308]. The use of larger numbers of time points is becoming more prevalent[303], but with the ability of modern thermal cameras to capture images at 300 frames per second, there is scope to further improve the accuracy of data captured. IRT has now been validated against PET-CT, with ΔTrel being the measure that most closely correlated with BAT activity measured as ^{18}FDG uptake. The use of relative temperature changes in response to cold enables participants to become their own controls, standardising results and allowing direct comparison of individuals. Thus, through utilisation of a validated tool, under standardised conditions (as outlined by the Delphi Statement on thermal imaging[305]), capturing large amounts of data in different countries, this study has been able to demonstrate accurately the differences in BAT activity between ethnicities.

There is a wealth of evidence on the recruitment and increase in activity of BAT in response to prolonged cold acclimation, with some studies involving days or weeks of acclimation to lower temperatures[257, 258, 309]. It can therefore be surmised that any period of time living within a colder climate (Nottingham average annual temperature 10°C vs Trivandrum average annual temperature 30°C), is likely to result in higher BAT activity, as demonstrated by this work. These differences in BAT activity remain apparent, even when the internal study environment is maintained between cohorts, regardless of country. A potential concern about this work could be that the study environment, although consistent across cohorts, may be felt as cold to the native Indian participants, resulting in pre-activation of BAT at baseline. The comparable work from the Dutch group found that South Asians had a higher shiver temperature (temperature at which an individual starts shivering), despite living in the same country. They did however note that the temperature at which shivering ceased (at which BAT activity/non-shivering thermogenesis could be measured) was the same between groups[179]. Within this study, the relative temperature was higher in the Indian cohort, even at baseline, which could be thought to represent pre-activation of BAT in response to a “cold” acclimation environment. It is however more likely to be a result of reduced skin perfusion in response to the perceived “cold” rather than BAT activation, as the temperature of both the supraclavicular fossa and reference point were significantly lower in this group compared to the UK cohorts. If BAT had been active from the outset, the supraclavicular temperature would be expected to be higher. If anything, this strengthens the argument that native Indians have lower BAT activity, as even when

exposed to a perceived cold environment, the temperature of their supraclavicular fossa was lower than the UK cohorts. In addition, the fact that the native Indians were “warmer” from the outset may mean that they were cooled by the same stimulus to a lesser extent. One way this could have been identified would have been to use IRT to measure the arm temperature after the protocol was completed, comparing the degree to which it was cooled between cohorts.

Whilst pragmatic in its approach, with a study design focussed specifically on the differences between populations with components and sequelae of the metabolic syndrome, this study did involve a very heterogeneous group of participants, many of whom were on lipid-lowering or anti-hypertensive medications, even within the control groups (data presented in Chapter 4). Whilst those with conditions or who are on treatments that may influence BAT activity (e.g. thyroid dysfunction, beta-blockers etc.) were excluded, little is known about the impact of, for example, diabetic medications on BAT activity – particularly as BAT utilises glucose as a substrate.

Within this study, participants were sampled from similar places. The UK participants were recruited from a study designed to identify NAFLD within the community, and the Indian participants from a large NAFLD cohort. This was done with the aim to reduce variability between groups, resulting in similar rates of metabolic factors between cohorts. This was not entirely successful, as the UK Caucasian cohort was significantly older, were of higher BMI and had higher rates of diabetes. These confounding factors were however adjusted for within the logistic regression models, and despite these differences, the Indian cohort were shown to have lower rates of BAT activity.

5.5.3 Study implications and future work

This study showed that BAT activity is influenced principally by environment, not just by ethnicity. Living in a colder environment is likely to increase BAT activity, as supported by this work and published literature. Whilst BAT activity does not directly influence NAFLD or diabetes risk, the physiological role it plays suggests that it does have an impact on metabolic factors, such as insulin resistance, which in turn affect NAFLD risk. Further work is required to analyse the relationship between BAT activity and individual components of the metabolic syndrome (and their treatment), and to identify whether there is any potential for an intervention that increases BAT activity. This will require larger studies of patients of specific phenotypes, using more detailed measures of insulin resistance and lipid physiology, but in which IRT may prove a useful tool.

Whilst IRT has been validated against the current gold standard as a tool to measure BAT activity, this validation work was done using healthy, lean volunteers; and therefore its use is an important limitation to this study, which focussed on participants with complications of obesity. Whilst these data showed that subclavicular skinfold thickness (as a surrogate measure of neck adiposity) correlates closely to BMI (which was therefore included within the adjusted model), the

thermal data were not adjusted to incorporate factors such as anatomical differences in neck shape and size, tissue depth and skin perfusion. Further methodological studies are therefore required, to compare IRT to PET-CT in this group, and to refine the IRT technique for use with patients of different phenotypes.

Finally, further work is also required to understand the role of genetics on BAT activity. Whilst these findings indicate that lower BAT activity among Indians in India is related to their environment, genetic factors may determine cold perception and hence, thermoregulation through BAT function. As previously described, allele frequency of an SNP in the Transient Receptor Potential (TRP) cation channel subfamily member 8 gene with a central role in cold sensation varies significantly with the latitude[265], and through positive selection its frequency in the Eurasian population during the last 25,000 years may have increased. Further work using data from this study to explore whether TRPM8 alleles have a key role in determining BAT activity is underway (preliminary work has been accepted for poster presentation – see Relevant Publications section).

5.6 Conclusion

In reference to the hypotheses addressed within this chapter, this study demonstrated that NAFLD risk is not associated with decreased BAT activity directly. Although BAT activity was lower in native Indians, it was due to environment rather than ethnicity, as BAT activity was higher in South Asians who live in the UK. Whilst there is overwhelming evidence in the literature, as well as presented within this thesis (Chapter 2.) that people of Indian ethnicity are at increased risk of diabetes and NAFLD, this study has suggested that reduced BAT activity is not a key determinant. Further work is required to analyse the role of BAT activity on individual components of the metabolic syndrome and to understand the genetic factors that underpin the environmental differences demonstrated here.

CHAPTER 6: Conclusion

The purpose of this thesis was to add to the current evidence base regarding NAFLD risk in people of Indian origin, to estimate accurately the NAFLD prevalence across a large Indian district and confirm which factors influence disease risk. This chapter will summarise the main findings of this work, the extent to which the work confirmed or refuted the original hypotheses (Section 1.5), its research implications and potential future studies.

6.1 Principal findings

6.1.1 Increased prevalence of NAFLD within India

The prevalence of NAFLD is rising on a global scale, with a worldwide pooled prevalence estimated at 25.24%[1]. Data from large population-based studies in India are however lacking, with quoted rates of disease at a population level varying from 8.7-32% depending on study sample, setting and diagnostic tool used (Table 1-2). Through the work of Chapter 2, it was demonstrated that the prevalence of NAFLD within a large South Indian population is significantly higher than global estimates (49.8%), confirming the first hypothesis. Despite having higher prevalence, the severity of disease appeared to be comparable to the West[206], with 4.5% of those with NAFLD having evidence of advanced chronic liver disease. Those with evidence of fibrosis were more likely to be male, of higher BMI and have diabetes.

6.1.2 Commonality of NAFLD risk

In addition to establishing the increased prevalence of NAFLD within India, this work also demonstrated the commonality of NAFLD risk between India and the Western world, in line with the second hypothesis.

Obesity has long been an established risk factor for NAFLD, with a global pooled obesity prevalence amongst NAFLD patients of 51.34%[1]. Within Asian populations, obesity has been identified as the strongest risk factor for disease due to the so-called Asian Paradox – where small increases in BMI equate to much larger increases in disease risk among populations who are in general of lower BMI overall. This has resulted in the reclassification of BMI cut-offs for Asian populations[4]. The work of Chapter 2 confirms this link, with 64.9% of those with NAFLD being categorised as obese, and presence of obesity being the strongest risk factor for NAFLD within an Indian population (adjusted OR 2.81 2.15-3.68, $p<0.001$). The impact of obesity on NAFLD risk appears to be the same for both native and migrant Indian populations, as shown by the data from Chapter 4. Here it was demonstrated that, regardless of ethnicity, presence of obesity increased NAFLD risk 4-fold (adjusted OR 4.50, 1.92-10.56, $p=0.001$).

In addition to obesity, insulin resistance and diabetes are also closely linked to NAFLD. The global pooled prevalence of diabetes in those with NAFLD is 22.5%[1], which is lower than the rates within the Indian population sampled in Chapter 3 (34.1%). Prevalence of diabetes within India is much higher than the Western world (>2 million vs. 31,000 mean annual increment[5]) and as such is another strong contributing factor to the increased NAFLD prevalence within this population. Interestingly, diabetes did not appear to impact NAFLD risk within the work of Chapter 4, which may be due to sampling error within the native Indian group, who had equal numbers of diabetics in case and control groups.

6.1.3 Dietary habits may influence NAFLD risk in India

Changing dietary habits and physical activity levels over the last century are also thought to be contributing to the rising burden of NAFLD worldwide. Within India, there has been a shift away from consumption of coarse grains towards consumption of rice and wheat[217]. More pertinently, there has also been an increase in fat intake, often in the form of highly saturated ghee[218]. Within Chapter 3, it was demonstrated that, in addition to the traditional clinical risk factors outlined above, increased dietary fat (in particular saturated fat) is independently associated with increased NAFLD risk within this Indian population (adjusted OR 1.02 1.00-1.03, p=0.019). This confirmed in part, the hypothesis relating to dietary composition and NAFLD risk. Contrary to evidence in the literature however, dietary carbohydrate intake does not appear to influence NAFLD risk within this Indian population

In relation to the hypothesis that acculturation can alter NAFLD risk profile, the work presented in Chapter 4 suggested a degree of dietary acculturation in UK-migrant South Asians, who consumed a diet of similar total calorie and macronutrient composition as UK-Caucasians, and who appeared to eat less total calories and carbohydrate than native Indians. Dietary habits did not however appear to be associated with NAFLD risk in either native or migrant Indian groups, with the strongest association being with presence of obesity regardless of ethnicity.

6.1.4 BAT activity does not influence NAFLD risk

There has been increasing interest in BAT activity and its link to obesity and metabolic disease including NAFLD. The hypothesis that reduced BAT activity in people of Indian ethnicity may contribute to their adverse metabolic profile and increased NAFLD riskwas examined through the work of Chapter 5. It was shown that, whilst BAT activity was lower in native Indian compared to UK-migrant South Asian and UK Caucasian cohorts, the difference appeared to be due to environment.

Reduced BAT activity was not directly associated with increased NAFLD risk, and that any differences in BAT activity were related to increasing BMI, which is itself a risk factor for NAFLD.

6.2 Implications and future work

6.2.1 Screening programmes

This body of work demonstrates that NAFLD prevalence is significantly higher in India than both national and global estimates. It is already well documented that people of Indian ethnicity have much higher rates of diabetes, have a predisposition to central adiposity and are at increased risk of obesity-related morbidity and mortality at a lower BMI than other ethnic groups. NAFLD screening programmes could therefore focus the limited resources on those populations who are at highest risk – in particular the urban, Southern districts who have the highest documented rates of obesity[201]. The Wilson-Junger criteria for a valid screening programme include the need for the condition to be an important health problem, with a natural history that is known, that has an acceptable and suitable test that can detect disease early and can enable early intervention that is more beneficial than management of late disease[310]. Through population-level sampling, this work has proven that large-scale screening for NAFLD can be achieved using ultrasound. Consideration does however need to be taken to ensure accurate reproducibility of NAFLD diagnosis using this modality. Inter/intra-observer variability could have been improved throughout this work if scans were performed by a single operator, with findings confirmed by a second reviewer. FibroScan®, with the ability to measure controlled attenuated parameter score (a measure of steatosis), may be an even better alternative. Its portable nature, ability to identify and stratify disease and provide opportunity for point-of-care intervention/education within the community make it particularly suitable, as a large proportion of the Indian population resides within the rural regions where access to healthcare is limited. Future work should ensure validity of FibroScan® data, through collection of 10 valid readings, aiming for an IQR that is less than 30% of the median reading[207]. This was not achieved within the work of Chapter 2, reducing confidence in data relating to severity of disease within this population.

In addition to screening for NAFLD within at-risk populations in India, it may also be worthwhile including Indian (or South Asian) ethnicity as a factor when considering risk of NAFLD within the UK. This is particularly pertinent as use of Asia-specific BMI cut-offs within the UK is not always widespread and may result in UK-Asians being incorrectly categorised as normal BMI and thus low risk for disease. Again, suitable screening programmes may best be instigated within areas of high ethnic diversity, where it has already been shown that particular South Asian populations have higher rates of NAFLD than their Caucasian neighbours[228].

6.2.2 Dietary intervention

Having identified that increased intake of fat is independently associated with increased NAFLD risk within India (particularly saturated fat), this could become a focus for interventional strategies for both disease prevention and treatment. Interventional strategies that focus on particular dietary macronutrients are not

pragmatic at a population level. Basic food group analysis appeared to show that the higher fat intake came from increased intake of edible oils and animal protein; however further work is required to identify which dietary patterns are linked to NAFLD risk in this population. Alterations in the types of grains eaten, cooking oils used and consumption of meat are more realistic alternatives for future interventional strategies. Indeed, pilot studies of this kind are already appearing in the literature (canola oil vs ghee[311], red rice vs white rice ClinicalTrials.gov NCT03844165).

6.2.3 Indian NAFLD genome-wide association study

The work of this thesis has focussed mainly on the impact of lifestyle on the increased susceptibility to NAFLD in people of Indian ethnicity. In addition to these "environmental" factors, genetics are also well known to play a role. Through GWAS and candidate gene studies, multiple genetic variants have been shown to influence NAFLD risk, some of which have also been shown to impact NAFLD risk in Indian populations. To date however, there has not been a NAFLD GWAS performed in an Indian population. During the creation of the Trivandrum NAFLD cohort, whole blood aliquots were taken from each participant, from which DNA was extracted. Future work within this cohort will therefore include a GWAS to confirm/identify variants that affect NAFLD risk within this group, and will enable examination of the interaction between lifestyle and genetic factors within this population.

6.3 Conclusion

In conclusion, prevalence of NAFLD is significantly higher in India than current global estimates. There is a commonality of risk between India and the West – namely presence of obesity and diabetes. Whilst there may be a degree of acculturation to a Western lifestyle with UK-migration, NAFLD risk factors are unchanged across ethnic groups. Through detailed dietary analysis, it was shown that an increased intake of saturated fat as cooking oil and animal protein might be a contributory factor to increased NAFLD risk in India, which could be the focus for further research and future interventional studies. Finally, having shown that BAT activity is lower in native Indians, this is due to differences in environment, and that reduced BAT activity is not directly linked to NAFLD risk, but is associated with increased BMI, which is itself a risk factor for NAFLD.

REFERENCES

1. Younossi, Z.M., et al., *Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes.* Hepatology, 2016. **64**(1): p. 73-84.
2. Younossi, Z., et al., *Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention.* Nature Reviews Gastroenterology & Hepatology, 2018. **15**(1): p. 11-20.
3. Allen, A.M., et al., *Nonalcoholic fatty liver disease incidence and impact on metabolic burden and death: A 20 year-community study.* Hepatology, 2018. **67**(5): p. 1726-1736.
4. WHO expert consultation, *Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies.* Lancet, 2004. **363**(9403): p. 157-63.
5. Whiting, D.R., et al., *IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030.* Diabetes Res Clin Pract, 2011. **94**(3): p. 311-21.
6. Popkin, B.M., *The nutrition transition and obesity in the developing world.* J Nutr, 2001. **131**(3): p. 871s-873s.
7. Mohan, V., et al., *Prevalence of non-alcoholic fatty liver disease in urban south Indians in relation to different grades of glucose intolerance and metabolic syndrome.* Diabetes Res Clin Pract, 2009. **84**(1): p. 84-91.
8. Majumdar, A., et al., *Prevalence of nonalcoholic fatty liver disease in an adult population in a rural community of Haryana, India.* Indian J Public Health, 2016. **60**(1): p. 26-33.
9. National Institute for Health and Care Excellence, *Non-alcoholic fatty lvier disease (NAFLD): assessment and management.* https://www.nice.org.uk/guidance/ng49, 2016.
10. Angulo, P., *Nonalcoholic fatty liver disease.* N Engl J Med, 2002. **346**(16): p. 1221-31.
11. Kleiner, D.E., et al., *Design and validation of a histological scoring system for nonalcoholic fatty liver disease.* Hepatology, 2005. **41**(6): p. 1313-21.
12. Bugianesi, E., et al., *Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma.* Gastroenterology, 2002. **123**(1): p. 134-40.
13. Feldstein, A.E., et al., *The natural history of non-alcoholic fatty liver disease in children: a follow-up study for up to 20 years.* Gut, 2009. **58**(11): p. 1538-44.
14. Marchesini, G., et al., *Nonalcoholic fatty liver disease: a feature of the metabolic syndrome.* Diabetes, 2001. **50**(8): p. 1844-50.
15. Eckel, R.H., S.M. Grundy, and P.Z. Zimmet, *The metabolic syndrome.* Lancet, 2005. **365**(9468): p. 1415-28.
16. Flegal, K.M., et al., *Prevalence and trends in obesity among US adults, 1999-2008.* Jama, 2010. **303**(3): p. 235-41.
17. Misra, A. and L. Khurana, *Obesity and the metabolic syndrome in developing countries.* J Clin Endocrinol Metab, 2008. **93**(11 Suppl 1): p. S9-30.
18. Fan, J.G., S.U. Kim, and V.W. Wong, *New trends on obesity and NAFLD in Asia.* J Hepatol, 2017. **67**(4): p. 862-873.
19. Vernon, G., A. Baranova, and Z.M. Younossi, *Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults.* Aliment Pharmacol Ther, 2011. **34**(3): p. 274-85.

20. Sayiner, M., et al., *Epidemiology of Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis in the United States and the Rest of the World.* Clin Liver Dis, 2016. **20**(2): p. 205-14.
21. Chitturi, S., V.W. Wong, and G. Farrell, *Nonalcoholic fatty liver in Asia: Firmly entrenched and rapidly gaining ground.* J Gastroenterol Hepatol, 2011. **26 Suppl 1**: p. 163-72.
22. Fan, J.G., *Epidemiology of alcoholic and nonalcoholic fatty liver disease in China.* J Gastroenterol Hepatol, 2013. **28 Suppl 1**: p. 11-7.
23. Armstrong, M.J., et al., *Presence and severity of non-alcoholic fatty liver disease in a large prospective primary care cohort.* J Hepatol, 2012. **56**(1): p. 234-40.
24. Skelly, M.M., P.D. James, and S.D. Ryder, *Findings on liver biopsy to investigate abnormal liver function tests in the absence of diagnostic serology.* Journal of Hepatology, 2001. **35**(2): p. 195-199.
25. Wilman, H.R., et al., *Characterisation of liver fat in the UK Biobank cohort.* PLOS ONE, 2017. **12**(2): p. e0172921.
26. Williams, R., et al., *Addressing liver disease in the UK: a blueprint for attaining excellence in health care and reducing premature mortality from lifestyle issues of excess consumption of alcohol, obesity, and viral hepatitis.* The Lancet. **384**(9958): p. 1953-1997.
27. Estes, C., et al., *Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016-2030.* J Hepatol, 2018. **69**(4): p. 896-904.
28. McPherson, S., et al., *Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease.* Gut, 2010. **59**(9): p. 1265.
29. Ong, J.P., A. Pitts, and Z.M. Younossi, *Increased overall mortality and liver-related mortality in non-alcoholic fatty liver disease.* J Hepatol, 2008. **49**(4): p. 608-12.
30. Allen, A.M., et al., *Healthcare Cost and Utilization in Nonalcoholic Fatty Liver Disease: Real-World Data From a Large U.S. Claims Database.* Hepatology, 2018. **68**(6): p. 2230-2238.
31. Younossi, Z.M., et al., *The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe.* Hepatology, 2016. **64**(5): p. 1577-1586.
32. Weinstein, A.A., et al., *Depression in patients with nonalcoholic fatty liver disease and chronic viral hepatitis B and C.* Psychosomatics, 2011. **52**(2): p. 127-32.
33. Golabi, P., et al., *Non-alcoholic Fatty Liver Disease (NAFLD) is associated with impairment of Health Related Quality of Life (HRQOL).* Health Qual Life Outcomes, 2016. **14**: p. 18.
34. Tapper, E.B. and M. Lai, *Weight loss results in significant improvements in quality of life for patients with nonalcoholic fatty liver disease: A prospective cohort study.* Hepatology (Baltimore, Md.), 2016. **63**(4): p. 1184-1189.
35. Ratziu, V., et al., *Sampling variability of liver biopsy in nonalcoholic fatty liver disease.* Gastroenterology, 2005. **128**(7): p. 1898-906.
36. Saadeh, S., et al., *The utility of radiological imaging in nonalcoholic fatty liver disease.* Gastroenterology, 2002. **123**(3): p. 745-50.
37. Dasarathy, S., et al., *Validity of real time ultrasound in the diagnosis of hepatic steatosis: a prospective study.* J Hepatol, 2009. **51**(6): p. 1061-7.
38. Ekstedt, M., et al., *Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up.* Hepatology, 2015. **61**(5): p. 1547-54.

39. Angulo, P., et al., *Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease.* Gastroenterology, 2015. **149**(2): p. 389-97.e10.
40. West, J., et al., *Risk of hepatocellular carcinoma among individuals with different aetiologies of cirrhosis: a population-based cohort study.* Alimentary Pharmacology & Therapeutics, 2017. **45**(7): p. 983-990.
41. Jiang, W., et al., *Diagnostic accuracy of point shear wave elastography and transient elastography for staging hepatic fibrosis in patients with non-alcoholic fatty liver disease: a meta-analysis.* BMJ open, 2018. **8**(8): p. e021787-e021787.
42. Guha, I.N., et al., *Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers.* Hepatology, 2008. **47**(2): p. 455-60.
43. EASL-EASD-EASO, *Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease.* Journal of Hepatology, 2016. **64**(6): p. 1388-1402.
44. Coppell, K.J., et al., *Obesity and the extent of liver damage among adult New Zealanders: findings from a national survey.* Obes Sci Pract, 2015. **1**(2): p. 67-77.
45. Ascha, M.S., et al., *The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis.* Hepatology, 2010. **51**(6): p. 1972-8.
46. Younossi, Z., et al., *Global Perspectives on Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis.* Hepatology, 2019. **69**(6): p. 2672-2682.
47. Eguchi, Y., et al., *Prevalence and associated metabolic factors of nonalcoholic fatty liver disease in the general population from 2009 to 2010 in Japan: a multicenter large retrospective study.* J Gastroenterol, 2012. **47**(5): p. 586-95.
48. Caballería, L., et al., *Prevalence and factors associated with the presence of nonalcoholic fatty liver disease in an adult population in Spain.* Eur J Gastroenterol Hepatol, 2010. **22**(1): p. 24-32.
49. Adams, L.A., et al., *The natural history of nonalcoholic fatty liver disease: a population-based cohort study.* Gastroenterology, 2005. **129**(1): p. 113-21.
50. Shulman, G.I., *Ectopic Fat in Insulin Resistance, Dyslipidemia, and Cardiometabolic Disease.* New England Journal of Medicine, 2014. **371**(12): p. 1131-1141.
51. Marra, F. and C. Bertolani, *Adipokines in liver diseases.* Hepatology, 2009. **50**(3): p. 957-69.
52. Kabir, M., et al., *Molecular evidence supporting the portal theory: a causative link between visceral adiposity and hepatic insulin resistance.* Am J Physiol Endocrinol Metab, 2005. **288**(2): p. E454-61.
53. Park, B.J., et al., *Visceral adipose tissue area is an independent risk factor for hepatic steatosis.* J Gastroenterol Hepatol, 2008. **23**(6): p. 900-7.
54. Eguchi, Y., et al., *Visceral fat accumulation and insulin resistance are important factors in nonalcoholic fatty liver disease.* Journal of Gastroenterology, 2006. **41**(5): p. 462-469.
55. Guerrero, R., et al., *Ethnic differences in hepatic steatosis: an insulin resistance paradox?* Hepatology, 2009. **49**(3): p. 791-801.
56. van der Poorten, D., et al., *Visceral fat: a key mediator of steatohepatitis in metabolic liver disease.* Hepatology, 2008. **48**(2): p. 449-57.
57. Dyson, J.K., Q.M. Anstee, and S. McPherson, *Non-alcoholic fatty liver disease: a practical approach to diagnosis and staging.* Frontline Gastroenterol, 2014. **5**(3): p. 211-218.
58. Hu, F.B., R.M. van Dam, and S. Liu, *Diet and risk of Type II diabetes: the role of types of fat and carbohydrate.* Diabetologia, 2001. **44**(7): p. 805-17.

59. Musso, G., et al., *Dietary habits and their relations to insulin resistance and postprandial lipemia in nonalcoholic steatohepatitis.* Hepatology, 2003. **37**(4): p. 909-16.
60. Heath, R., et al., *Selective partitioning of dietary fatty acids into the VLDL TG pool in the early postprandial period.* Journal of lipid research, 2003. **44**(11): p. 2065-2072.
61. Fernández, M.I., et al., *Steatosis and Collagen Content in Experimental Liver Cirrhosis Are Affected by Dietary Monounsaturated and Polyunsaturated Fatty Acids.* Scandinavian Journal of Gastroenterology, 1997. **32**(4): p. 350-356.
62. Kim, S.P., et al., *Primacy of hepatic insulin resistance in the development of the metabolic syndrome induced by an isocaloric moderate-fat diet in the dog.* Diabetes, 2003. **52**(10): p. 2453-60.
63. Westerbacka, J., et al., *Dietary fat content modifies liver fat in overweight nondiabetic subjects.* J Clin Endocrinol Metab, 2005. **90**(5): p. 2804-9.
64. Lau, C., et al., *Dietary Glycemic Index, Glycemic Load, Fiber, Simple Sugars, and Insulin Resistance.* The Inter99 study, 2005. **28**(6): p. 1397-1403.
65. Herman, R.H., D. Zakim, and F.B. Stifel, *Effect of diet on lipid metabolism in experimental animals and man.* Fed Proc, 1970. **29**(3): p. 1302-7.
66. Goodyear, L.J. and B.B. Kahn, *Exercise, glucose transport, and insulin sensitivity.* Annu Rev Med, 1998. **49**: p. 235-61.
67. Miller, W.J., W.M. Sherman, and J.L. Ivy, *Effect of strength training on glucose tolerance and post-glucose insulin response.* Med Sci Sports Exerc, 1984. **16**(6): p. 539-43.
68. Hannukainen, J.C., et al., *Increased physical activity decreases hepatic free fatty acid uptake: a study in human monozygotic twins.* J Physiol, 2007. **578**(Pt 1): p. 347-58.
69. Dixon, J.B., et al., *Nonalcoholic fatty liver disease: Improvement in liver histological analysis with weight loss.* Hepatology, 2004. **39**(6): p. 1647-54.
70. Shah, K., et al., *Diet and exercise interventions reduce intrahepatic fat content and improve insulin sensitivity in obese older adults.* Obesity (Silver Spring), 2009. **17**(12): p. 2162-8.
71. Andersen, T., et al., *Hepatic effects of dietary weight loss in morbidly obese subjects.* J Hepatol, 1991. **12**(2): p. 224-9.
72. Toshimitsu, K., et al., *Dietary habits and nutrient intake in non-alcoholic steatohepatitis.* Nutrition, 2007. **23**(1): p. 46-52.
73. Solga, S., et al., *Dietary composition and nonalcoholic fatty liver disease.* Dig Dis Sci, 2004. **49**(10): p. 1578-83.
74. Machado, M.V., et al., *Blood oxidative stress markers in non-alcoholic steatohepatitis and how it correlates with diet.* Scand J Gastroenterol, 2008. **43**(1): p. 95-102.
75. Li, Y.H., et al., *Efficacy of poly-unsaturated fatty acid therapy on patients with nonalcoholic steatohepatitis.* World J Gastroenterol, 2015. **21**(22): p. 7008-13.
76. Abdelmalek, M.F., et al., *Increased fructose consumption is associated with fibrosis severity in patients with nonalcoholic fatty liver disease.* Hepatology, 2010. **51**(6): p. 1961-71.
77. Zelber-Sagi, S., et al., *Long term nutritional intake and the risk for non-alcoholic fatty liver disease (NAFLD): a population based study.* J Hepatol, 2007. **47**(5): p. 711-7.
78. Cortez-Pinto, H., et al., *How different is the dietary pattern in non-alcoholic steatohepatitis patients?* Clin Nutr, 2006. **25**(5): p. 816-23.
79. Oddy, W.H., et al., *The Western dietary pattern is prospectively associated with nonalcoholic fatty liver disease in adolescence.* Am J Gastroenterol, 2013. **108**(5): p. 778-85.
80. Yang, C.Q., et al., *Dietary Patterns Modulate the Risk of Non-Alcoholic Fatty Liver Disease in Chinese Adults.* Nutrients, 2015. **7**(6): p. 4778-91.

81. Chung, G.E., et al., *Dietary patterns are associated with the prevalence of nonalcoholic fatty liver disease in Korean adults.* Nutrition, 2019. **62**: p. 32-38.
82. Lim, J.S., et al., *The role of fructose in the pathogenesis of NAFLD and the metabolic syndrome.* Nat Rev Gastroenterol Hepatol, 2010. **7**(5): p. 251-64.
83. DiNicolantonio, J.J., A.M. Subramonian, and J.H. O'Keefe, *Added fructose as a principal driver of non-alcoholic fatty liver disease: a public health crisis.* Open Heart, 2017. **4**(2): p. e000631.
84. Capanni, M., et al., *Prolonged n-3 polyunsaturated fatty acid supplementation ameliorates hepatic steatosis in patients with non-alcoholic fatty liver disease: a pilot study.* Aliment Pharmacol Ther, 2006. **23**(8): p. 1143-51.
85. Tanaka, N., et al., *Highly purified eicosapentaenoic acid treatment improves nonalcoholic steatohepatitis.* J Clin Gastroenterol, 2008. **42**(4): p. 413-8.
86. Trovato, F.M., et al., *Neglected features of lifestyle: Their relevance in non-alcoholic fatty liver disease.* World journal of hepatology, 2016. **8**(33): p. 1459-1465.
87. Kontogianni, M.D., et al., *Adherence to the Mediterranean diet is associated with the severity of non-alcoholic fatty liver disease.* Clin Nutr, 2014. **33**(4): p. 678-83.
88. Perseghin, G., et al., *Habitual physical activity is associated with intrahepatic fat content in humans.* Diabetes Care, 2007. **30**(3): p. 683-8.
89. Zelber-Sagi, S., et al., *Role of leisure-time physical activity in nonalcoholic fatty liver disease: a population-based study.* Hepatology, 2008. **48**(6): p. 1791-8.
90. Kistler, K.D., et al., *Physical activity recommendations, exercise intensity, and histological severity of nonalcoholic fatty liver disease.* Am J Gastroenterol, 2011. **106**(3): p. 460-8; quiz 469.
91. Church, T.S., et al., *Association of cardiorespiratory fitness, body mass index, and waist circumference to nonalcoholic fatty liver disease.* Gastroenterology, 2006. **130**(7): p. 2023-30.
92. Hallsworth, K., et al., *Non-alcoholic fatty liver disease is associated with higher levels of objectively measured sedentary behaviour and lower levels of physical activity than matched healthy controls.* Frontline Gastroenterol, 2015. **6**(1): p. 44-51.
93. Sung, K.C., et al., *Effect of exercise on the development of new fatty liver and the resolution of existing fatty liver.* J Hepatol, 2016. **65**(4): p. 791-797.
94. Cannon, B. and J. Nedergaard, *Brown adipose tissue: function and physiological significance.* Physiological Reviews, 2004. **84**(1): p. 277-359.
95. Nedergaard, J., T. Bengtsson, and B. Cannon, *Unexpected evidence for active brown adipose tissue in adult humans.* American Journal of Physiology - Endocrinology And Metabolism, 2007. **293**(2): p. E444-E452.
96. Gesner, C., *Medici tigurini historiae animalium liber i de quadrupedibus uiuiparis*. 1551.
97. Tseng, Y.H., et al., *New role of bone morphogenetic protein 7 in brown adipogenesis and energy expenditure.* Nature, 2008. **454**(7207): p. 1000-4.
98. Heaton, J.M., *The distribution of brown adipose tissue in the human.* Journal of Anatomy, 1972. **112**(Pt 1): p. 35-39.
99. Leitner, B.P., et al., *Mapping of human brown adipose tissue in lean and obese young men.* Proceedings of the National Academy of Sciences, 2017.
100. Smith, R.E. and B.A. Horwitz, *Brown fat and thermogenesis.* Physiological Reviews, 1969. **49**(2): p. 330-425.
101. Hany, T.F., et al., *Brown adipose tissue: a factor to consider in symmetrical tracer uptake in the neck and upper chest region.* Eur J Nucl Med Mol Imaging, 2002. **29**(10): p. 1393-8.
102. Bronnikov, G., et al., *beta1 to beta3 switch in control of cyclic adenosine monophosphate during brown adipocyte development explains distinct beta-*

adrenoceptor subtype mediation of proliferation and differentiation. Endocrinology, 1999. **140**(9): p. 4185-97.
103. Madden, C.J., E.P. Santos da Conceicao, and S.F. Morrison, *Vagal afferent activation decreases brown adipose tissue (BAT) sympathetic nerve activity and BAT thermogenesis.* Temperature (Austin), 2017. **4**(1): p. 89-96.
104. Bronnikov, G., J. Houstěk, and J. Nedergaard, *Beta-adrenergic, cAMP-mediated stimulation of proliferation of brown fat cells in primary culture. Mediation via beta 1 but not via beta 3 adrenoceptors.* J Biol Chem, 1992. **267**(3): p. 2006-13.
105. Jacobsson, A., et al., *Mitochondrial uncoupling protein from mouse brown fat. Molecular cloning, genetic mapping, and mRNA expression.* Journal of Biological Chemistry, 1985. **260**(30): p. 16250-16254.
106. Nechad, M., J. Nedergaard, and B. Cannon, *Noradrenergic stimulation of mitochondriogenesis in brown adipocytes differentiating in culture.* American Journal of Physiology - Cell Physiology, 1987. **253**(6): p. C889-C894.
107. Lindquist, J.M., et al., *β3- and α1-Adrenergic Erk1/2 Activation Is Src- but Not Gi-mediated in Brown Adipocytes.* Journal of Biological Chemistry, 2000. **275**(30): p. 22670-22677.
108. Rothwell, N.J. and M.J. Stock, *Similarities between cold- and diet-induced thermogenesis in the rat.* Canadian Journal of Physiology and Pharmacology, 1980. **58**(7): p. 842-848.
109. Klaus, S., *Adipose tissue as a regulator of energy balance.* Curr Drug Targets, 2004. **5**(3): p. 241-50.
110. Feldmann, H.M., et al., *UCP1 ablation induces obesity and abolishes diet-induced thermogenesis in mice exempt from thermal stress by living at thermoneutrality.* Cell Metab, 2009. **9**(2): p. 203-9.
111. Giles, D.A., et al., *Thermoneutral housing exacerbates nonalcoholic fatty liver disease in mice and allows for sex-independent disease modeling.* Nat Med, 2017. **23**(7): p. 829-838.
112. Poekes, L., et al., *Defective adaptive thermogenesis contributes to metabolic syndrome and liver steatosis in obese mice.* Clinical Science, 2017. **131**(4): p. 285-296.
113. Ouellet, V., et al., *Outdoor temperature, age, sex, body mass index, and diabetic status determine the prevalence, mass, and glucose-uptake activity of 18F-FDG-detected BAT in humans.* J Clin Endocrinol Metab, 2011. **96**(1): p. 192-9.
114. Koksharova, E., et al., *The Relationship Between Brown Adipose Tissue Content in Supraclavicular Fat Depots and Insulin Sensitivity in Patients with Type 2 Diabetes Mellitus and Prediabetes.* Diabetes Technology & Therapeutics, 2017. **19**(2): p. 96-102.
115. Bartelt, A., et al., *Brown adipose tissue activity controls triglyceride clearance.* Nat Med, 2011. **17**(2): p. 200-5.
116. Chondronikola, M., et al., *Brown adipose tissue improves whole-body glucose homeostasis and insulin sensitivity in humans.* Diabetes, 2014. **63**(12): p. 4089-99.
117. Hanssen, M.J.W., et al., *Short-term cold acclimation improves insulin sensitivity in patients with type 2 diabetes mellitus.* Nature Medicine, 2015. **21**(8): p. 863-865.
118. Meex, R.C., et al., *Restoration of muscle mitochondrial function and metabolic flexibility in type 2 diabetes by exercise training is paralleled by increased myocellular fat storage and improved insulin sensitivity.* Diabetes, 2010. **59**(3): p. 572-9.
119. Hanssen, M.J.W., et al., *Short-term Cold Acclimation Recruits Brown Adipose Tissue in Obese Humans.* Diabetes, 2016. **65**(5): p. 1179-1189.
120. Orava, J., et al., *Blunted metabolic responses to cold and insulin stimulation in brown adipose tissue of obese humans.* Obesity (Silver Spring), 2013. **21**(11): p. 2279-87.

121. Yilmaz, Y., et al., *Association between the presence of brown adipose tissue and non-alcoholic fatty liver disease in adult humans.* Aliment Pharmacol Ther, 2011. **34**(3): p. 318-23.
122. Zhang, Z., et al., *The prevalence and predictors of active brown adipose tissue in Chinese adults.* 2014. **170**(3): p. 359.
123. Romeo, S., et al., *Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease.* Nat Genet, 2008. **40**(12): p. 1461-5.
124. Yuan, X., et al., *Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes.* Am J Hum Genet, 2008. **83**(4): p. 520-8.
125. Chalasani, N., et al., *Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease.* Gastroenterology, 2010. **139**(5): p. 1567-76, 1576.e1-6.
126. Speliotes, E.K., et al., *Genome-Wide Association Analysis Identifies Variants Associated with Nonalcoholic Fatty Liver Disease That Have Distinct Effects on Metabolic Traits.* PLOS Genetics, 2011. **7**(3): p. e1001324.
127. Chambers, J.C., et al., *Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma.* Nat Genet, 2011. **43**(11): p. 1131-8.
128. Kawaguchi, T., et al., *Genetic Polymorphisms of the Human PNPLA3 Gene Are Strongly Associated with Severity of Non-Alcoholic Fatty Liver Disease in Japanese.* PLoS ONE, 2012. **7**(6): p. e38322.
129. Anstee, Q.M., R. Darlay, and J.B. Leathart, *Genome-wide association analysis confirms importance of PNPLA3 and identifies novel variants associated with histologically progressive steatohepatitis in NAFLD.* Hepatology, 2012. **56**: p. 265A-266A.
130. Feitosa, M.F., et al., *The ERLIN1-CHUK-CWF19L1 gene cluster influences liver fat deposition and hepatic inflammation in the NHLBI Family Heart Study.* Atherosclerosis, 2013. **228**(1): p. 175-180.
131. Kozlitina, J., et al., *Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease.* Nat Genet, 2014. **46**(4): p. 352-6.
132. DiStefano, J.K., et al., *Genome-wide analysis of hepatic lipid content in extreme obesity.* Acta Diabetol, 2015. **52**(2): p. 373-82.
133. Sookoian, S. and C.J. Pirola, *Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease.* Hepatology, 2011. **53**(6): p. 1883-94.
134. Kersten, S., B. Desvergne, and W. Wahli, *Roles of PPARs in health and disease.* Nature, 2000. **405**(6785): p. 421-4.
135. van Dijk, K.W., et al., *The role and mode of action of apolipoproteins CIII and AV: synergistic actors in triglyceride metabolism?* Curr Opin Lipidol, 2004. **15**(3): p. 239-46.
136. Zain, S.M., Z. Mohamed, and R. Mohamed, *Common variant in the glucokinase regulatory gene rs780094 and risk of nonalcoholic fatty liver disease: a meta-analysis.* J Gastroenterol Hepatol, 2015. **30**(1): p. 21-7.
137. Bambha, K., et al., *Ethnicity and nonalcoholic fatty liver disease.* Hepatology, 2012. **55**(3): p. 769-80.
138. Browning, J.D., et al., *Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity.* Hepatology, 2004. **40**(6): p. 1387-95.
139. Kallwitz, E.R., et al., *Ethnicity and Nonalcoholic Fatty Liver Disease in an Obesity Clinic: The Impact of Triglycerides.* Digestive Diseases and Sciences, 2008. **53**(5): p. 1358-1363.

140. Schwimmer, J.B., et al., *Influence of gender, race, and ethnicity on suspected fatty liver in obese adolescents.* Pediatrics, 2005. **115**(5): p. e561-5.
141. Anurag, L., et al., *Non-alcoholic fatty liver disease prevalence and associated risk factors--A study from rural sector of Maharashtra.* Trop Gastroenterol, 2015. **36**(1): p. 25-30.
142. Das, K., et al., *Nonobese population in a developing country has a high prevalence of nonalcoholic fatty liver and significant liver disease.* Hepatology, 2010. **51**(5): p. 1593-602.
143. Prashanth, M., et al., *Prevalence of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus.* J Assoc Physicians India, 2009. **57**: p. 205-10.
144. Amarapurkar, D., et al., *Prevalence of non-alcoholic fatty liver disease: population based study.* Ann Hepatol, 2007. **6**(3): p. 161-3.
145. Chan, W.-k., et al., *Non-alcoholic fatty liver disease in a young multiracial Asian population: a worrying ethnic predilection in Malay and Indian males.* Hepatology international., 2014. **8**(1): p. 121-127.
146. Singh, S.P., et al., *Prevalence of nonalcoholic fatty liver disease in coastal eastern India: a preliminary ultrasonographic survey.* Trop Gastroenterol, 2004. **25**(2): p. 76-9.
147. Praveenraj, P., et al., *Prevalence and Predictors of Non-Alcoholic Fatty Liver Disease in Morbidly Obese South Indian Patients Undergoing Bariatric Surgery.* Obes Surg, 2015. **25**(11): p. 2078-87.
148. Gupte, P., et al., *Non-alcoholic steatohepatitis in type 2 diabetes mellitus.* J Gastroenterol Hepatol, 2004. **19**(8): p. 854-8.
149. Agarwal, A.K., et al., *Prevalence of non-alcoholic fatty liver disease and its correlation with coronary risk factors in patients with type 2 diabetes.* J Assoc Physicians India, 2011. **59**: p. 351-4.
150. Hajong, R., et al., *Incidence of nonalcoholic fatty liver disease in patients undergoing laparoscopic cholecystectomy.* Journal of family medicine and primary care, 2018. **7**(6): p. 1375-1378.
151. Duseja, A., et al., *High prevalence of non-alcoholic fatty liver disease among healthy male blood donors of urban India.* JGH Open, 2019. **3**: p. 1-7.
152. Uchil, D., et al., *Non-alcoholic fatty liver disease (NAFLD)--the hepatic component of metabolic syndrome.* J Assoc Physicians India, 2009. **57**: p. 201-4.
153. Bajaj, S., et al., *A case-control study on insulin resistance, metabolic co-variates & prediction score in non-alcoholic fatty liver disease.* Indian J Med Res, 2009. **129**(3): p. 285-92.
154. GBD 2015 Risk Factors Collaborators, *Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015.* Lancet, 2016. **388**(10053): p. 1659-1724.
155. Wasir, J.S. and A. Misra, *The Metabolic Syndrome in Asian Indians: Impact of Nutritional and Socio-economic Transition in India.* Metab Syndr Relat Disord, 2004. **2**(1): p. 14-23.
156. http://www.ncdrisc.org/index.html.
157. Lear, S.A., et al., *Modification of the relationship between simple anthropometric indices and risk factors by ethnic background.* Metabolism, 2003. **52**(10): p. 1295-301.
158. Wong, R.J. and A. Ahmed, *Obesity and non-alcoholic fatty liver disease: Disparate associations among Asian populations.* World Journal of Hepatology, 2014. **6**(5): p. 263-273.

159. Fan, J.G., et al., *What are the risk factors and settings for non-alcoholic fatty liver disease in Asia-Pacific?* J Gastroenterol Hepatol, 2007. **22**(6): p. 794-800.
160. Wei, J.L., et al., *Prevalence and Severity of Nonalcoholic Fatty Liver Disease in Non-Obese Patients: A Population Study Using Proton-Magnetic Resonance Spectroscopy.* Am J Gastroenterol, 2015. **110**(9): p. 1306-14; quiz 1315.
161. Younossi, Z.M., et al., *Nonalcoholic fatty liver disease in lean individuals in the United States.* Medicine (Baltimore), 2012. **91**(6): p. 319-27.
162. Singh, D.K., et al., *Independent predictors of steatohepatitis and fibrosis in Asian Indian patients with non-alcoholic steatohepatitis.* Dig Dis Sci, 2008. **53**(7): p. 1967-76.
163. Madan, K., et al., *Non-alcoholic fatty liver disease may not be a severe disease at presentation among Asian Indians.* World Journal of Gastroenterology : WJG, 2006. **12**(21): p. 3400-3405.
164. Agarwal, S.R., et al., *Clinical, biochemical and histological profile of nonalcoholic steatohepatitis.* Indian J Gastroenterol, 2001. **20**(5): p. 183-6.
165. Sookoian, S. and C.J. Pirola, *Systematic review with meta-analysis: risk factors for non-alcoholic fatty liver disease suggest a shared altered metabolic and cardiovascular profile between lean and obese patients.* Aliment Pharmacol Ther, 2017. **46**(2): p. 85-95.
166. Cruz, A.C.D., et al., *Characteristics and Long-Term Prognosis of Lean Patients With Nonalcoholic Fatty Liver Disease.* Gastroenterology, 2014. **146**(5): p. S-909.
167. Raji, A., et al., *Body fat distribution and insulin resistance in healthy Asian Indians and Caucasians.* J Clin Endocrinol Metab, 2001. **86**(11): p. 5366-71.
168. Hughes, K., et al., *Central obesity, insulin resistance, syndrome X, lipoprotein(a), and cardiovascular risk in Indians, Malays, and Chinese in Singapore.* J Epidemiol Community Health, 1997. **51**(4): p. 394-9.
169. McKeigue, P.M., B. Shah, and M.G. Marmot, *Relation of central obesity and insulin resistance with high diabetes prevalence and cardiovascular risk in South Asians.* Lancet, 1991. **337**(8738): p. 382-6.
170. Banerji, M.A., et al., *Body composition, visceral fat, leptin, and insulin resistance in Asian Indian men.* J Clin Endocrinol Metab, 1999. **84**(1): p. 137-44.
171. Mohan, V., et al., *Intra-urban differences in the prevalence of the metabolic syndrome in southern India -- the Chennai Urban Population Study (CUPS No. 4).* Diabet Med, 2001. **18**(4): p. 280-7.
172. Diez-Roux, A.V., et al., *Prevalence and social correlates of cardiovascular disease risk factors in Harlem.* Am J Public Health, 1999. **89**(3): p. 302-7.
173. Petersen, K.F., et al., *Increased prevalence of insulin resistance and nonalcoholic fatty liver disease in Asian-Indian men.* Proc Natl Acad Sci U S A, 2006. **103**(48): p. 18273-7.
174. Council on Clinical Cardiology, *Exercise and physical activity in the prevention and treatment of atherosclerotic cardiovascular disease: a statement from the Council on Clinical Cardiology (Subcommittee on Exercise, Rehabilitation, and Prevention) and the Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity).* Circulation., 2003. **107**(24): p. 3109-16.
175. Sudha, V., G. Radhika, and V. Mohan, *Current dietary trends in the management of diabetes.* Indian J Med Res, 2004. **120**(1): p. 4-8.
176. Sathiaraj, E., et al., *A case-control study on nutritional risk factors in non-alcoholic fatty liver disease in Indian population.* Eur J Clin Nutr, 2011. **65**(4): p. 533-537.
177. Singh, S.P., et al., *Risk Factors Associated With Non-Alcoholic Fatty Liver Disease in Indians: A Case–Control Study.* Journal of Clinical and Experimental Hepatology, 2015. **5**(4): p. 295-302.

178. Admiraal, W.M., et al., *Cold-induced activity of brown adipose tissue in young lean men of South-Asian and European origin.* Diabetologia, 2013. **56**(10): p. 2231-7.
179. Bakker, L.E., et al., *Brown adipose tissue volume in healthy lean south Asian adults compared with white Caucasians: a prospective, case-controlled observational study.* Lancet Diabetes Endocrinol, 2014. **2**(3): p. 210-7.
180. Zain, S.M., et al., *A multi-ethnic study of a PNPLA3 gene variant and its association with disease severity in non-alcoholic fatty liver disease.* Hum Genet, 2012. **131**(7): p. 1145-52.
181. Kanth, V.V.R., et al., *Pooled genetic analysis in ultrasound measured non-alcoholic fatty liver disease in Indian subjects: A pilot study.* World Journal of Hepatology, 2014. **6**(6): p. 435-442.
182. Bhatt, S.P., et al., *Association of peroxisome proliferator activated receptor-gamma gene with non-alcoholic fatty liver disease in Asian Indians residing in north India.* Gene, 2013. **512**(1): p. 143-7.
183. Petersen , K.F., et al., *Apolipoprotein C3 Gene Variants in Nonalcoholic Fatty Liver Disease.* New England Journal of Medicine, 2010. **362**(12): p. 1082-1089.
184. Puppala, J., et al., *Apolipoprotein C3 gene polymorphisms in Southern Indian patients with nonalcoholic fatty liver disease.* Indian J Gastroenterol, 2014. **33**(6): p. 524-9.
185. Tan, H.L., et al., *Association of glucokinase regulatory gene polymorphisms with risk and severity of non-alcoholic fatty liver disease: an interaction study with adiponutrin gene.* J Gastroenterol, 2014. **49**(6): p. 1056-64.
186. Bale, G., et al., *Regional differences in genetic susceptibility to non-alcoholic liver disease in two distinct Indian ethnicities.* World J Hepatol, 2017. **9**(26): p. 1101-1107.
187. Pingali, P., *Westernization of Asian diets and the transformation of food systems: Implications for research and policy.* Food Policy, 2007. **32**(3): p. 281-298.
188. Brown, N.G., et al., *Advanced ERCP techniques for the extraction of complex biliary stones: a single referral center's 12-year experience.* Scand J Gastroenterol, 2018. **53**(5): p. 626-631.
189. Anstee, Q.M. and C.P. Day, *The genetics of NAFLD.* Nat Rev Gastroenterol Hepatol, 2013. **10**(11): p. 645-55.
190. Kahali, B., B. Halligan, and E.K. Speliotes, *Insights from Genome-Wide Association Analyses of Nonalcoholic Fatty Liver Disease.* Semin Liver Dis, 2015. **35**(4): p. 375-91.
191. K T Shenoy, e.a. *Prevalence and risk factors for non-alcoholic fatty liver in a transition society with high incidence of diabetes mellitus*. in *The 15th Annual Conference of the INASAL*. 2007. Vellore.
192. Liberatos, P., B.G. Link, and J.L. Kelsey, *The measurement of social class in epidemiology.* Epidemiol Rev, 1988. **10**: p. 87-121.
193. Davey Smith, G., et al., *Education and occupational social class: which is the more important indicator of mortality risk?* J Epidemiol Community Health, 1998. **52**(3): p. 153-60.
194. Misra, A., et al., *Consensus statement for diagnosis of obesity, abdominal obesity and the metabolic syndrome for Asian Indians and recommendations for physical activity, medical and surgical management.* J Assoc Physicians India, 2009. **57**: p. 163-70.
195. Das, K., et al., *"Normal" liver stiffness measure (LSM) values are higher in both lean and obese individuals: a population-based study from a developing country.* Hepatology, 2012. **55**(2): p. 584-93.
196. Fraquelli, M., et al., *Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease.* Gut, 2007. **56**(7): p. 968-73.
197. Altman, D.G. and J.M. Bland, *Statistics notes: the normal distribution.* BMJ (Clinical research ed.), 1995. **310**(6975): p. 298-298.

198. Altman, D.G. and J.M. Bland, *Statistics notes: the normal distribution.* Bmj, 1995. **310**(6975): p. 298.
199. Bellentani, S., et al., *Epidemiology of non-alcoholic fatty liver disease.* Dig Dis, 2010. **28**(1): p. 155-61.
200. Abarca-Gómez, L., et al., *Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults.* The Lancet, 2017. **390**(10113): p. 2627-2642.
201. The Ministry of Health and Family Welfare, *National Family Health Survey (NFHS-4).* 2016.
202. Pourshams, A., et al., *Prevalence and etiology of persistently elevated alanine aminotransferase levels in healthy Iranian blood donors.* J Gastroenterol Hepatol, 2005. **20**(2): p. 229-33.
203. Wong, V.W., et al., *Metabolic and histological features of non-alcoholic fatty liver disease patients with different serum alanine aminotransferase levels.* Aliment Pharmacol Ther, 2009. **29**(4): p. 387-96.
204. Mofrad, P., et al., *Clinical and histologic spectrum of nonalcoholic fatty liver disease associated with normal ALT values.* Hepatology, 2003. **37**(6): p. 1286-92.
205. Fracanzani, A.L., et al., *Risk of severe liver disease in nonalcoholic fatty liver disease with normal aminotransferase levels: a role for insulin resistance and diabetes.* Hepatology, 2008. **48**(3): p. 792-8.
206. Chalmers, J., et al., *Development and implementation of a commissioned pathway for the identification and stratification of liver disease in the community.* Frontline Gastroenterology, 2019: p. flgastro-2019-101177.
207. Lucidarme, D., et al., *Factors of accuracy of transient elastography (fibroscan) for the diagnosis of liver fibrosis in chronic hepatitis C.* Hepatology, 2009. **49**(4): p. 1083-9.
208. Chiu, S., et al., *Effect of fructose on markers of non-alcoholic fatty liver disease (NAFLD): a systematic review and meta-analysis of controlled feeding trials.* Eur J Clin Nutr, 2014. **68**(4): p. 416-23.
209. Simon, T.G., et al., *Physical Activity Compared to Adiposity and Risk of Liver-Related Mortality: Results from Two Prospective, Nationwide Cohorts.* J Hepatol, 2020.
210. Public Health England, *Government Dietary Recommendations*. 2016: UK.
211. Vijay, A., et al., *The Evaluation and Use of a Food Frequency Questionnaire Among the Population in Trivandrum, South Kerala, India.* Nutrients, 2020. **12**(2): p. 383.
212. C. Gopalan, B.V.R.S., S.C. Balasubramanian, *Nutritive Value of Indian Foods*. 2011: Indian Council of Medical Research.
213. Cox, D.R. and N. Wermuth, *A Comment on the Coefficient of Determination for Binary Responses.* The American Statistician, 1992. **46**(1): p. 1-4.
214. Kroke, A., et al., *Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods.* Am J Clin Nutr, 1999. **70**(4): p. 439-47.
215. Kipnis, V., et al., *Bias in dietary-report instruments and its implications for nutritional epidemiology.* Public Health Nutr, 2002. **5**(6a): p. 915-23.
216. Lear, S.A., et al., *The effect of physical activity on mortality and cardiovascular disease in 130 000 people from 17 high-income, middle-income, and low-income countries: the PURE study.* Lancet, 2017. **390**(10113): p. 2643-2654.
217. DeFries, R., et al., *Impact of Historical Changes in Coarse Cereals Consumption in India on Micronutrient Intake and Anemia Prevalence.* Food and Nutrition Bulletin, 2018. **39**(3): p. 377-392.

218. Popkin, B.M., et al., *Trends in diet, nutritional status, and diet-related noncommunicable diseases in China and India: the economic costs of the nutrition transition.* Nutr Rev, 2001. **59**(12): p. 379-90.
219. Popkin, B.M. and P. Gordon-Larsen, *The nutrition transition: worldwide obesity dynamics and their determinants.* Int J Obes Relat Metab Disord, 2004. **28 Suppl 3**: p. S2-9.
220. Amarapurkar, D.N., et al., *How common is non-alcoholic fatty liver disease in the Asia-Pacific region and are there local differences?* J Gastroenterol Hepatol, 2007. **22**(6): p. 788-93.
221. Bennett, J., *The SAGE Encyclopedia of Intercultural Competence*. 2015, SAGE Publications, Inc.: Thousand Oaks, California.
222. Patel, J.V., et al., *Impact of migration on coronary heart disease risk factors: comparison of Gujaratis in Britain and their contemporaries in villages of origin in India.* Atherosclerosis, 2006. **185**(2): p. 297-306.
223. Holmboe-Ottesen, G. and M. Wandel, *Changes in dietary habits after migration and consequences for health: a focus on South Asians in Europe.* Food & Nutrition Research, 2012. **56**(1): p. 18891.
224. Garduno-Diaz, S.D. and S. Khokhar, *South Asian dietary patterns and their association with risk factors for the metabolic syndrome.* J Hum Nutr Diet, 2013. **26**(2): p. 145-55.
225. Delavari, M., et al., *Acculturation and obesity among migrant populations in high income countries – a systematic review.* BMC Public Health, 2013. **13**(1): p. 458.
226. Harding, S., *Mortality of migrants from the Indian subcontinent to England and Wales: effect of duration of residence.* Epidemiology, 2003. **14**(3): p. 287-92.
227. Balakrishnan, M., et al., *Acculturation and Nonalcoholic Fatty Liver Disease Risk Among Hispanics of Mexican Origin: Findings From the National Health and Nutrition Examination Survey.* Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association, 2017. **15**(2): p. 310-312.
228. Alazawi, W., et al., *Ethnicity and the diagnosis gap in liver disease: a population-based study.* Br J Gen Pract, 2014. **64**(628): p. e694-702.
229. Neukam, K., et al., *High prevalence of non-alcoholic fatty liver disease (NAFLD) among Gujarati Indians in North London: a population-based study.* Clinical Lipidology, 2017. **12**(1): p. 33-39.
230. Matthews, D.R., et al., *Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man.* Diabetologia, 1985. **28**(7): p. 412-9.
231. Lansang, M.C., G.H. Williams, and J.S. Carroll, *Correlation between the glucose clamp technique and the homeostasis model assessment in hypertension.* Am J Hypertens, 2001. **14**(1): p. 51-3.
232. Aller, R., et al., *Influence of insulin resistance and adipokines in the grade of steatosis of nonalcoholic fatty liver disease.* Dig Dis Sci, 2008. **53**(4): p. 1088-92.
233. Hill, R.J. and P.S. Davies, *The validity of self-reported energy intake as determined using the doubly labelled water technique.* Br J Nutr, 2001. **85**(4): p. 415-30.
234. Thielecke, F., et al., *Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods.* The American Journal of Clinical Nutrition, 1999. **70**(4): p. 439-447.

235. Caspersen, C.J., K.E. Powell, and G.M. Christenson, *Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research.* Public Health Rep, 1985. **100**(2): p. 126-31.
236. Prince, S.A., et al., *A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review.* Int J Behav Nutr Phys Act, 2008. **5**: p. 56.
237. Armstrong, T. and F. Bull, *Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ).* Journal of Public Health, 2006. **14**(2): p. 66-70.
238. Herrmann, S.D., et al., *Validity and Reliability of the Global Physical Activity Questionnaire (GPAQ).* Measurement in Physical Education and Exercise Science, 2013. **17**(3): p. 221-235.
239. Public Health England, *Government Dietary Recommendations.* 2016.
240. National Institute for Health and Care Excellence, *BMI: preventing ill health and premature death in black, Asian and other minority ethnic groups (Public health guideline PH46).* https://www.nice.org.uk/guidance/ph46/chapter/1-Recommendations, 2013.
241. Grundy, S.M., et al., *Definition of metabolic syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition.* Circulation, 2004. **109**(3): p. 433-8.
242. Xia, M.-F., et al., *Standardized ultrasound hepatic/renal ratio and hepatic attenuation rate to quantify liver fat content: an improvement method.* Obesity (Silver Spring, Md.), 2012. **20**(2): p. 444-452.
243. Bland, J.M. and D.G. Altman, *Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet, 1986. **1**(8476): p. 307-10.
244. Chalmers, J., et al., *Cohort profile: the Trivandrum non-alcoholic fatty liver disease (NAFLD) cohort.* BMJ Open, 2019. **9**(5): p. e027244.
245. Shen, H., et al., *Association between caffeine consumption and nonalcoholic fatty liver disease: a systemic review and meta-analysis.* Therapeutic advances in gastroenterology, 2016. **9**(1): p. 113-120.
246. Office for National Statistics. *www.ons.gov.uk*. 2011.
247. Yoneshiro, T., et al., *Brown adipose tissue, whole-body energy expenditure, and thermogenesis in healthy adult men.* Obesity (Silver Spring), 2011. **19**(1): p. 13-6.
248. van Marken Lichtenbelt, W.D., et al., *Cold-activated brown adipose tissue in healthy men.* N Engl J Med, 2009. **360**(15): p. 1500-8.
249. Ouellet, V., et al., *Brown adipose tissue oxidative metabolism contributes to energy expenditure during acute cold exposure in humans.* Journal of Clinical Investigation, 2012. **122**(2): p. 545-552.
250. Cannon, B. and J. Nedergaard, *Brown adipose tissue: function and physiological significance.* Physiol Rev, 2004. **84**(1): p. 277-359.
251. Stanford, K.I., et al., *Brown adipose tissue regulates glucose homeostasis and insulin sensitivity.* The Journal of Clinical Investigation, 2013. **123**(1): p. 215-223.
252. Cohade, C., K.A. Mourtzikos, and R.L. Wahl, *"USA-Fat": prevalence is related to ambient outdoor temperature-evaluation with 18F-FDG PET/CT.* J Nucl Med, 2003. **44**(8): p. 1267-70.
253. Au-Yong, I.T., et al., *Brown adipose tissue and seasonal variation in humans.* Diabetes, 2009. **58**(11): p. 2583-7.
254. Cronin, C.G., et al., *Brown fat at PET/CT: correlation with patient characteristics.* Radiology, 2012. **263**(3): p. 836-42.

255. Chen, K.Y., et al., *Brown Adipose Reporting Criteria in Imaging STudies (BARCIST 1.0): Recommendations for Standardized FDG-PET/CT Experiments in Humans.* Cell metabolism, 2016. **24**(2): p. 210-222.
256. Kim, S., et al., *Temporal relation between temperature change and FDG uptake in brown adipose tissue.* Eur J Nucl Med Mol Imaging, 2008. **35**(5): p. 984-9.
257. Davis, T.R., *Chamber cold acclimatization in man.* J Appl Physiol, 1961. **16**: p. 1011-5.
258. van der Lans, A.A., et al., *Cold acclimation recruits human brown fat and increases nonshivering thermogenesis.* J Clin Invest, 2013. **123**(8): p. 3395-403.
259. Sellayah, D., *The Impact of Early Human Migration on Brown Adipose Tissue Evolution and Its Relevance to the Modern Obesity Pandemic.* Journal of the Endocrine Society, 2018. **3**(2): p. 372-386.
260. Nishimura, T., et al., *Experimental evidence reveals the UCP1 genotype changes the oxygen consumption attributed to non-shivering thermogenesis in humans.* Scientific reports, 2017. **7**(1): p. 5570-5570.
261. Flouris, A.D., et al., *Role of UCP1 Gene Variants in Interethnic Differences in the Development of Cardio-Metabolic Diseases.* Frontiers in Genetics, 2017. **8**(7).
262. Masuo, K. and G.W. Lambert, *Relationships of adrenoceptor polymorphisms with obesity.* J Obes, 2011. **2011**: p. 609485.
263. Frayling, T.M., et al., *A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity.* Science, 2007. **316**(5826): p. 889-94.
264. Claussnitzer, M., et al., *FTO Obesity Variant Circuitry and Adipocyte Browning in Humans.* New England Journal of Medicine, 2015. **373**(10): p. 895-907.
265. Key, F.M., et al., *Human local adaptation of the TRPM8 cold receptor along a latitudinal cline.* PLOS Genetics, 2018. **14**(5): p. e1007298.
266. Madsen, L., et al., *UCP1 induction during recruitment of brown adipocytes in white adipose tissue is dependent on cyclooxygenase activity.* PLoS ONE [Electronic Resource], 2010. **5**(6).
267. Mitchell, J.R., et al., *Regulation of expression of the lipoprotein lipase gene in brown adipose tissue.* American Journal of Physiology, 1992. **263**(3 Pt 1): p. E500-6.
268. Vergnes, L., et al., *Adipocyte Browning and Higher Mitochondrial Function in Periadrenal But Not SC Fat in Pheochromocytoma.* Journal of Clinical Endocrinology and Metabolism, 2016. **101**(11): p. 4440-4448.
269. Porter, C., et al., *Human and mouse brown adipose tissue mitochondria have comparable UCP1 function.* Cell metabolism, 2016. **24**(2): p. 246-255.
270. Wang, Q., et al., *Brown adipose tissue in humans is activated by elevated plasma catecholamines levels and is inversely related to central obesity.* PLoS ONE, 2011. **6**(6).
271. Ukropec, J., et al., *UCP1-independent thermogenesis in white adipose tissue of cold-acclimated Ucp1-/-mice.* Journal of Biological Chemistry, 2006. **281**(42): p. 31894-31908.
272. Meyer, C.W., et al., *Adaptive thermogenesis and thermal conductance in wild-type and UCP1-KO mice.* American Journal of Physiology. Regulatory, Integrative and Comparative Physiology, 2010. **299**(5): p. R1396-406.
273. Hany, T.F., et al., *Brown adipose tissue: a factor to consider in symmetrical tracer uptake in the neck and upper chest region.* European Journal of Nuclear Medicine & Molecular Imaging, 2002. **29**(10): p. 1393-8.
274. Sacks, H.S., *The importance of brown adipose tissue.* New England Journal of Medicine, 2009. **361**(4): p. author reply 418-20.
275. Celi, F.S., *Brown adipose tissue--when it pays to be inefficient.* New England Journal of Medicine, 2009. **360**(15): p. 1553-6.

276. Virtanen, K.A., et al., *Functional brown adipose tissue in healthy adults.* New England Journal of Medicine, 2009. **360**(15): p. 1518-25.
277. Cypess, A.M., et al., *Identification and importance of brown adipose tissue in adult humans.* New England Journal of Medicine, 2009. **360**(15): p. 1509-17.
278. van Marken Lichtenbelt, W.D., et al., *Cold-activated brown adipose tissue in healthy men.* New England Journal of Medicine, 2009. **360**(15): p. 1500-8.
279. Gatidis, S., et al., *Is it possible to detect activated brown adipose tissue in humans using single-time-point infrared thermography under thermoneutral conditions? Impact of BMI and subcutaneous adipose tissue thickness.* PLoS ONE, 2016. **11**(3): p. e0151152.
280. Hankir, M.K., et al., *Dissociation between brown adipose tissue 18F-FDG uptake and thermogenesis in uncoupling protein 1 deficient mice.* Journal of Nuclear Medicine, 2017.
281. Khedoe, P.P., et al., *Brown adipose tissue takes up plasma triglycerides mostly after lipolysis.* J Lipid Res., 2015. **56**.
282. Avram, A.S., M.M. Avram, and W.D. James, *Subcutaneous fat in normal and diseased states: 2. Anatomy and physiology of white and brown adipose tissue.* J Am Acad Dermatol, 2005. **53**(4): p. 671-83.
283. Gifford, A., et al., *Characterizing active and inactive brown adipose tissue in adult humans using PET-CT and MR imaging.* American journal of physiology. Endocrinology and metabolism, 2016. **311**(1): p. E95-E104.
284. Chen, Y.C., et al., *Measurement of human brown adipose tissue volume and activity using anatomic MR imaging and functional MR imaging.* J Nucl Med, 2013. **54**(9): p. 1584-7.
285. Raiko, J., et al., *Brown adipose tissue triglyceride content is associated with decreased insulin sensitivity, independently of age and obesity.* Diabetes Obes Metab, 2015. **17**(5): p. 516-9.
286. Planck M, *The theory of heat radiation*. 1914, Philadelphia, US: P. Bllakiston's Son & Co.
287. Rothwell, N.J. and M.J. Stock, *A role for brown adipose tissue in diet-induced thermogenesis.* Nature, 1979. **281**(5726): p. 31-5.
288. Carter, E.A., et al., *Association of heat production with 18F-FDG accumulation in murine brown adipose tissue after stress.* J Nucl Med, 2011. **52**(10): p. 1616-20.
289. Crane, J.D., et al., *A standardized infrared imaging technique that specifically detects UCP1-mediated thermogenesis in vivo.* Mol Metab, 2014. **3**(4): p. 490-4.
290. Symonds, M.E., et al., *Thermal Imaging to Assess Age-Related Changes of Skin Temperature within the Supraclavicular Region Co-Locating with Brown Adipose Tissue in Healthy Children.* The Journal of Pediatrics, 2012. **161**(5): p. 892-898.
291. Lee, P., et al., *Hot fat in a cool man: infrared thermography and brown adipose tissue.* Diabetes Obes Metab, 2011. **13**(1): p. 92-3.
292. Jang, C., et al., *Infrared thermography in the detection of brown adipose tissue in humans.* Physiological reports, 2014. **2**(11): p. e12167.
293. Salem, V., et al., *Glucagon increases energy expenditure independently of brown adipose tissue activation in humans.* Diabetes, Obesity and Metabolism, 2016. **18**(1): p. 72-81.
294. Veicsteinas, A., G. Ferretti, and D.W. Rennie, *Superficial shell insulation in resting and exercising men in cold water.* J Appl Physiol Respir Environ Exerc Physiol, 1982. **52**(6): p. 1557-64.
295. Salamunes, A.C.C., A.M.W. Stadnik, and E.B. Neves, *The effect of body fat percentage and body fat distribution on skin surface temperature with infrared thermography.* J Therm Biol, 2017. **66**: p. 1-9.

296. Yoneshiro, T., et al., *Age-related decrease in cold-activated brown adipose tissue and accumulation of body fat in healthy humans.* Obesity (Silver Spring), 2011. **19**(9): p. 1755-60.
297. Robinson, L., et al., *Body Mass Index as a Determinant of Brown Adipose Tissue Function in Healthy Children.* The Journal of Pediatrics, 2014. **164**(2): p. 318-322.e1.
298. Saito, M., et al., *High Incidence of Metabolically Active Brown Adipose Tissue in Healthy Adult Humans: Effects of Cold Exposure and Adiposity.* Diabetes, 2009. **58**(7): p. 1526-1531.
299. Hartwig, V., et al., *Multimodal Imaging for the Detection of Brown Adipose Tissue Activation in Women: A Pilot Study Using NIRS and Infrared Thermography.* Journal of Healthcare Engineering, 2017. **2017**: p. 6.
300. Kong, W., et al., *A role for arcuate cocaine and amphetamine-regulated transcript in hyperphagia, thermogenesis, and cold adaptation.* Faseb j, 2003. **17**(12): p. 1688-90.
301. Hibi, M., et al., *Brown adipose tissue is involved in diet-induced thermogenesis and whole-body fat utilization in healthy humans.* International Journal of Obesity, 2016.
302. Robinson, L.J., et al., *Brown adipose tissue activation as measured by infrared thermography by mild anticipatory psychological stress in lean healthy females.* Experimental Physiology, 2016. **101**(4): p. 549-557.
303. Law, J.M., et al., *Thermal imaging is a non-invasive alternative to PET-CT for measurement of brown adipose tissue activity in humans.* J Nucl Med, 2017.
304. Bahler, L., et al., *Hot heads & cool bodies: The conundrums of human brown adipose tissue (BAT) activity research.* Eur J Intern Med, 2017. **40**: p. 26-29.
305. Moreira, D.G., et al., *Thermographic imaging in sports and exercise medicine: A Delphi study and consensus statement on the measurement of human skin temperature.* J Therm Biol, 2017. **69**: p. 155-162.
306. Hardy, J.D. and T.W. Oppel, *Studies in temperature sensation. IV. The stimulation of cold sensation by radiation.* The Journal of clinical investigation, 1938. **17**(6): p. 771-778.
307. El Hadi, H., et al., *Infrared thermography for indirect assessment of activation of brown adipose tissue in lean and obese male subjects.* Physiol Meas, 2016. **37**(12): p. N118-n128.
308. Velickovic, K., et al., *Caffeine exposure induces browning features in adipose tissue in vitro and in vivo.* Scientific Reports, 2019. **9**(1): p. 9104.
309. Chen, K.Y., et al., *Brown Fat Activation Mediates Cold-Induced Thermogenesis in Adult Humans in Response to a Mild Decrease in Ambient Temperature.* The Journal of Clinical Endocrinology & Metabolism, 2013. **98**(7): p. E1218-E1223.
310. Wilson, J.M.G., G. Jungner, and O. World Health, *Principles and practice of screening for disease* 1968, World Health Organization: Geneva.
311. Nigam, P., et al., *Effect of a 6-month intervention with cooking oils containing a high concentration of monounsaturated fatty acids (olive and canola oils) compared with control oil in male Asian Indians with nonalcoholic fatty liver disease.* Diabetes Technol Ther, 2014. **16**(4): p. 255-61.

APPENDICES

Appendix 1. Indian FFQ item list and average item portion intake per day (n=2047)

FFQ Item	Food Recipe	Average portion per day (n=2047)
1	Coffee with milk with sugar	0.203
2	Coffee without milk with sugar	0.208
3	Coffee without sugar with milk	1.423
4	Coffee without milk without sugar	0.168
5	Milk with sugar	0.576
6	Milk without sugar	0.894
7	Tea with milk with sugar	2.557
8	Tea without milk with sugar	1.802
9	Tea without sugar with milk	2.323
10	Tea without sugar without milk	1.482
11	Ada (Jaggery and Coconut)	0.024
12	Ada (Sugar and coconut)	0.074
13	Ada (Coconut)	0.037
14	Appam (with coconut)	0.496
15	Appam (without coconut)	0.475
16	Bread Toast	0.127
17	Chappathi	0.781
18	Dosai (Maida)	0.072
19	Dosai (P.R.+B.G)	0.847
20	Dosai (R.R.+B.G)	0.824
21	Dosai (R.R.+Coconut)	0.355
22	Dosai (R.R. + P.R. + B.G.)	0.916

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
23	Dosai (Wheat flour)	0.317
24	Dosai (Raw rice)	0.434
25	Iddli (P.B. + R.R. + B.G.)	0.464
26	Iddli (R.R.+B.G.)	0.892
27	Idiappam	0.321
28	Kozhukatta (Coconut+jaggery)	0.075
29	Kozhukatta (Coconut+sugar)	0.069
30	Kozhukatta (Coconut)	0.097
31	Oothappam	0.088
32	Oratti	0.104
33	Poori (Maida)	0.170
34	Poori (Wheat flour)	0.138
35	Porotta	0.065
36	Puttu (Raw rice)	0.243
37	Puttu (Wheat flour)	0.174
38	Puttu (Semolina)	0.040
39	Sandwitch	0.094
40	Uppuma(rice)	0.211
41	Uppuma (Semolina)	0.056
42	Uppuma (Wheat)	0.156
43	Biriyani (chicken)	0.103
44	Biriyani (Beef)	0.047
45	Biriyani(Egg)	0.029
46	Biriyani (mutton)	0.045
47	Biriyani(veg)	0.036

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
48	Curd Rice	0.040
49	Fried rice	0.046
50	Fried rice (Beef)	0.055
51	Fried rice(Chicken)	0.058
52	Fried rice(Egg)	0.107
53	Fried rice (fish)	0.028
54	Fried rice(mutton)	0.030
55	Lemon rice	0.423
56	Pazhamkanji	4.567
57	Rice (P.B.)	2.013
58	Rice(Raw)	0.142
59	Rice gruel	0.100
60	Tomato rice	0.143
61	Wheat gruel	0.066
62	Amaranth (red) curry	0.092
63	Amaranth (green) curry	0.221
64	Amaranth (red) curry with dhal	0.106
65	Amaranth (green) curry with dhal	0.022
66	Ash gourd curry	0.217
67	Aviyal with curd	0.324
68	Aviyal without curd	0.030
69	Beet root Pachadi	0.109
70	Bengal gram dhal	0.031
71	Bitter gourd kichadi	0.035
72	Bilimbi Pachadi	0.077

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
73	Butter milk curry with coconut	0.048
74	Butter milk curry without coconut	0.138
75	Buttermilk	0.117
76	Cabbage Aviyal	0.057
77	Cauliflower curry	0.076
78	Chammanthi podi	0.082
79	Cherkkurmani's curry	0.212
80	Coconut Chutney (dry)	0.390
81	Coconut Chutney (wet)	0.110
82	Coconut ozhichucurry	0.040
83	Cow pea curry	0.070
84	Cucumber kichadi	0.035
85	Cucumber pachadi	0.116
86	Curd	0.054
87	Drum stick curry	0.089
88	Drum stick curry with dhal	0.046
89	Drum stick leaves curry	0.041
90	Erisseri with pulses	0.028
91	Erisseri without pulses	0.036
92	Garlic chutney	0.116
93	Green gram curry	0.042
94	Green gram dhal curry	0.020
95	Jack Aviyal	0.020
96	Jack fruit curry	0.064
97	Kuttu curry	0.158

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
98	Ladies finger kichadi	0.171
99	Mixed vegetable Kuruma	0.141
100	Mulaku chutney	0.020
101	Olan	0.099
102	Onion curry	0.058
103	Palak Curry	0.029
104	Papaya curry	0.087
105	Peas curry	0.040
106	Plantain stem pachady	0.148
107	Potato curry	0.132
108	Pulissery	0.091
109	Pulin curry	0.178
110	Rasam with dhal	0.242
111	Rasam without dhal	0.051
112	Raw banana curry	0.640
113	Red gram dhal curry	0.305
114	Sambar (with dhal)	0.159
115	Sambar (without dhal)	0.029
116	Soya chunks curry	0.033
117	Stew	0.036
118	Tapioca Aviyal	0.109
119	Tapioca curry with coconut	0.182
120	Tapioca curry without coconut	0.066
121	Thiyyal	0.190
122	Tomato chutney	0.129

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
123	Tomato curry	0.065
124	Tomato Kichadi	0.091
125	Tomato Onion curry	0.045
126	Tomato Pachadi	0.016
127	Agathi Thoran	0.101
128	Amaranth (green) thoran	0.073
129	Amaranth (red) thoran with dhal jack seed	0.030
130	Amaranth thoran (red)	0.046
131	Banana peel thoran	0.052
132	Banana peel thoran (dhal)	0.109
133	Beans thoran	0.219
134	Beet root thoran	0.156
135	Bengal gram dhal thoran	0.059
136	Bitter gourd thoran	0.135
137	Black gram dhal thoran	0.065
138	Brinjal thoran	0.097
139	Cabbage thoran	0.097
140	Carrot thoran	0.060
141	Cauliflower thoran	0.032
142	Chekkurumanis thoran	0.108
143	Cluster beans thoran	0.064
144	Cow pea pod thoran	0.075
145	Cow pea thoran	0.038
146	Drumstick leaves thoran	0.048

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
147	Drumstick thoran	0.037
148	Drumstick jack seed thoran	0.058
149	Green gram thoran	0.085
150	Horse gram	0.021
151	Jack seed thoran	0.018
152	Jack tender thoran	0.051
153	Kovai thoran	0.164
154	Ladies finger thoran	0.062
155	Onion thoran	0.034
156	Onion stalk thoran	0.030
157	Papaya thoran	0.025
158	Plantain flower thoran	0.024
159	Plantain flower thoran(dhal)	0.022
160	Plantain stem thoran	0.019
161	Plantain stem thoran with dhal	0.069
162	Potato thoran	0.033
163	Pappad thoran	0.045
164	Raw banana Thoran	0.040
165	Snake gourd thoran	0.030
166	Tapioca thoran	0.089
167	Thakara leaf thoran	0.026
168	Yam thoran	0.107
169	Yam leaf thoran	0.022
170	Amaranth green (sauté)	0.072
171	Amaranth red (sauté)	0.032

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
172	Beans sauté	0.054
173	Beet root sauté	0.036
174	Bengal gram chundal	0.077
175	Bitter gourd sauté	0.061
176	Brinjal sauté	0.075
177	Cabbage sauté	0.051
178	Carrot sauté	0.049
179	Chekkurmanis sauté	0.028
180	Cluster beans sauté	0.020
181	Cow pea pod sauté	0.084
182	Cow pea sauté	0.031
183	Drumstick leaves sauté	0.043
184	Green gram sauté	0.017
185	Jack fruit seed sauté	0.016
186	Koorka sauté	0.045
187	Kovai sauté	0.112
188	Ladies finger sauté	0.028
189	Papaya sauté	0.057
190	Potato sauté	0.038
191	Raw banana sauté	0.036
192	Snake gourd sauté	0.030
193	Tapioca sauté	0.035
194	Yam Sauté	0.035
195	Brinjal fry	0.138
196	Pappad without oil	0.082

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
197	Potato fry	0.019
198	Yam fry	0.071
199	Boild egg	0.031
200	Boiled egg thoran	0.072
201	Bull's eye	0.034
202	Egg Aviyal	0.052
203	Egg curry (duck)	0.056
204	Egg curry (hen)	0.105
205	Egg omlette	0.111
206	Egg roast (hen)	0.034
207	Egg thoran	0.023
208	Crab thiyyal	0.021
209	Crab thoran	0.068
210	Dry fish curry	0.045
211	Dry fish fry	1.622
212	Fish curry (with coconut)	0.778
213	Fish curry (without coconut)	0.584
214	Fish fry	0.030
215	Fish peera	0.038
216	Fish (Tuna) thiyyal	0.028
217	Fish Thoran	0.020
218	Prawn curry	0.021
219	Prawn fry	0.020
220	Prawn thiyyal	0.023
221	Prawn thoran	0.034

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
222	Beef curry (Buffalo)	0.060
223	Beef curry (cow)	0.038
224	Beef fry (Buffalo)	0.027
225	Beef fry (cow)	0.067
226	Chicken curry (2 small pieces)	0.056
227	Chicken fry	0.039
228	Mutton curry	0.077
229	Mutton fry	0.120
230	Tomato onion salad	0.085
231	Tomato onion raita	0.032
232	Carrot cucumber salad	0.025
233	Carrot raita	0.037
234	Cucumber raita	0.069
235	Onion salad	0.014
236	Sprouted Green Gram Salad	0.019
237	Bilimbi pickle	0.021
238	Garlic pickle	0.027
239	Ginger curry	0.054
240	Ginger pickle	0.022
241	Goose berry pickle	0.105
242	Lime pickle	0.061
243	Mango pickle	0.173
244	Pappad	0.031
245	Achappam	0.019
246	Avalosu podi	0.043

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
247	Baji-Capsicum	0.051
248	Banana Baji (raw plantain)	0.051
249	Banana fry	0.056
250	Bengal gram vada	0.253
251	Biscuit	0.119
252	Biscuit (Cream)	0.016
253	Boli	0.033
254	Bonda	0.070
255	Bread	0.063
256	Bread with jam	0.028
257	Bun	0.034
258	Cake	0.025
259	Chakka Appam	0.044
260	Chicken roll	0.068
261	Chips (banana)	0.023
262	Chips (Jack)	0.114
263	Chips (potato)	0.036
264	Chips (Tapioca)	0.028
265	Cutlet (Veg)	0.029
266	Cutlet (non veg.)	0.029
267	Diamond Cut	0.025
268	Dilkush	0.022
269	Fruit Salad	0.055
270	Fruity bread	0.134
271	Ground nut	0.023

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
272	Ground nut candy	0.016
273	Halwa	0.022
274	Ice cream	0.019
275	Jilebi	0.058
276	Karachavu	0.034
277	Kuzhalappam	0.018
278	Laddu	0.018
279	Madhura Seva	0.021
280	Maladdoo	0.120
281	Masala Biscuits	0.022
282	Masala Dosai	0.018
283	Meat roll	0.029
284	Milk shake	0.105
285	Mixture	0.036
286	Modhakam	0.036
287	Munthirikothu	0.052
288	Murukku	0.015
289	Mysore pak	0.045
290	Nancut	0.018
291	Neyyappam	0.017
292	Noodles	0.291
293	Oats	0.036
294	Omappodi	0.046
295	Onion vada	0.050
296	Pakkavada	0.015

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
297	Pastry	0.018
298	Payasam	0.091
299	Peas fry	0.015
300	Peda	0.039
301	Pudding	0.022
302	Puffs (Egg)	0.048
303	Puffs (non - veg.)	0.222
304	Puffs (Veg.)	0.017
305	Rava laddoo	0.057
306	Rice flakes with jaggery	0.144
307	Rice balls	0.056
308	Rusk	0.025
309	Samosa – Veg.	0.087
310	Samosa – (non-Veg)	0.025
311	Tea – Rusk	0.037
312	Thirali	0.121
313	Till balls (small)	0.042
314	Unniappam	0.015
315	Urad Vadai	0.015
316	Vatta Appam	0.017
317	Beef soup	0.155
318	Corn soup	0.590
319	Mutton soup	0.111
320	Chicken soup	0.103
321	Rice soup	0.020

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
322	Tomato soup	0.020
323	Vegetable soup	0.027
324	Apple juice	0.059
325	Carrot juice	0.109
326	Grape juice	0.030
327	Lime ginger Juice	0.048
328	Lime juice	0.030
329	Mango juice	0.022
330	Mussambi	0.017
331	Orange juice	0.016
332	Pineapple juice	0.023
333	Tender coconut water	0.244
334	Water melon juice	0.030
335	Apple	0.039
336	Banana	0.016
337	Dates	0.036
338	Grapes	0.033
339	Guava	0.019
340	Jack fruit	0.041
341	Jamboo	0.019
342	Mango	0.024
343	Orange	0.665
344	Papaya	0.018
345	Pineapple	0.039
346	Plantain	0.015

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
347	Plum	0.020
348	Robesta	0.033
349	Seethafal	0.015
350	Water melon	0.026
351	Ghee rice	0.020
352	Kalan	0.021
353	Amaranth with moru curry	0.015
354	Bilimbi with dhal curry	0.015
355	Bread fruit curry	0.020
356	Boiled Kachil	0.011
357	Cluster beans juice	0.014
358	Soaked fenugreek	0.015
359	Pomegranate	0.077
360	Cashew nuts	0.140
361	Green mango	0.042
362	Noodles soup	0.015
363	Curd chilli	0.020
364	Gooseberry juice	0.014
365	Cucumber salad	0.015
366	Bullock's heart	0.210
367	Ghee roast	0.016
368	Drumstick leaves juice	0.867
369	Semolina gruel	0.064
370	Sugar	0.020
371	Yam curry	0.029

(continued)

FFQ Item	Food Recipe	Average portion per day (n=2047)
372	Green chille	0.016
373	Rasavada	0.015
374	Boiled Colocasia	0.016
375	Carrot salad	0.025
376	Peas onion sauté	0.022
377	Bread fruit thoran	0.019
378	Butter	0.035
379	Brinjal Kichadi	0.012
380	Coconut milk with sugar	0.017
381	Mulakoshum curry	0.038
382	Coffee with jaggery	0.015
383	Almond	0.034
384	Bread fruit thiyyal	0.059
385	Rice gruel with green gram	0.047
386	Colocasia leaf thoran	0.203
387	Dry fish chutney	0.208
388	Jack Fruit	1.423
389	Jack Pappad	0.168

Appendix 2. Modified Global Physical Activity Questionnaire (GPAQ)

Section A: Occupational Physical Activity

For every 8 hours spent at work, on a typical day how many hours do you spend on:

- Activities involving mainly sitting or standing with only a little walking
- Activities that require the same effort as heavy walking, cleaning etc.
- Activities that require the same effort as heavy lifting or heavy construction work

Section B: Non-occupational Physical Activity

For every 8 hours spent not at work, on a typical day how many hours do you spend on:

- Activities involving mainly sitting or standing with only a little walking
- Activities that require the same effort as heavy walking, cleaning etc.
- Activities that require the same effort as heavy lifting or heavy construction work

Section C: Transportation Physical Activity

On an average day, how much time do you spend on:

- Getting around in a car, bus or a motorcycle
- Getting around by walking or cycling