Applying Bayesian-networks to model multi-organ pathologies



Thomas H Waddell
Magdalen College
University of Oxford

Submitted Michaelmas 2023

This Thesis is submitted to the Department of Engineering Science, University of Oxford, in partial fulfilment of the requirements for the degree of Doctor of Philosophy

DECLARATION

I declare that this thesis is entirely my own work, and except where otherwise stated, describes my own research.

Thomas H Waddell

Magdalen College

DEDICATION

To my parents for their unrelenting support - no surrender; and to Wilson, for always putting things into perspective.

ACKNOWLEDGEMENTS

I am extremely appreciative for all of the support and guidance I have received throughout my DPhil studies and in the preparation of this Thesis. Firstly, I am very privileged to have been supervised and mentored by Sir Michael Brady who guided me throughout this Thesis, keeping me on the straight and narrow with his “size 11 boots”, shaped the narrative and the ‘why’ of this Thesis and allowed me to see the wood for the trees on many occasions. He has been extremally kind and selfless with his valuable time. Secondly, to Ana Namburete, my supervisor within the department of Computer Science who has been a focal catalyst for this Thesis and in shaping my understanding machine learning and artificial intelligence. I am again extremely grateful for her time and for welcoming me into her thriving lab. To my fellow lab mates at the Oxford Machine Learning in Neuroimaging Lab whose good humour and guidance were the most welcome introduction to the world of machine learning and artificial intelligence.

I am, and always will be, hugely appreciative to Adam Lewandowski whose kindness and belief in me is what started my entire journey at Oxford. None of this would have been possible without you – thank you.

My colleagues and mentors at Perspectum have been focal in allowing this DPhil to happen. Firstly, Banjo for not only his guidance and valuable time throughout the completion of this DPhil, but for is mentorship since I began at Perspectum and in shaping the person I am today. I am also extremely grateful for my fellow colleagues at Perspectum who have selflessly offered up their time, opinions and help during my time at Perspectum.

My supervisors within the Oxford Robotics Institute, where I started my DPhil, were extremely patient and helpful in getting my journey into the field of AI and machine learning off the ground – Nick Hawes, Paul Duckworth.

Lastly, I would like to acknowledge to Royal Commission for the Exhibition of 1851 whose funding and Industrial Fellowship award has allowed me to pursue these doctoral studies. Your belief in both me and my DPhil project have allowed this Thesis to come together.

Table of Contents

DECLARATION 2

DEDICATION 3

ACKNOWLEDGEMENTS 4

PUBLICATIONS ARISING FROM THESIS WORK 10

ACRONYMS AND ABBREVIATIONS 12

LIST OF FIGURES 14

LIST OF TABLES 16

PREFACE 17

ABSTRACT 18

CHAPTER 1 - OBESITY, METABOLIC DISEASE & MULTI-ORGAN PATHOLOGIES 21

1.1 BRIEF HISTORY AND DEFINITION OF OBESITY 21

1.2. EPIDEMIOLOGY OF OBESITY 22

1.3. OBESITY-RELATED DISORDERS 25

1.4. ECTOPIC FAT DEPOSITION 28

1.5. IMAGING BIOMARKERS 30

1.6. INTERVENTIONS FOR THE MANAGEMENT OF METABOLIC DISEASE 35

1.7. THE CLINICAL CHALLENGE 36

1.8. EXPLORING T2D PHENOTYPES IN THE UK BIOBANK 38

1.8.1. DATA SELECTION 39
1.8.2. STATISTICAL ANALYSIS 40
1.8.3. RESULTS 41
1.8.4. DISCUSSION 46

CHAPTER 2 - LIMITATIONS OF TRADITIONAL STATISTICAL APPROACHES FOR STUDYING BIOLOGICAL SYSTEMS 50

2.1. CAUSE, EFFECT, AND SPURIOUS CORRELATIONS.....50

2.2. SPURIOUS CORRELATIONS 52

2.3. COMPLEXITY OF BIOLOGICAL SYSTEMS 54

2.4. FOUNDATIONS OF BAYESIAN-NETWORKS 59

CHAPTER 3 - BAYESIAN-NETWORKS AND PROBABILISTIC INFERENCE 62

3.1. STRUCTURE OF BAYESIAN-NETWORKS 62

3.2.DERIVING THE STRUCTURE OF A BAYESIAN-NETWORK.....67

3.3. PROBABILITY DISTRIBUTIONS & PARAMETERISATION OF BAYESIAN-NETWORKS 73

3.4. ESTIMATING TREATMENT OUTCOMES AND SAMPLING FROM THE BAYESIAN-NETWORK.....78

3.6. EXISTING APPLICATIONS OF BAYESIAN-NETWORKS FOR STUDYING BIOLOGICAL SYSTEMS.....87

CHAPTER 4 - MODELLING FAT DEPOSITION AND HYPERGLYCAEMIA: SEMANTIC NETWORKS AND INSIGHTS FROM BAYESIAN-NETWORKS 94

4.1. INTRODUCTION.....94

4.2. METHODOLOGY 96

4.2.1. DATA COLLECTION & PREPARATION.....96
4.2.2. DERIVING THE BAYESIAN-NETWORK STRUCTURE.....97
4.2.3. COMPARISON OF AUTOMATED AND EXPERT-LED NETWORK CONSTRUCTION 100
4.2.4. NETWORK COMPARISON METRICS.....100
4.2.5. RESULTS FROM NETWORK COMPARISON.....102

4.3. APPLICATION OF PROBABILISTIC REASONING 103

4.4. RESULTS.....104

4.4.1. DESCRIPTIVE STATISTICS OF POPULATION 104

4.4.2. ESTIMATING THE PROBABILITY OF HYPERGLYCEMIA 105

4.5. DISCUSSION 108

4.6. CONCLUSION 111

CHAPTER 5 - APPLYING BAYESIAN-NETWORKS TO STUDY THE IMPACT OF PROJECTED INCREASES IN OBESITY PREVALENCE ON INCIDENT LIVER DISEASE IN THE UK 113

5.1. INTRODUCTION 113

5.2. METHODS 115

5.2.1. DATA COLLECTION AND PREPARATION 115
5.2.2 BAYESIAN-NETWORK CONSTRUCTION AND INFERENCE 116

5.3. RESULTS 117

5.3.1. POPULATION CHARACTERISTICS 117
5.3.2. ESTIMATING CONSEQUENCES OF INCREASED OBESITY PREVALENCE 119

5.4. DISCUSSION 123

5.5 CONCLUSION 125

CHAPTER 6 - APPLYING BAYESIAN-NETWORKS TO STUDY CONCOMITANT NON-ALCOHOLIC STEATOHEPATITIS IN PEOPLE WITH TYPE-2 DIABETES 127

6.1. INTRODUCTION 127

6.2. PROJECT METHODOLOGY 129

6.2.1. DATA COLLECTION & PREPARATION 129
6.2.2. CLINICAL DETAILS AND MRI IMAGE ACQUISITION AND ANALYSIS 129
6.2.3. BAYESIAN-NETWORK CONSTRUCTION 130
6.2.4. STATISTICAL ANALYSIS 130

6.3. RESULTS 132

6.4. DISCUSSION 138

6.5. CONCLUSION....140

CHAPTER 7 - APPLYING BAYESIAN-NETWORKS TO STUDY HOSPITALISATION RISK FOLLOWING ACUTE COVID-19....142

7.1. INTRODUCTION....142

7.2. PROJECT METHODOLOGY....144

7.2.1. DATA COLLECTION & PREPARATION....144
7.2.2. BAYESIAN-NETWORK CONSTRUCTION....145
7.2.3. PROBABILISTIC INFERENCE....147
7.2.4. PREDICTION OF HOSPITALISATION STATUS....148
7.2.5. STATISTICAL ANALYSIS....148

7.3. RESULTS....149

7.4. DISCUSSION....152

CHAPTER 8 - CONCLUSIONS & FURTHER WORK....157

8.1. SUMMARY OF KEY RESEARCH FINDINGS AND THESIS CONTRIBUTIONS....160

8.2 AREA OF FURTHER WORK 1 – DYNAMIC BAYESIAN-NETWORKS....163

8.2.1. PARTICIPANT DATA....166
8.2.2. BAYESIAN-NETWORK CONSTRUCTION....166
8.2.3. RESULTS....168
8.2.4. DISCUSSION....170

8.3. AREA OF FURTHER WORK 2 – PRE-SURGICAL DECISION SUPPORT....171

8.3. FINAL SUMMARY OF THESIS....173

REFERENCES....175

APPENDIX....186

PUBLICATIONS ARISING FROM THESIS WORK

Journal publications

Waddell, T., Bagur, A., Cunha, D., Thomaides-Brears, H., Banerjee, R., Cuthbertson, D.J., Brown, E., Cusi, K., Després, J.P. and Brady, M., 2022. Greater ectopic fat deposition and liver fibroinflammation, and lower skeletal muscle mass in people with type 2 diabetes. *Obesity*.

Waddell, T., Namburete, A., Duckworth, P., Fichera, A., Telford, A., Thomaides-Brears, H., Cuthbertson, D.J. and Brady, M., 2023. Poor glycaemic control and ectopic fat deposition mediates the increased risk of non-alcoholic steatohepatitis in high-risk populations with type 2 diabetes: Insights from Bayesian-network modelling. *Frontiers in endocrinology*, *14*.

Waddell, T.H., Namburete, A., Duckworth, P., Eichert, N., Thomaides-Brears, H., Cuthbertson, D., Despres, J.P., Brady, M. Bayesian-networks and imaging derived phenotypes highlight the role of fat deposition on hospitalisation risk with COVID-19. *Frontiers in Bioinformatics*, *3*, p.41.

Journal publications under review

Waddell, T., Uazman, A., Cuthbertson, D., Brady, M. (2023). The emerging applications of Bayesian-networks in clinical medicine. *Clinical Medicine, Royal College of Physicians.* (under review).

Conference publications

Waddell, T., Bagur, A., Cunha, D., Thomaides-Brears, H., Banerjee, R. and Brady, M., 2021, September. Multi-organ multiparametric magnetic resonance imaging reveals distinct ectopic fat distribution in type 2 diabetics with and without co-existing obesity. *Diabetologia (*Vol. 64, pp. 50-50).

Waddell, T., Roca-Fernandez, A., McGonigle, J., Jandor, A., Carapella, V., Thomaides-Brears, H., Dennis, A., Brady, M. and Banerjee, R., 2021. Elevated visceral and liver fat content significantly increase risk of hospitalisation following COVID-19 infection. *Hepatology*, pp.317A-317A.

Waddell, T., Roca-Fernandez, A., McGonigle, J., Jandor, A., Carapella, V., Thomaides-Brears, H., Dennis, A., Brady, J.M. and Banerjee, R., 2022. Ectopic fat deposition in the liver and pancreas and the presence of excess visceral adipose tissue significantly increase hospitalisation risk following COVID-19. *Journal of Hepatology*, *77*, p.S424.

Y Chun Wang, Salvani, R., Connel, J., Var Zig, N., **Waddell, T.,** Bulte, D., Brady, M. (2022). Bayesian Networks reveal the Interplay between quantitative liver MRI metrics. ISMRM.

Waddell, T., Namburete, A., Duckworth, P., Cuthbertson, D., McMracken, C., Brady, M. The impact of projected increases in obesity prevalence on incident liver disease in the UK: Insights from Bayesian-network modelling (2023). *International Conference on Medical and Health Informatics.* (Oral presentation).

Waddell, T., Namburete, A., Duckworth, P., Cuthbertson, D., McMracken, C., Brady, M. The impact of projected increases in obesity prevalence on incident liver disease in the UK: Insights from Bayesian-network modelling (2023). *European Association for the Study of Liver Disease.*

Waddell, T., Namburete, A., Brady, M. Modelling fat deposition and hyperglycaemia: Insights from Bayesian-networks (2023). Oral Presentation. *European Association for the study of Diabetes.*

ACRONYMS AND ABBREVIATIONS

AI: Artificial Intelligence

AIC: Akaike's information criterion

ALT: Alanine Transaminase

AST: Aspartate Transferase

BC: Body composition

BIC: Bayesian information criterion

BMI: Body mass index

BN: Bayesian-network

CI: constraint based

cT1: corrected T1

CVD: Cardiovascular disease

DAG: Directed acyclic graph

DHG: diabetic-status hyperglycaemia

GGT = Gamma-glutamyl Transferase

HbA1c: Glycated haemoglobin

HC: Hill-climbing

HDL-c: High density lipoprotein cholesterol

HTN: Hypertension

IAMB: Incremental Association Markov Blanket

Int: Intercept

LDL-c: Low density lipoprotein cholesterol

MI: Myocardial infarction

MMHC: Max-min hill climbing

MRI: Magnetic resonance imaging

MS: Metabolic syndrome

NAFLD: Non-alcoholic fatty liver disease

NASH: Non-alcoholic steatohepatitis

NICE: National Institute for health and care excellence

PC: PC stable

PDFF: Proton density fat fraction

RF: Radio frequency

SAT: Subcutaneous adipose tissue

SHD: Structural hamming distance

SMI: Skeletal muscle index

SS: Search and score

T2D: Type-2 diabetes

Tchol: Total cholesterol

TE: Time to echo

TR: Repetition time

TRIG: Triglycerides

TS: Tabu search

UKBB: UK Biobank

VAT: Visceral adipose tissue

WC: Waist circumference

LIST OF FIGURES

	Number	Figure title
Chapter 1	1.1	Obesity statistics in England, taken from the 2023 Research Briefing (obesity statistics), House of Commons Library
	1.2	J-shaped curve association between obesity and cardiovascular disease.
	1.3	Example LMS images with reference PDFF and cT1 values.
	1.4	L3 slice positioning (numbered) for extraction of body composition segmentation.
	1.5	Example manual body composition segmentation.
	1.6	Liver cT1 (ms) in participants stratified by obesity and diabetes status.
	1.7	Liver fat (PDFF %) in participants stratified by obesity and diabetes status.
Chapter 2	2.1	Illustration of the spurious correlation between a nation's chocolate consumption and number of Nobel prize winners.
	2.2	Potential spurious correlation between Vitamin D consumption and COVID-19 related infection and mortality.
	2.3	Moving from association (left) to explicit direction of conditional dependence.
	2.4	Moving from a simple 1-1 interaction between two biomarkers (left), to explicitly encoding multiple conditional dependencies.
	2.5	Pearl's ladder of causality, from 1) association to 3) counterfactuals.
Chapter 3	3.1	a) example chain connection between three variables: *A, B* and *C*; b) example fork connection; c) example collider connection
	3.2	Example fork connection, whereby the value of A is a common cause of both B and C.
	3.3	Example Bayesian-network with corresponding conditional probability tables.
	3.4	Representation of Bayes' Theorem.
	3.5.	Full joint probability distribution of example Bayesian-network with conditional probability tables. M = male, F = female; -0 = non-obese, +O = obese; -D = non-diabetic, +D = diabetic; -H = no heart attack, +H confirmed heart attack and '|' denotes 'given'.

	3.6.	Full conditional probability distribution with highlighted (green) sample and highlighted (red) evidence variables and corresponding weights. M = male, F = female; -0 = non-obese, +O = obese; -D = non-diabetic, +D = diabetic; -H = no heart attack, +H confirmed heart attack and '|' denotes 'given'.
Chapter 4	4.1	Final 'semantic' Bayesian-network structure following manual adjustment. Values within each variable denote the coefficient value, given the parents of that specific variable
	4.2	Illustration showing the simplified Bayesian-network structure (left) and corresponding probability increases of DHG (right)
Chapter 5	5.1	Final Bayesian-network structure following manual adjustment.
	5.2	Illustration of the estimated key consequences of increased obesity on liver and metabolic health parameters, per ~108k people.
	5.3	Illustration of the estimated impact of increased obesity prevalence in areas of low socio-economic status on liver disease, per ~108k people.
Chapter 6	6.1	Bayesian-network structures from the T2D (top) and non-T2D (bottom) cohorts.
	6.2	Probability of NASH (%) by obesity category (BMI) in the non-T2D (red) and T2D (blue) cohorts.
	6.3	Probability of NASH (%) under VAT (top left), SAT (top right), SMI (bottom left) and HbA1c status (bottom right) conditions in the non-T2D (red) and T2D (blue) cohorts.
	6.4	Illustration of main project findings.
Chapter 7	7.1	Final semantic Bayesian-network structure.
	7.2	ROC plots indicating the performance of classification models in predicting the hospitalisation status AUC [95% CI].
Chapter 8	8.1	Illustration of the twin-cycle hypothesis taken from Al-Mrabeh et al. (2020). Elevated plasma glucose is driven by the liver and pancreas cycles, whereby elevated fat deposition drives de-novo lipogenesis and dysregulated insulin secretion.
	8.2	Example dynamic Bayesian-network with three time slices. Dashed and solid arrows represent intra and inter time slice edges, respectively.
	8.3	Dynamic Bayesian-network structure.
	8.4	Example Bayesian-network for liver surgical planning; FLR = future liver remnant.

LIST OF TABLES

	Number	Table title
Chapter 1	1.1	Anthropometric and demographics of study population. Data is median [IQR], significant p-values in bold.
	1.2	Biochemical and imaging characteristics. Data is median [IQR], significant p-values in bold.
	1.3	Abnormalities of organ structure/function significant p-values in bold.
	1.4	Spearman's correlation analysis, significant p-values in bold.
Chapter 2	1.1	Example correlations from *Obesity* publication.
Chapter 3	3.1	Overview example automated network construction algorithms.
	3.2.	Example distribution from 10 random sample instantiations.
	3.3.	Example distribution with corresponding weights from three random samples.
Chapter 4	1.1	Example edge connections and supporting literature inserted into the semantic network.
	1.2	Comparison metrics of automated network structures against semantic network.
	1.3	Probability of DHG after fixing the value (evidence) of specific network variables.
Chapter 5	5.1	Descriptive statistics of entire population cohort. Data is Median [IQR].
	5.2	Baseline and estimated probabilities of variable states with corresponding change in probability and population number after intervening on the variable 'Obesity'.
Chapter 6	6.1	Baseline characteristics of T2D and non-T2D cohorts. Data is median [IQR]. Bold denotes significance.
	6.2	Probability of NASH (%) under different variable settings in the T2D and non-T2D cohorts.
Chapter 7	7.1	Population characteristics between hospitalised and non-hospitalised participants. Data is median [IQR], significant p-values in bold.
	7.2	Probability of hospitalisation (%) and probability change (+/- %) across BMI-categories.
Chapter 8	8.1	Descriptive statistics of population. Data is mean ± standard deviation.
	8.2	Probability of obesity status and liver events between lifestyle and GLP-1RA treatment arms.

PREFACE

In 2019 I began working at Perspectum, an imaging diagnostics company developing cutting edge technology for characterising multi-organ health. My role here was primarily to study multi-organ dysfunction in metabolic disease and the phenotypes associated with concurrent fatty liver disease, type-2 diabetes, and cardiovascular disease. With a wealth of information at my fingertips, I explored such data using the classical statistical techniques that my previous education had afforded me – the pairwise comparisons, correlations, and significance thresholds that we're all too familiar with. However, I was deeply unsatisfied with such approaches, their limitations obvious to me. I was then handed a copy of Judea Pearl's book 'The Book of Why' by Perspectum's chairman and founder, and my eventual DPhil supervisor, Sir Michael Brady. This book, listed below for the inclined reader, transformed the way I observed data and the potential inferences I could glean from it, and is nicely illustrated as the 'Ladder of Causality' found here in Chapter 3. This book also introduced me to the focal contribution of this Thesis – Bayesian-networks. This Thesis is the written account of the work that followed, specifically applying Bayesian-networks to study multi-organ imaging data to explore phenotypes associated with metabolic disease and to 'speak' the language of clinicians who are making complex decisions about their patients' care.

The Book of Why - Pearl, J. and Mackenzie, D., 2018. *The book of why: the new science of cause and effect*. Basic books.

ABSTRACT

The care and management of people with metabolic disease is incredibly complex, and increasingly so. Clinicians must not only assess their patient's current health and scope of comorbidities, but also seek to deliver an optimal treatment plan while balancing the risk profiles and contraindications of individual therapies. For example, obesity is associated with an increased risk of type-2 diabetes and non-alcoholic fatty liver disease, however, there exists a synergistic risk of liver related complications in people with concomitant obesity and T2D. This further compounds the complexity of clinical decision making, which is inherently counterfactual in nature, where clinicians must post questions such as 'what would happen to my patients' risk of advanced liver disease, given that we are able to control their blood glucose'? Traditional statistical techniques such as correlations and pairwise comparisons do not possess the necessary language for posing and answering such questions.

Furthermore, the body mass index is intrinsically limited for the study of metabolic disease, for it confounds all facets of *body composition* (visceral and subcutaneous adiposity, lean muscle) into a single measure. Rather, the *distribution* of fat, namely in ectopic sites such as the liver and viscera, is the primary determinant of metabolic health.

This Thesis applies Bayesian-networks to magnetic resonance imaging derived measures of fat deposition and organ health, alongside routinely collected biomarkers such as anthropometrics and circulating biomarkers, to explore phenotypes associated with a high risk of metabolic-related diseases. This Thesis is centred around four central questions though are connected by a single shared theme – can Bayesian-networks assist clinicians with their

increasingly complex decision making by unravelling high-risk phenotypes and potential treatment outcomes?

Chapters 1-3 introduce the clinical need, limitations of current statistical approaches and the field of Bayesian-networks, respectively, illustrating the gap in the literature that this Thesis seeks to address. Chapter 4 sets out the initial development of a Bayesian-network in collaboration with Prof Cuthbertson and Prof Despres and compares a semantic network against those returned by automated structure learning algorithms, revealing their limitations in returning biologically plausible networks and the need to adopt semantic networks in Chapters 5-7. This Chapter then reveals how diabetic-status hyperglycaemia (DHG) was driven by fat deposition in multiple, as opposed to single, sites. That is, elevated visceral adipose tissue (VAT) or subcutaneous adipose tissue (SAT) in isolation only elicited a 4% and 2% increase in the probability of DHG, respectively, however an example phenotype with elevated subcutaneous adipose tissue (SAT), visceral adipose tissue (VAT) and liver fat resulted in a marked increase of DHG (39%).

Chapters 6 and 7 continue to study the effect of body composition and fat deposition on clinical outcomes, where Chapter 6 explores the probability of non-alcoholic steatohepatitis (NASH) in participants with type-2 diabetes, revealing that elevated VAT, liver fat and blood glucose independently caused a 20%, 31% and 22% increase in the probability of NASH, respectively. Chapter 7 then demonstrates how the existing Bayesian-network can be pivoted to explore the association between obesity and an increased risk of hospitalisation from acute COVID-19. This analysis revealed that although moving from ‘normal weight’ to ‘obesity’ caused an 18% increase in the probability of hospitalisation, this was primarily driven by

elevated fat deposition within VAT, and not SAT, sites. Furthermore, across all obesity classifications, elevated liver fat (>10%) caused a 39% average increase in the probability of hospitalisation.

This Thesis illustrates how Bayesian-networks can be applied to explore the unique phenotypes associated with metabolic-related disease and overcome the limitations of the BMI through modelling *fat deposition*. For example, in Chapter 4 we report how a phenotype with low SAT and SMI, but elevated VAT and liver fat, resulted in the highest probability of DG (52%); while in Chapter 7 reducing liver fat under 'normal weight' settings reduced the probability of hospitalisation by 29%. These finding emphasises the importance of considering how the *weight* of someone is constituted and how the usual treatment recommendations (5-10% weight loss) may not be suitable for certain people considered normal weight. This Thesis also illustrates how Bayesian-networks are an effective method for modelling the complexity of biological systems and can be applied to ask the type of *counterfactual* questions posed by clinicians in their clinical decision making.

1

OBESITY, METABOLIC DISEASE & MULTI-ORGAN PATHOLOGIES

This Thesis is concerned with modelling and distilling the complexity of metabolic disease into concise, clinically meaningful questions. To this end, the Thesis encompasses several areas of science and engineering including imaging, Artificial Intelligence (AI) and medicine. This Chapter introduces the medical context and imaging aspects of such work, motivating the Thesis and making explicit the clinical challenge that it seeks to tackle.

1.1 Brief history and definition of obesity

Once praised as a sign of wealth and largely limited to the aristocracy, obesity was uncommon for the most part of human history, only growing in prevalence to epidemic proportions in the past ~50 years. Although considered a modern-day disease, with the term 'obesity' only appearing in the English language during the seventeenth century [1], the consequences of excess weight have long been documented, with Hippocrates first noting its impact on health in 400 BC:

"It is injurious to health to take in more food than the constitution will bear, when at the same time one uses no exercise to carry off this excess…"

- Translated from Hippocrates, circa 400 BC.

However, it took until the nineteenth century for obesity to be recognised as a *cause* of poor health, and until the first decades of the twentieth century for the complications of obesity to be officially documented [1]. While the full effects are still being unravelled, the consequences of obesity and obesity-related diseases have been extensively documented for many years. We next overview multi-organ pathologies, highlighting the scale of obesity and obesity-related disorders, as well as the role that ectopic fat deposition plays in driving metabolic disease, particularly type-2 diabetes (T2D).

1.2. Epidemiology of obesity

Obesity is highly complex, with its causes spanning biological, cultural, and sociodemographic domains. That said, the 'Occam's razor' of the obesity crisis is that nowadays people in (typically) developed countries are generally less active yet consume considerably more calories than they require. For example, people living in the UK are an estimated 20% less active now than in the 1960s and this is estimated to increase to 35% by 2030 [2]. Children are also significantly less active, with only 22% of 5–15-year-olds meeting the physical activity guidelines set by the UK Government, a modest target of being moderately active for 60 minutes per day.

This increase in sedentariness is coupled with an increased availability (and consumption) of cheap, energy dense foods. More than one quarter of all adults are reported to eat from food outlets at least once a week, equating to 18% of total meals being eaten outside of the home [2]. Such foods are typically higher in refined fats and carbohydrates yet contain less micronutrient content, resulting in a higher calorie meal with less nutritional value [3]. For

example, a large UK Biobank analysis of 210,106 participants revealed that 32% of men and 42% of women consume more energy than the recommended daily amount [4]. This analysis revealed that despite this excessive calorie intake, both men and women were significantly deficient in dietary fibre intake, while exceeding the recommended intake of saturated fat and sugar. Combined with a generally higher societal acceptance of larger phenotypes [5], we are now living in what is considered an 'obesogenic' environment which is catalysing the obesity and obesity-related disorder crisis.

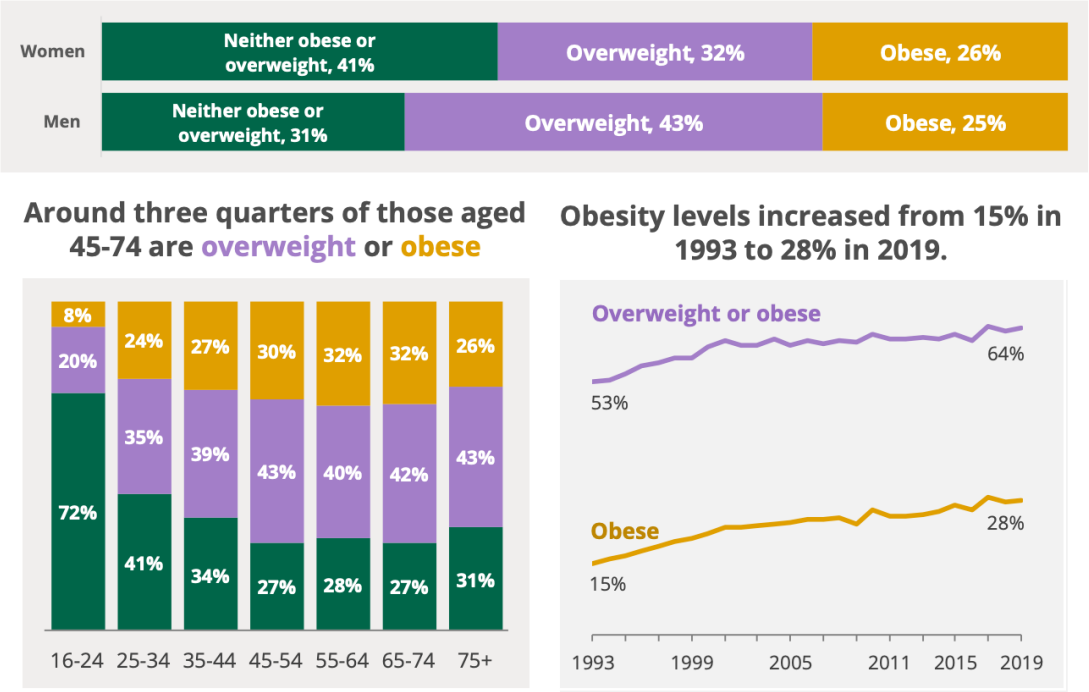


Figure 1.1. Obesity statistics in England, taken from the 2023 Research Briefing (obesity statistics), House of Commons Library [6].

In 2021, the WHO announced that worldwide obesity had nearly tripled since 1975, with 39% and 13% of adults (18 years and older) now classified as overweight and obese, respectively

[7]. In the UK, the proportion of adults classified as overweight or obese rose from 52.9% to 64.3% between 1993-2022, with people aged 45-74 years most likely to be overweight or obese [6]. The influence of sociodemographic factors on excess weight in England is illustrated through the UK Government's 2020/21 report [6], showing that the prevalence of overweight or obesity was 72% in the most deprived areas, 20% higher than the least deprived (52% vs 72%). Furthermore, this report showed that black ethnic groups were also most likely to be overweight or obese (72%), while disabled people were 12% more likely than non-disabled (61% vs 73%).

Despite putative governmental strategies, such as restrictions on the sale and advertisement of sugary foods, Society's response to the growing obesity crisis has been inadequate. Global childhood obesity has risen more than 10-fold over the past 40 years, from an estimated 11 million to 124 million in 2016 [8]. In the UK, 14.3% and 23.4% of children aged 10-11 were recognised as overweight and obese, respectively, in 2021 [6]. Similarly, children aged 10-11 living in the most deprived areas were 17.8% more likely to be obese or overweight compared those living in the least deprived areas (13.5% vs 31.3%), with total obesity rates in the most deprived rising by 10 percentage points in the last 15 years. Childhood obesity poses challenges beyond physical health, with research showing an association with a higher likelihood of being bullied, poorer school attendance, achievements, and employment prospects; and salary achievements in adulthood [9], [10].

Recent estimates from the Cancer Research UK institute [11] suggest that, if current trends continue, by 2040 71% of the adult population in England (36 million people) will be overweight or living with obesity, with the proportion of the population living with obesity

(37%) surpassing that of the population deemed to have a healthy weight (29%). This represents an 11% increase in the prevalence of overweight and obesity from 64% in 2019. Furthermore, while the current rate of severe obesity (body mass index [BMI] >40kg/m^2) is relatively low (3%), it is predicted that this figure will increase to 7% (102% relative increase) by 2040, with the deprivation gap for obesity between the most and least deprived areas projected to more than double by 2040 [11]. Such projections threaten to engulf already-overstretched public health system. It is already estimated that obesity costs the NHS £6 billion annually and this is projected to reach £9.7 billion by 2050 [2], an astonishing 385% relative increase from £2 billion since 2001 [12]. However, the true societal costs of excess weight are considerably higher, with the wider societal cost of obesity estimated at £27 billion in 2017 due to factors such as time spent out of work [2].

1.3. Obesity-related disorders

Obesity and excess weight elicit several endocrine, mechanical, and metabolic effects on the body and is associated with a sequalae of metabolic-related disorders, including type-2 diabetes (T2D), non-alcoholic fatty liver disease (NAFLD) and cardiovascular disease (CVD) [13]. In the United States (US) and UK, an estimated 30 and 4.3 million adults, respectively, have been diagnosed with type-2 diabetes (T2D), with disease prevalence expected to double by 2030 [14]. Recent estimates suggest that the global prevalence of NAFLD in the general population is ~25%, with US-specific projections estimating a 63% prevalence of the advanced form of liver disease, non-alcoholic steatohepatitis (NASH), by 2030 [15].

Overweight and obese individuals are three and seven times more likely, respectively, to develop T2D compared to those with a 'normal' BMI [16]. There is a J-shaped curve between CVD morbidity and increasing obesity status [17], see Fig. 1.2. Obesity is an independent risk factor for NAFLD, with or without concomitant inflammation and fibrosis (NASH) [18]. However, in people with both obesity and T2D, there appears to be an increased synergistic risk of liver-related complications. Worldwide, NAFLD is present in 55.5% of people with T2D, 37.3% of whom demonstrate co-existing NASH and 17.3% with biopsy-confirmed advanced fibrosis [19]. In BMI matched people, those with T2D reportedly have significantly elevated liver fat content and liver disease activity (as measured by corrected T1, introduced later in this Chapter) [20]. People with co-existing T2D and NAFLD are significantly more likely to develop fibroinflammatory change (NASH) compared to those with NAFLD alone [19], [21] and are at a greater risk of non-liver complications such as CVD [22]–[24]. Furthermore, both obesity and T2D are independent risk factors for CVD, with T2D estimated to cause over 590 heart attacks per week just in the UK [25]. In the US, cardiovascular disease (CVD) is the leading cause of death for men and women, with an estimated one person dying every 34 seconds from CVD [26].

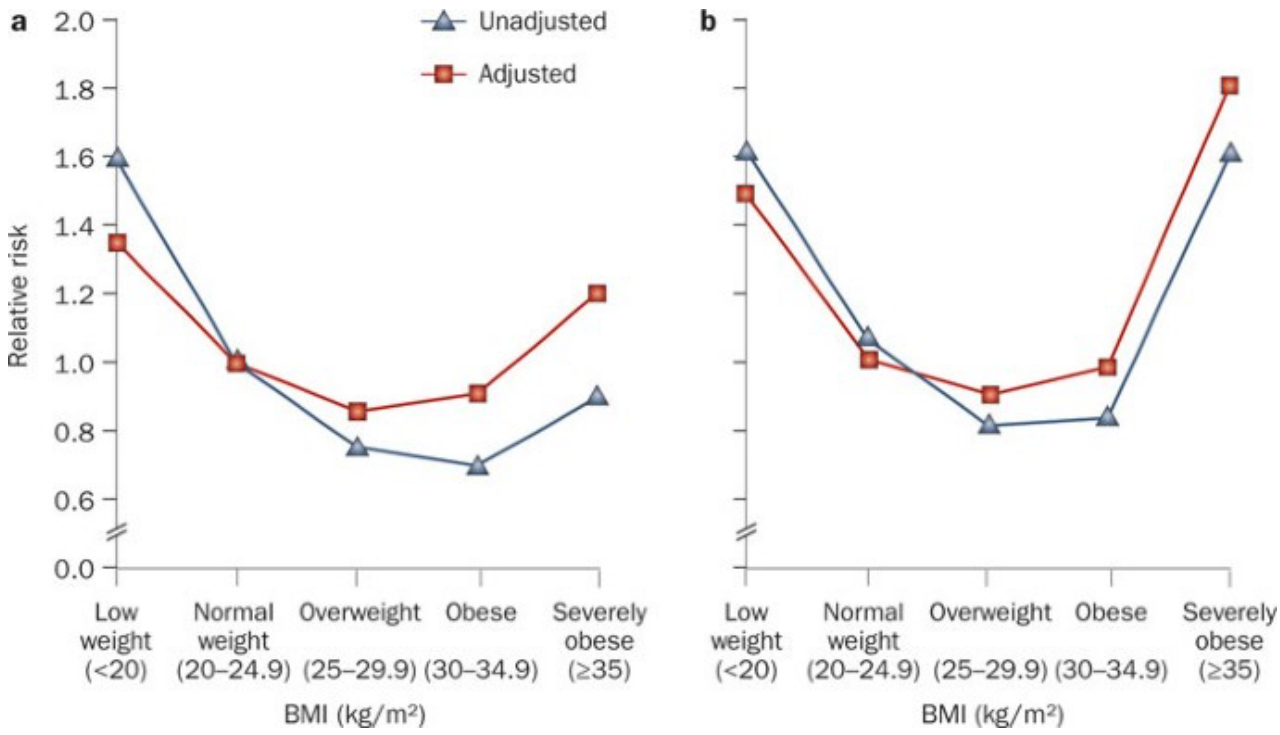


Figure 1.2. J-shaped curve between cardiovascular mortality and BMI. A) unadjusted and adjusted relative risk according to BMI on total mortality; b) unadjusted and adjusted relative risk according to BMI on coronary artery disease. Taken from Valery et al. (2011) [27].

The early stages of metabolic disorders may be 'silent', for example fatty liver disease and early-stage NASH are typically asymptomatic, with an estimated 79.8% of the prevalent UK NASH population being undiagnosed in 2018 [28]. Consequently, people may otherwise appear 'healthy' and may remain asymptomatic until they experience a major clinical event or develop a cluster of comorbidities. This Thesis is concerned with the mechanistic drivers and phenotypes of such obesity-related complications with the aim of developing early preventative strategies.

1.4. Ectopic fat deposition

Obesity status is generally classified using the BMI, a continuous measure calculated by taking a person's weight (kg) and dividing by the square of their height (m^2). Thresholds at various values of the BMI, result in a categorical ranking from 'underweight' to 'severe obesity', have become well established to determine one's obesity status. The BMI was enrolled into mainstream health and policy in the 1950s and continues to be the standard for defining obesity, driving obesity-related research, and informing public health policies. However, the BMI is intrinsically limited, fundamentally because it confounds all facets of body composition (lean muscle, adipose tissue, musculoskeletal tissue) into a single measure, disregarding the effect of *fat deposition* on metabolic health.

Ectopic fat deposition can develop when a person is in a state of chronic, positive energy balance. Excess calories in the form of free fatty acids (FFAs) are initially 'stored' within subcutaneous adipose tissue (SAT). However, once the storage capacity of SAT is exceeded, FFAs shift towards ectopic sites, notably the heart, viscera, liver, and vasculature [29], [30]. The mystery behind the preferential storage sites of FFAs has not been fully unravelled, with ethnicity, hormonal balance, sex, and age all considered to exert an influence [18]. For example, South Asian individuals have been shown to accumulate greater VAT and are significantly more likely to develop T2D than White European individuals of a similar BMI [31], [32]. Excess fat deposition within ectopic sites is a critical component of metabolic dysfunction and is independently associated with a greater risk of metabolic disease. Elevated fat deposition in the liver and pancreas is independently significantly associated with T2D incidence [33], [34], and elevated VAT, but not SAT, is also associated with a significant increase in circulating glucose and insulin in BMI-matched individuals [35], [36]. In addition

to T2D, elevated VAT and fat deposition within epicardial and pericardial sites and within the liver are associated with acute myocardial events and significant CVD risk factors such as hypertension and dyslipidaemia [37], [38]. Of course, fat deposition within the liver in itself is a condition – non-alcoholic fatty liver disease – and if left untreated can progress to NASH, cirrhosis, and hepatocellular carcinoma [39].

The role of fat deposition is highlighted through the large variation in metabolic disease risk in people with overweight or obesity. For example, individuals with overweight or obesity, yet low levels of VAT, are associated with a more favourable CVD risk profile and lower risk of secondary complications compared to those of a similar BMI [18]. Such a phenotype is commonly titled 'metabolically healthy obesity' (MHO) and refers to a subgroup of people who, despite presenting with obesity, do not exhibit overt comorbidities or elements of the metabolic syndrome, such as impaired insulin action or dyslipidaemia. The MHO phenotype is typically characterised by a phenotype with lower measures of liver fat and higher subcutaneous leg fat content; lower measures of inflammatory markers, greater insulin sensitivity and a superior cardiorespiratory fitness [40]. However, there exists an open question as to whether people with MHO are simply a group in transition to developing metabolic impairments, termed metabolically unhealthy obesity (MUO). Still, fat deposition is a critical factor of metabolic health, irrespective of obesity status, and therefore to truly understand metabolic disease risk we must consider how the *weight* of an individual is constituted.

Fat distribution of VAT, SAT, and liver fat using magnetic resonance imaging (MRI) has revealed six unique sub phenotypes of patients with T2D [41]. The authors reported that the

'cluster' with the highest risk of vascular disease, T2D and renal disease included insulin resistant patients with obesity but also elevated liver fat deposition. Wagner et al applied soft clustering to 32 clinical variables and highlighted how the heterogeneity of T2D can be translated to individual etiological processes [42]. Including newly diagnosed T2D patients, the authors identified four archetypical profiles, revealing those with the fastest disease progression had concurrent obesity, dyslipidaemia, impaired beta cell sensitivity, and insulin resistance.

1.5. Imaging biomarkers

The interplay between ectopic fat deposition and metabolic-disease risk is highly complex, with many variables interacting in myriad ways. This complexity is compounded with the merging of several different sources of data such as imaging, circulating biomarkers and patient demographics that the clinician must take account of in their clinical decision making. This Thesis was completed in partnership with Perspectum Ltd, an image analysis company that leverages magnetic resonance imaging (MRI) to provide a quantitative assessment of organ characteristics. Below I provide a brief overview of the key imaging biomarkers developed by Perspectum that are used throughout this Thesis.

MRI works by placing a participant into a static magnetic field (B_0), the strength of which is measured in tesla (T). The nucleus of hydrogen atoms found in water (of which the body is approximately 60% constituted of), consist of a single positively charged proton that possess an intrinsic property called spin angular momentum. Such 'spin' means protons have a magnetic moment, and when placed in an external magnetic field (B_0), will orientate

themselves either towards or against this external magnetic field. As given by the Larmor equation, the protons also precess around the axis of the magnetic field, denoted as $\omega = \gamma B_0$ where ω is the angular precissional frequency, γ is the gyromagnetic ratio and B_0 is the external magnetic field strength.

To create a signal, MRI applies a strong radio frequency (RF) pulse with the Larmor frequency perpendicular to the B_0 field, causing the protons to precess around the new field in addition to B_0. Once the radio frequency is switched off, the protons 'relax' and become increasingly out of phase, the rate of which is influenced by tissue-specific spin-spin interactions and B_0 inhomogeneities. Such re-equilibrium results in an exponential decay of (net) magnetization within the transverse plane and is referred to as T2*. By applying a second pulse we can counteract the B_0 inhomogeneities and also measure the amplitude of signal decay caused only by tissue-specific spin-spin interactions, resulting in a T2 measure in the transverse plane. In addition to measuring T2 and T2* decay within the transverse plane, we can also measure the growth of magnetization back to the original equilibrium state within the longitudinal plane. This recovery of decay within the longitudinal plane takes longer, compared to that of the transverse, and provides a measure of T1 relaxation which is the time constant of the exponential growth of the signal.

Specifically, T1 and T2 weighted images are sensitive to the different tissue characteristics, accentuating the different contrasts through variations in spin-lattice interactions (T1-weighted) and spin-spin interactions (T2-weighted). This is achieved by performing acquisition protocols that use specific repetition times (TR) and time to echo (TE) to alter the weighting of the image, where multiple images with different weightings are then used to fit

the signal intensities to each individual pixels to form a map. T1 and T2* maximise the contrast between different types of tissue and can provide a measure of intracellular fluid and iron content, respectively. T1 provides an MRI indicator of regional tissue water and lengthens with accordance to increases in extracellular fluid and reflects the degree of tissue inflammation and fibrosis. However, T1 is influenced by the presence of iron. Corrected T1 (cT1) is an algorithm developed by the imaging analysis company Perspectum (Oxford, UK) that removes the bias introduced by elevated iron on T1 by correcting for iron derived from the T2* signal, see Fig. 1.3 for an example cT1 map of the liver [43]. Proton density fat fraction (PDFF) is another MRI-derived biomarker that provides a quantitative measure of fat infiltration (steatosis), see Fig. 1.3. PDFF is quantified by measuring the ratio of the density of triglycerides in each voxel against mobile water and is expressed as an absolute percentage (%).

cT1 and PDFF have been shown to correlate well with the characteristic histological features of NAFLD [39], [44], predict liver-related outcomes [45], and accurately stratify patients with NASH and those at risk of developing NASH [46], [47]. Research published by Perspectum has revealed a liver PDFF >5.6% and liver cT1 >800ms demonstrates high diagnostic accuracy for stratifying patients with NASH, predicting liver-related outcomes and classifying between NAFLD and NASH [47], [48].

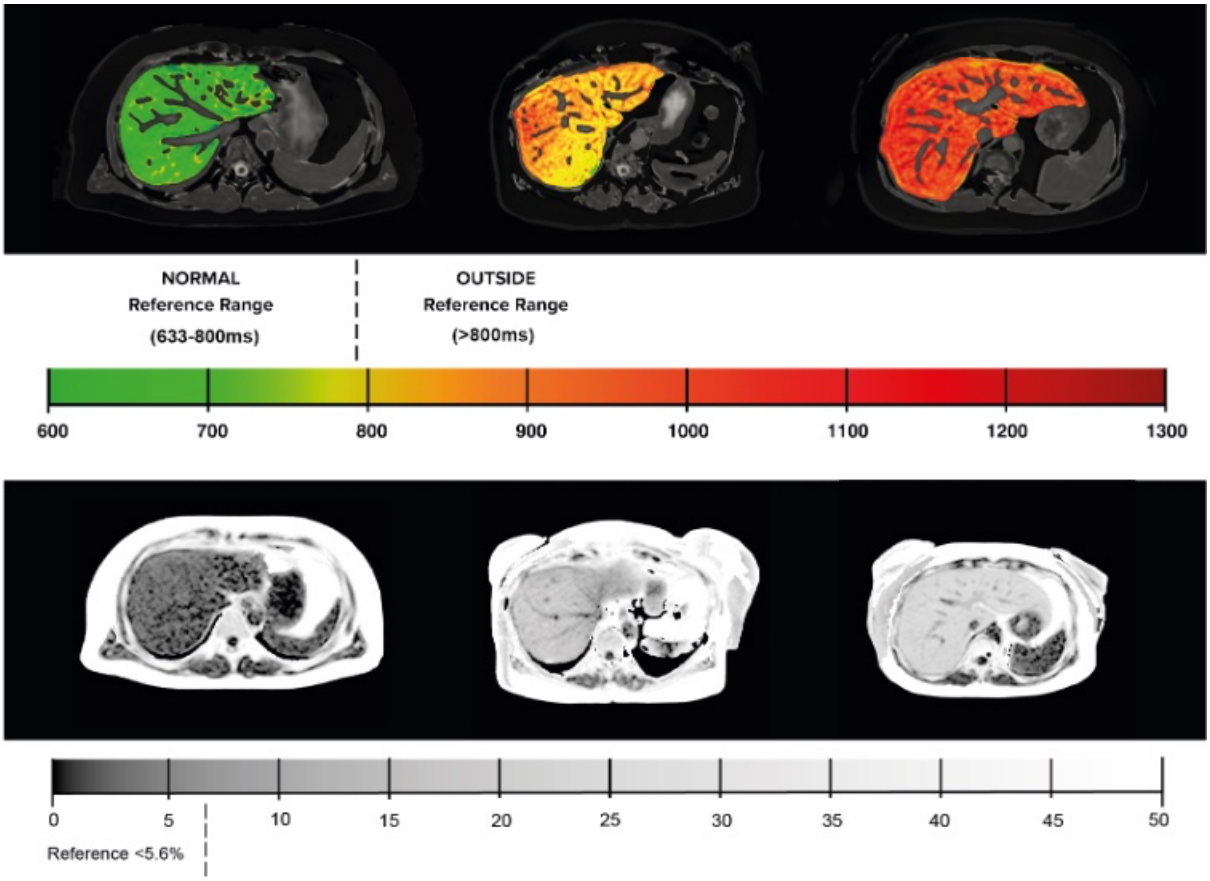


Figure 1.3. Example cT1 (top) and PDFF (bottom) images of the liver, with reference ranges.

Body composition can be delineated (or *segmented*) from a single two-dimensional MRI slice positioned at the third lumbar (L3) vertebrae, see Fig. 1.4. The L3 section is typically selected because it has been shown to strongly correlate with whole-body skeletal muscle distribution and accurately estimate total subcutaneous (SAT) and visceral adipose tissue (VAT) volumes [49]–[51]. In this Thesis, measures of SAT, VAT and skeletal muscle were assessed manually from the L3 slice and are reported as cm^2, see Fig. 1.5. Measures of skeletal muscle were then indexed to the height of the participant (cm^2/m^2), producing a measure of skeletal muscle index (SMI).

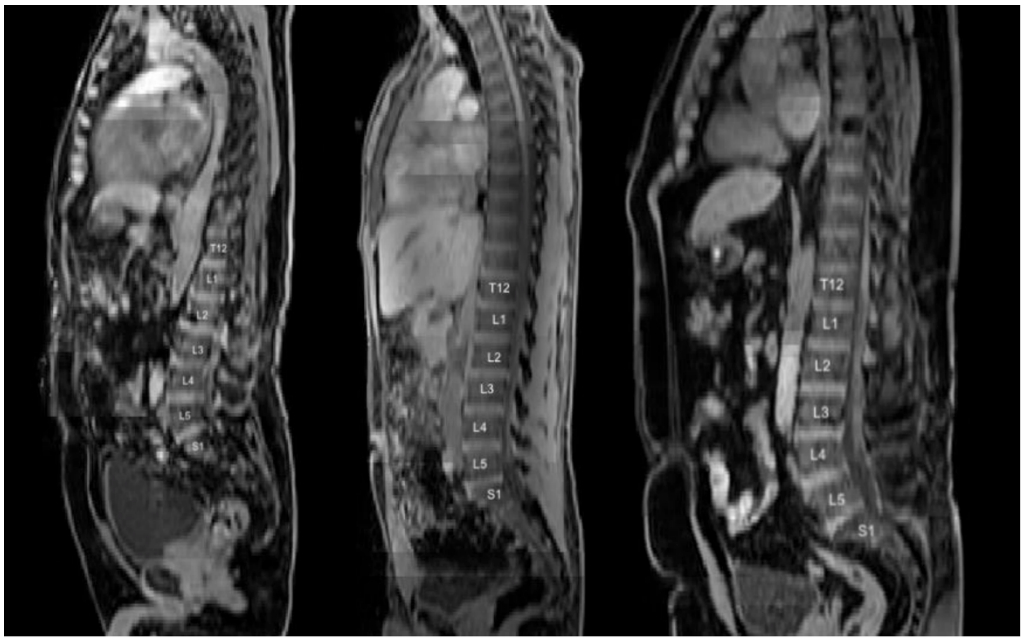


Figure 1.4. L3 slice positioning (numbered) for extraction of body composition segmentation.

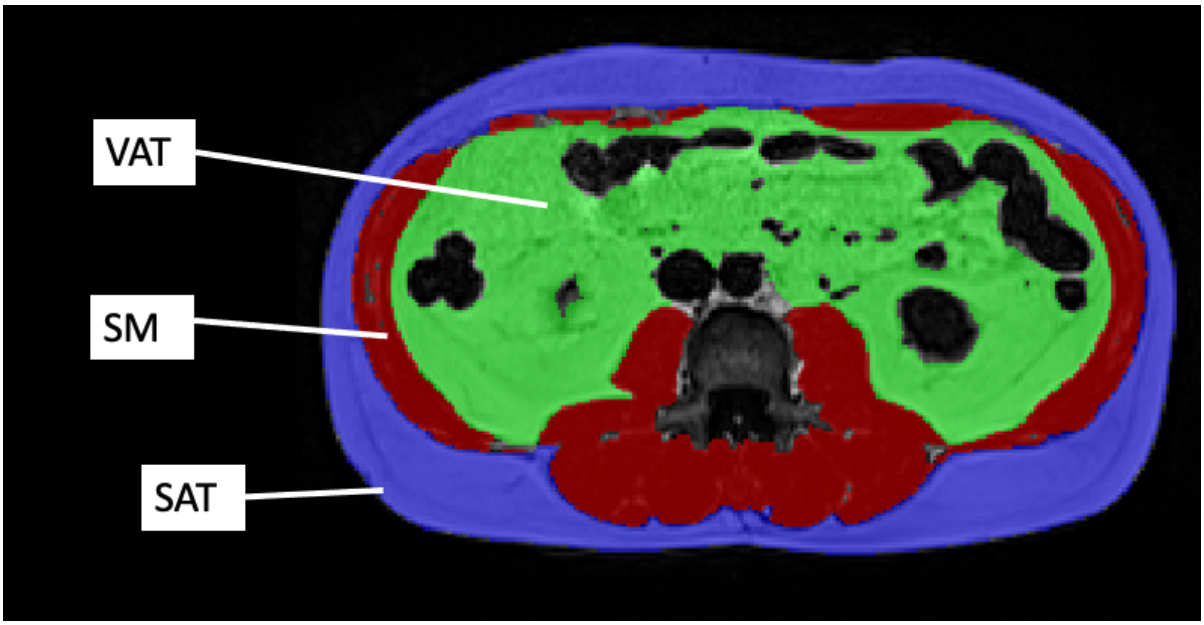


Figure 1.5. Example segmented body composition from a 2D-slice positioned at the third lumbar vertebrae. VAT = visceral adipose tissue; SAT = subcutaneous adipose tissue; SM = skeletal muscle.

1.6. Interventions for the management of metabolic disease

Although difficult, it is possible to manage or reverse metabolic disease through a combination of interventional strategies. The first line of treatment is typically energy modulation by increasing energy output and/or decreasing energy intake through lifestyle changes. Just a small reduction in body weight has been shown to elicit improvements in metabolic health. For example, a 5% weight loss results in a 25% to 43% reduction in intrahepatic triglycerides, while a 10% weight loss reduces NAFLD activity score and regression of liver fibrosis [52]–[54].

A 14-day carbohydrate (CHO) restriction intervention in people with obesity has reported to reduce liver fat and *de novo* lipogenesis while increasing serum B-hydroxybutyrate [55]. Furthermore, rather than calorie restriction, prescription of a Mediterranean diet (MEDd) has been proven through numerous large multicentre trials to reduce the incidence of NAFLD [56], CVD [57], T2D [58] and the metabolic syndrome [59]. The MEDd is high in monounsaturated fatty acids but low in sugars, achieved through consumption of foods such as fresh vegetables, nuts, oils, and fish. Modulating metabolism through increasing energy expenditure has also shown to improve both non-invasive measures and biopsy proven steatosis independent of significant weight loss [35], [60], [61], although the optimal mode of training is inconclusive. Moderate to vigorous physical activity has proven to reduce CVD risk and improve glycaemic control in people with T2D [62], while simply walking for at least 30 minutes per day has shown to reduce T2D risk by an approximately staggering 50% [63].

While the mechanisms mediating the improvements in metabolic health through energy modulation are not fully understood, there are similarities between the target mechanisms

used in pharmacological therapies. For example, physical activity and the most frequently prescribed first line T2D therapy – Metformin – both work by increasing mitochondrial oxidative capacity within the liver and adipose tissue, shifting substrate metabolism towards fat oxidation [64]. However, pharmaceutical interventions target direct mechanisms beyond those conferred by energy modulation alone. For example, two of the most common types of T2D therapies - Sodium-Glucose co-transporter-2 inhibitors (SGLT2-Is) and Glucagon-Like Peptide-1 receptor agonists (GLP1-Rs) - work by reducing the amount of glucose reabsorbed by the kidneys and increase the level of incretin hormones, respectively. There is also an increasing usage of traditional T2D therapies for the management of secondary disorders. For example, the T2D treatment class named SGLT2-Is have also demonstrated to reduce major CVD events including myocardial infarction and heart failure vs placebo, significantly reduce liver fat and promote greater histological resolution of NASH [65]–[67].

1.7. The clinical challenge

The care and management of people with metabolic disease is incredibly complex, and increasingly so. Clinicians must not only assess their patient's current health and scope of comorbidities, but also seek to deliver an optimal treatment plan while balancing the risk profiles and contraindications of individual therapies. Furthermore, people with T2D are often prescribed a combination of therapies, from first to fourth line prescriptions that constitute the layering of therapies (the first being their initial treatment), further compounding the complexity of clinical decision making. For example, Metformin as a first-line treatment was prescribed to 92% of patients between 2010-2017, while prescription of SGLT2-Is are prescribed in 17% of first to fourth line therapies [68]. When making clinical decisions for their

patient, clinicians are also often guided from outcome data derived from clinical trials where the participant demographic may not be representative of their patient. Consequently, clinicians are often forced to adopt a 'trial and error' approach to tailoring their patients' care.

The following examples are drawn from routine clinical practice as representative of the kinds of the daily reasoning and decision making by clinicians. They also illustrate the complexity of the reasoning tasks that the AI system developed in this Thesis seeks to contend with:

Example case study 1

40-year-old male living with obesity, diagnosed with type 2 diabetes (T2D) after a routine health assessment (bloods & anthropometrics) to monitor cardiovascular disease risk after an acute cardiac event 1 year prior.

This is a high-risk individual with newly diagnosed T2D. The treatment decision is whether to initiate single (Metformin, per guidelines) or dual (Metformin + SGLT2-I, shown to reduce CV risk) therapy from the outset.

Example case study 2

60-year-old female living with obesity with a 10-year history of T2D and 5-year history of concomitant fatty liver disease. Recent diagnostics reveal evidence of non-alcoholic steatohepatitis (NASH). The patient was initiated on Metformin for 8 years prior but failed to control blood glucose and has reported side effects of nausea and gastrointestinal (GI) upset.

The treatment decision is whether to move the patient onto combinatorial therapy (Metformin plus Semaglutide) or keep the patient on Metformin only. Common side effects of Semaglutide also include nausea and GI upset and the clinician must weigh up the cost/benefit of clinical inertia versus treatment intensification to manage glycaemia and NASH against the likelihood of worsening the patients' side effects.

In the above example case studies, the clinician must pose intrinsically counterfactual questions about their patient - what would happen to my patient's liver disease activity if they reduced their blood glucose concentrations? What is the risk that my patient will develop advanced liver disease if they continue to accumulate greater visceral adipose tissue and liver fat deposition? Here the term 'counterfactual', defined as 'counter to the facts', relates to a fictional outcome that could have occurred if a different direction or decision had preceded the outcome. For example, 'would I still have a headache, had I taken the aspirin?' is a counterfactual question, for the reality is that the aspirin was not taken, and the counterfactual is posing a question within an alternative world about the headache, had the aspirin been taken. As I discuss in further detail in Chapter 2, the current statistical techniques used in medical research are not equipped to address such questions. This Thesis is concerned with trying to unravel such complexity of patient trajectories and potential areas of intervention using a novel Bayesian-network approach, introduced in detail in Chapter 3.

1.8. Exploring T2D phenotypes in the UK Biobank

Tying together the contents of this Thesis thus far, the first project of my DPhil was to apply Perspectum's imaging capabilities to explore the phenotype of people with T2D. This project

also served as a means of generating a database for the first application of Bayesian-networks set out in Chapter 4. The aim of that study was to explore differences in MR-derived measures of fat deposition in BMI, age, and sex matched participants with and without T2D to see whether those with T2D presented with a distinct phenotype, despite a similar obesity status. This work was published in the journal *Obesity* in 2022. Much of the of the written publication has been omitted here, for the full paper please see the following citation:

Waddell, T., Bagur, A., Cunha, D., Thomaides-Brears, H., Banerjee, R., Cuthbertson, D.J., Brown, E., Cusi, K., Després, J.P. and Brady, M., 2022. Greater ectopic fat deposition and liver fibroinflammation and lower skeletal muscle mass in people with type 2 diabetes. *Obesity*, *30*(6), pp.1231-1238.

1.8.1. Data selection

MRI data and participant demographics were acquired retrospectively from the UK Biobank (UKBB) study. UKBB is a general population-based cohort that conducted MRI scanning of the brain, heart and abdomen in people aged 40 to 69 years from across the UK. Accompanying participant demographic data, collected through an online data access application (application number: 9914), included measures of age, sex, BMI, waist circumference, and routine biochemistry results, including serum/plasma high-density lipoprotein (HDL), triglycerides, aspartate aminotransferase (AST), alanine transaminase (ALT), and glycated haemoglobin (HbA1c) concentrations.

In total, there were 318 participants with self-reported T2D and accompanying MRI data enrolled in the UK Biobank study. Of these participants, 237 (162 male individuals, 75 female individuals) were selected, with an age range of 41 to 70 years and a BMI range of 18.5 to 45.9. Each participant with T2D was then paired with a participant with no self-reported diabetes of any type, matched for sex, age ($\pm$5 years), and BMI ($\pm$1 kg/m^2). Data from 474 participants (237 with T2D and 237 without T2D) were collected for MRI analysis. Though widely used throughout the UK Biobank, the collection of self-reported clinical outcomes can lead to inaccuracies within this analysis, such as under/over reporting of specific diseases or miss-labelling. This can result in a miss-represented cohort and therefore results not reflective of the true disease endotype/phenotype. Furthermore, it is important to highlight the low sample size as a key limitation of this analysis (and throughout this Thesis). Low sample sizes can lead to an increased risk of low statistical power and risk of type -2 error, resulting inaccurate results and inferences. Indeed, the results throughout this Thesis should be viewed with caution and instead seen as an example of how Bayesian-networks can be applied to model biological systems, as opposed to definitive answers regarding the specific phenotypes and endotype of disease. Lastly, the UK Biobank cohort itself is notably biased (for example, a majority white Caucasian demographic), meaning the results gleaned from such data may not be generalisable to the UK population.

1.8.2. Statistical analysis

Descriptive statistics, which report medians and interquartile ranges (IQR), are displayed in order to summarise participant characteristics. Wilcoxon and X^2 tests were performed to calculate the significance (p value) of differences in biomarkers between participants with T2D and those without T2D. To calculate the association between participant biomarkers and

T2D, we performed univariate and stepwise multivariate logistic regression modelling. Model inputs included the following variables: pancreatic fat (PDFF), liver fat (PDFF), liver cT1, VAT, SAT, SMI, HDL-c, triglycerides, AST, ALT, BMI, waist circumference, age, and sex. Risk scores and confidence intervals (CI) were calculated against the odds ratios of participants who had self-reported T2D. Spearman correlation tests were calculated to investigate the relationships between biomarkers.

1.8.3. Results

Biochemistry

As anticipated, participants in the group with T2D reported significantly greater HbA1c values ($p < 0.001$) compared to participants without T2D. We observed no significant differences in triglycerides ($p = 0.26$), AST ($p = 0.35$), or ALT ($p = 0.11$). Prevalence of elevated AST and ALT was similar between the groups with T2D and without T2D ($p = 0.71$, $p = 0.23$, respectively).

Table 1.1. Anthropometric and demographics of study population. Data is median [IQR], no significant differences were observed.

	T2D	Non-T2D	p-value
N	131	135	-
Age (yrs)	56 [52 – 63]	57 [53 – 63]	0.66
Ethnicity (% White British)	84%	96%	0.37
Sex (% male)	68%	73%	0.67
BMI (kg/m^2)	29.4 [27 – 33.9]	29.7 [26.8 – 33.5]	0.97
WC (cm)	99 [91 – 106]	97 [91 – 109]	0.56
DBP (mmHg)	84 [77 – 90]	85 [78 – 92]	0.34
SBP (mmHg)	142 [129 – 151]	142 [131 - 155]	0.42

Self-reported hypertension (%)	50	42	0.40

MRI-derived measurements: liver, pancreas, and body composition metrics

Participants in the group with T2D reported significantly higher MRI-measured liver fat (p = 0.011) and liver cT1 (p = 0.019). We observed no significant differences in MRI-measured pancreas fat (p = 0.22), VAT (p = 0.35), or SAT (p = 0.43). Prevalence of significant pancreas steatosis and liver steatosis was high but not significantly different between participants with or without T2D (38% vs. 40%, 64% vs. 47%, respectively). However, we observed that prevalence of NASH was significantly greater in participants with T2D, and obesity compared with participants without T2D. Furthermore, significantly lower SMI values were observed in the group with T2D (p = 0.003).

Table 1.2. Biochemical and imaging characteristics. Data is median [IQR], significant p-values in bold.

	T2D	**Non-T2D**	**p-value**
Biochemical			
HbA1c (mmol/L)	42.9 [38.6 – 51.3]	36.1 [32.9 – 38.3]	**<0.001**
HDL (mmol/L)	1.1 [1 – 1.3]	1.3 [1.1 – 1.5]	**<0.001**
Triglycerides (mmol/L)	1.8 [1.3 – 2.6]	1.7 [1.2 – 2.4]	0.26
ALT (IU/L)	29.9 [20.6 – 36.6]	25.9 [19.3 – 34.6]	0.11
AST (IU/L)	25.3 [21.6 – 30.7]	26.6 [22.9 – 32.5]	0.35

Imaging			
VAT (cm^2)	219.2 [170.9 – 297.4]	215.1 [149.4 – 278.8]	0.35
SAT (cm^2)	244.3 [185.5 – 322.9]	271.4 [181.3 – 366.1]	0.43
SMI (cm^2/m^2)	45.2 [38.1 – 52.9]	50.6 [42.6 – 55.9]	**0.003**
Liver cT1 (ms)	730 [685 – 786]	709 [671 – 753]	**0.019**
Liver fat (%)	7.4 [4.1 – 13.8]	5.3 [2.7 – 10.6]	**0.01**
Pancreas fat (%)	6.1 [4.2 – 9.6]	5.8 [3.1 – 8.9]	0.22

Table 1.3. Abnormalities of organ structure/function, no significant differences were observed.

	People with T2D	**Controls**	**p-value**
NAFLD (%)	64	47	0.10
NASH (%)	11	7	0.35
Elevated pancreatic fat (%)	38	40	0.34
Elevated AST (%)	16	14	0.71
Elevated ALT (%)	16	10	0.23

Correlation analysis

cT1 showed moderate positive correlations with BMI and VAT (rho 0.41, p = <0.001; rho 0.41, p <0.001, respectively). BMI showed moderate positive correlations with all MRI-measured metrics of fat, except for pancreas fat, which correlated weakly. VAT positively correlated with MRI-measured liver fat but correlated only weakly with pancreas fat.

Table 1.4. Spearman's correlation analysis, significant p-values in bold.

Biomarker interaction	Rho	p-value
BMI - VAT	0.52	**<0.001**
BMI – SAT	0.69	**<0.001**
BMI – SMI	0.34	**<0.001**
BMI – Liver fat	0.42	**<0.001**
BMI – Liver cT1	0.41	**<0.001**
BMI – Pancreas fat	0.25	**<0.001**
VAT – Liver fat	0.54	**<0.001**
VAT – Liver cT1	0.41	**<0.001**
VAT – Pancreas fat	0.28	**<0.001**
VAT - SAT	0.27	**<0.001**
SAT – Liver fat	0.33	**<0.001**
SAT – Liver cT1	0.30	**<0.001**
SAT – Pancreas fat	0.19	**0.001**
HbA1c – Liver fat	0.19	**0.001**
HbA1c – Liver cT1	0.24	**<0.001**
HbA1c – Pancreas fat	0.19	**0.001**
HbA1c – BMI	0.19	**0.001**
HbA1c – VAT	0.15	**0.016**
HbA1c – SAT	0.04	0.47
HbA1c – SMI	-0.08	0.21
HbA1c – HDL	-0.23	**<0.001**

AST – Liver cT1	0.16	**0.005**
ALT – Liver cT1	0.28	**<0.001**

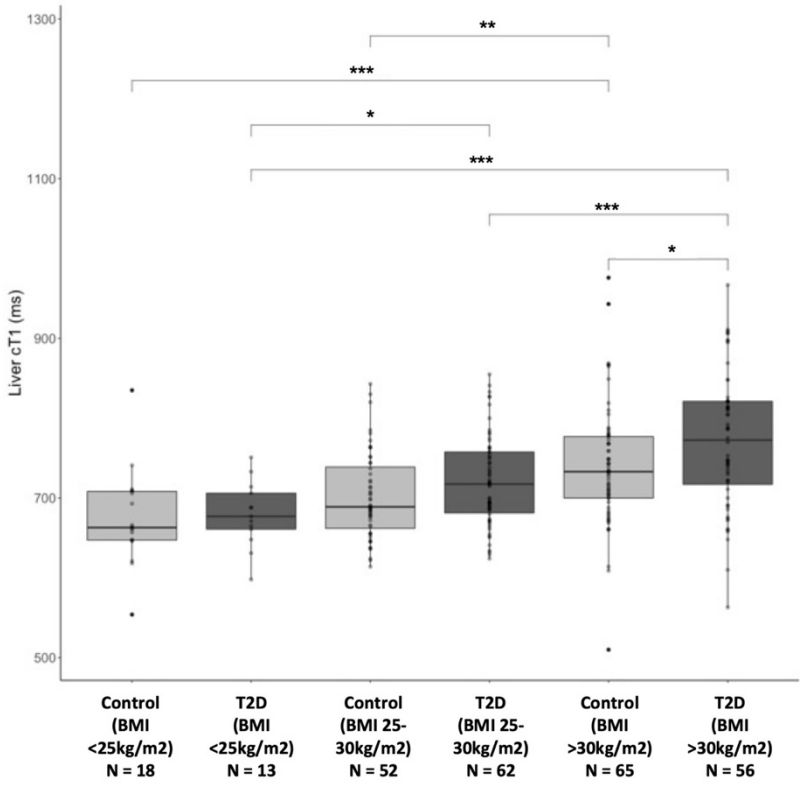


Figure 1.6. Liver cT1 (ms) in participants stratified by obesity and diabetes status.

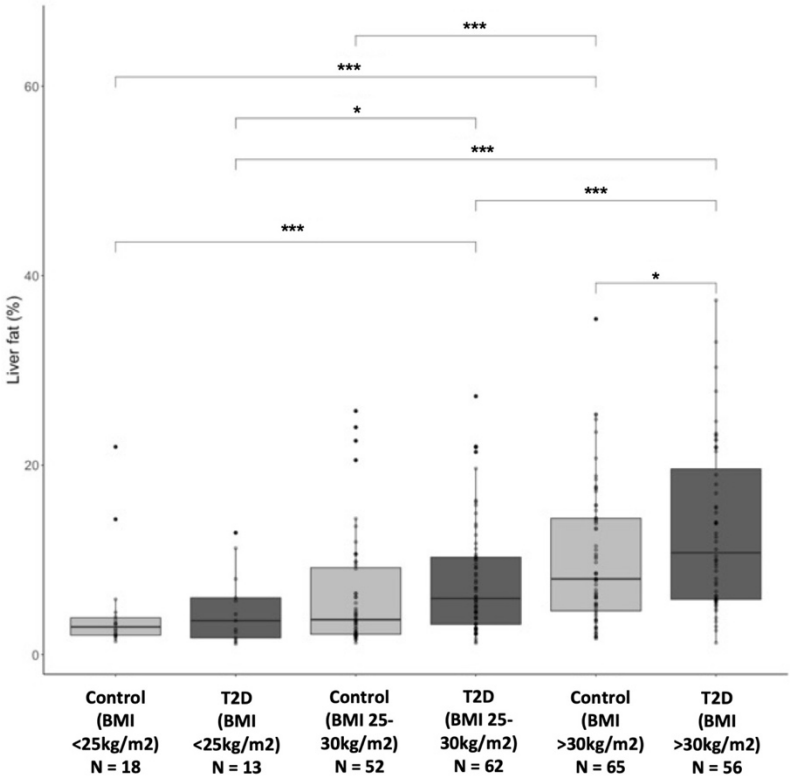


Figure 1.7. Liver fat (PDFF %) in participants stratified by obesity and diabetes status.

Univariate and multivariate analysis

Univariate modelling revealed that either liver fat >5% (p = 0.01) or >10% (p = 0.01) and elevated liver cT1 (p = 0.01) were associated with a diagnosis of T2D. We observed a significant reduction in SMI (p = 0.01) and HDL-c (p <0.001) with a diagnosis of T2D. Stepwise multivariate analysis, including the significant variables listed earlier, revealed that reduced SMI (p <0.001) and HDL-c (p <0.001) remained significantly associated with T2D, with an odds ratio of 0.94 (95% CI: 0.91-0.97) and 0.15 (95% CI: 0.06-0.40), respectively.

1.8.4. Discussion

In this project, we report MR-derived measures of the liver, pancreas, and body composition in participants with and without T2D from the UK Biobank. We highlight that participants with

self-reported T2D have evidence of significantly greater liver damage, as measured by significantly greater levels of liver fibroinflammation as determined by MRI-derived measures of liver fat and cT1, with significantly reduced skeletal muscle mass, as determined by MRI-derived measurement of skeletal muscle index (SMI) compared to BMI, age, and sex-matched participants.

In the following Chapter (Chapter 2), however, I illustrate how the traditional statistical techniques, such as those used in the *Obesity* publication, are intrinsically limited despite their extensive use in medical research. In doing so, I lay the foundations for introducing the alternative approach explored in this Thesis - the field of Bayesian-networks and probabilistic inference.

Chapter 3 details the methodology of Bayesian-networks, describing the semantics behind the graphical modelling of Bayesian-networks and the application of probabilistic inference for estimating counterfactuals. The literature surrounding Bayesian-networks within the medical domain is explored, highlighting their utility for modelling biological systems, revealing specific disease mechanisms, and identifying novel therapeutic strategies. This Chapter closes with highlighting the novelty of this Thesis work, specifically the application of Bayesian networks to study the complexity of fat deposition and metabolic disease risk.

Chapter 4 illustrates the first application of Bayesian-networks. This short project is a re-analysis of the *Obesity* publication and explores the role of fat deposition in driving diabetic-status hyperglycaemia. This work highlighted that hyperglycaemia was driven by ectopic fat deposition in multiple, as opposed to single, sites, and the potential protective role of

subcutaneous adipose tissue against impaired glucose metabolism. This project was accepted for oral presentation at the 2023 European Association for the Study of Diabetes.

Chapter 5 then applies Bayesian-networks to estimate the effect of projected increases in obesity on incident liver disease in the UK. By 2040, it is estimated that 37% of the UK population will be obese, surpassing that of healthy weight persons. In a cohort of 27,002 participants from the UK Biobank, this analysis estimated that such increases in obesity would lead to an 8%, 5% and 3% increase in the probability of severe fatty liver disease, moderate fatty liver disease and NASH, respectively. This work was accepted for oral presentation at the 2023 International Conference on Medical and Health Informatics and poster presentation at the 2023 European Association for the of Liver Disease conferences.

Chapter 6 dives deeper into the mechanistic drivers of liver disease by applying Bayesian-networks to a cohort of participants with and without type-2 diabetes (T2D). People with T2D are significantly more likely to progress from NAFLD to NASH and demonstrate greater liver disease activity compared to BMI-matched persons without T2D. This project reveals the role of elevated visceral adipose tissue and elevated blood glucose in driving the probability of NASH. This work was published in 2023 in the journal *Frontiers in Endocrinology.*

Chapter 7 applies Bayesian-networks to study the well documented association between obesity and the greater risk of hospitalisation following acute COVID-19. This work highlighted that ectopic fat deposition, namely within the liver, resulted in a marked increase of hospitalisation risk across obesity classifications and recommended potential therapeutic strategies. This work was published in 2023 in the journal *Frontiers in Bioinformatics.*

Chapter 8 begins by summarising and highlighting the key contributions of the work contained within this Thesis. I then detail areas of future work for the application of Bayesian-networks for modelling multi-organ pathologies and the potential barriers of adoption. I describe potential projects that could follow the work contained within this Thesis and their relative further contributions.

2

LIMITATIONS OF TRADITIONAL STATISTICAL APPROACHES FOR STUDYING BIOLOGICAL SYSTEMS

2.1. Cause, effect, and spurious correlations

It is hard to imagine any professional working in the medical field, or any field for that matter, *not* dealing with cause and effect. Our ability to identify cause and effect has shaped humankind for thousands of years, allowing the survival and proliferation of our race. From providing our ancestors with the ability to find nourishment, to the days of Hume and deriving causal connections from observation, identifying cause and effect has provided the framework for explaining how our world operates.

As intuitive humans, we know that the rooster's crow does not *cause* the sun to rise, but how easy is this to formulate into a mathematical construct? Moreover, without the appropriate techniques, how can we expect our decision support systems to know that two such variables are not *causally* connected? The inclusion of causality in science has, to say the least, not been without resistance. For example, in 1913 the philosopher Bertrand Russell stated:

> *"All philosophers, imagine that causation is one of the fundamental axioms of science, yet oddly enough, in advanced sciences, the word 'cause' never occurs… The law of causality, I*

believe, is a relic of bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm"

and more venomously, in 1911 Pearson stated:

"Beyond such discarded fundamentals as 'matter' and 'force' lies still another fetish amidst the inscrutable arcana of modern science, namely, the category of cause and effect" [69]

Traditional statistical concepts such as correlations, regression and standard deviation are credited to Sir Francis Galton, who alongside Karl Pearson sought to establish statistics as a solid, mathematics-based discipline. Followed by Ronald Fisher, who introduced the statistical constructs of variance and the 5% significance threshold, the field of statistics was tied to observing what could only be derived formally from data - the 'frequentist' approach to probability. From this perspective, the term 'cause' and the study of probability with an unknown event seemed unmathematical, judgemental, and was confined to the fringes of statistics, only rising in popularity again in the 1970's with Sir Adrian Smith's work on Bayesian statistics. The overarching theme which continues in force is denial of the need to study causation as a well-defined construct independently of correlation.

But why are correlations limited, and specifically for the study of biological systems? The answer is three-fold; firstly, biological systems are inherently *causal* in nature, not merely associational (as denoted by correlations); secondly, clinical biomarkers are often influenced by several other variables, as opposed to a simple 1-1 interaction, and thirdly; correlations do not possess the necessary language for asking the type of counterfactual questions posed by

clinicians. For example, clinicians are often seeking to assess their patients' health dependent on whether they were on one treatment or another – is treatment A or B better for their patient, given their current health status and medical history? Of course, they cannot administer both, and so clinicians are often weighing up inherently counterfactual outcomes; typically based on previous experience, current literature, or advice from peers.

2.2. Spurious correlations

We can illustrate the first two points by providing an example of a *spurious* correlation, a term given to two variables that appear correlated, despite no plausible *causal* connection. The technical foundation of a spurious correlation is explored in the next Chapter; however, a brief explanation is warranted here. The example provided here is the well-established correlation between a nation's per capita chocolate consumption and its number of Nobel prize winners (Fig. 2.1). Should a society initiate an education policy to provide students with a daily chocolate supply? Alas, such a policy would not work, because a nation's chocolate consumption and number of its Nobel Prizes are only *conditionally* dependent given a third variable – the wealth and location of that nation. This spurious correlation is formally illustrated in Fig 2.1.

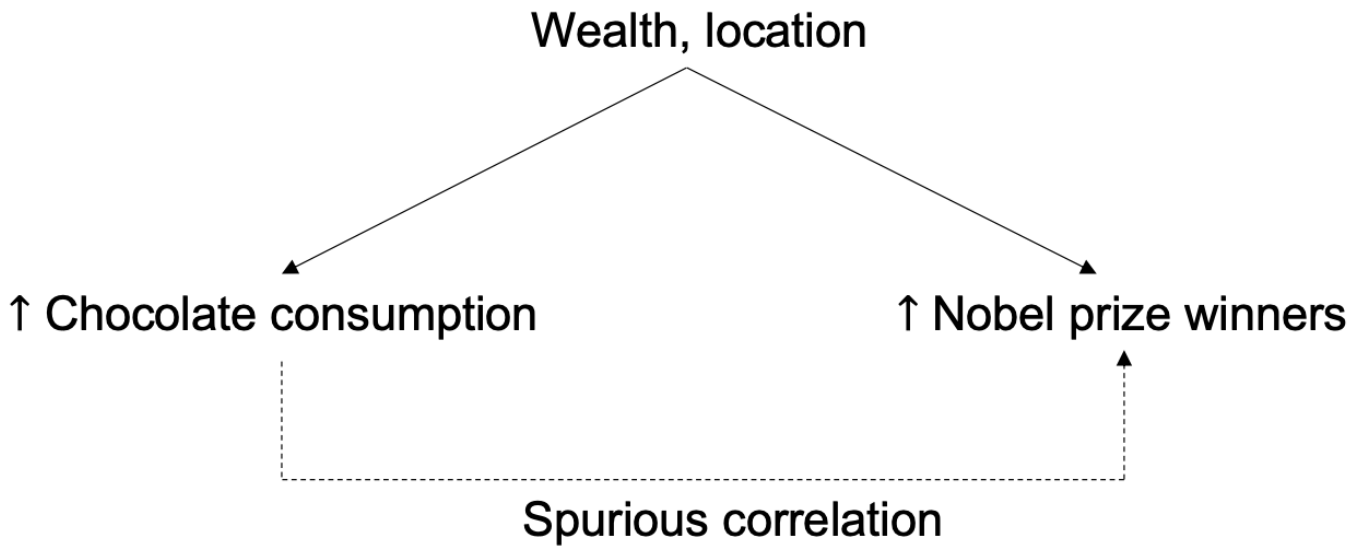


Figure 2.1. Illustration of the spurious correlation between a nation's chocolate consumption and number of Nobel prize winners.

It is possible that such spurious correlations are widely distributed throughout the medical literature, a recent example being the correlation between increased Vitamin D and lower COVID-19 infection and mortality [70]. Likely eager for some good news around the pandemic, this correlation was inflated by news outlets claiming that Vitamin D may reduce mortality risk by 50%. However, such a correlation is likely spurious, with [71] reporting that the coefficient of Vitamin D lost its significance when controlling for the life expectancy of the country. The authors concluded that life expectancy was a statistically significant predictor of both mortality and infections at the country level, noting that such association could have been driven by better healthcare systems and a greater reporting of COVID-19 related deaths in those countries. Of course, we cannot rule out that a causal connection does exist, but this

must be studied using appropriate techniques and not through secondary analysis from cross-sectional studies.

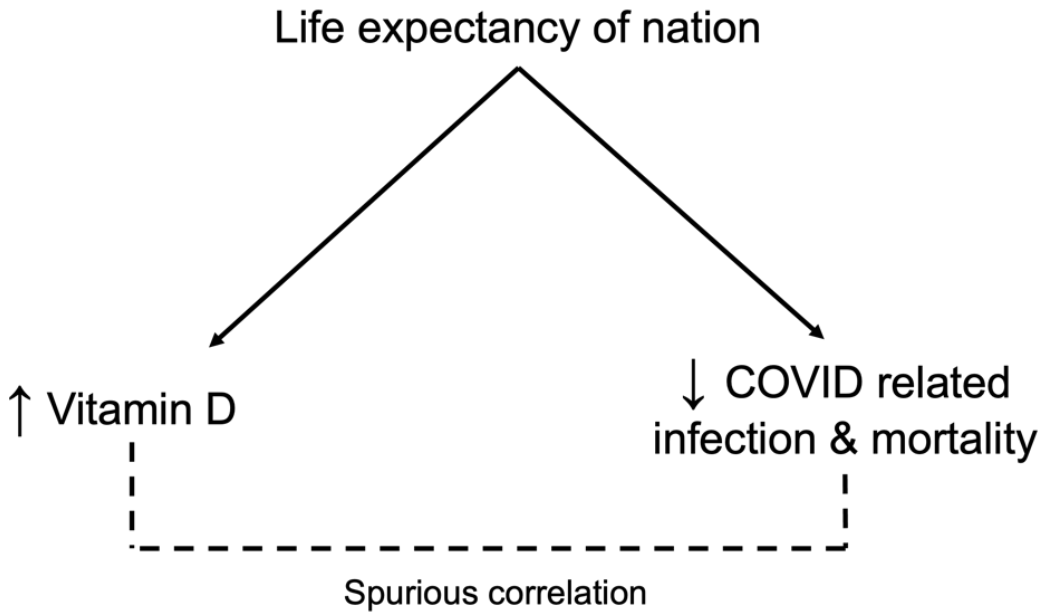


Figure 2.2. Potential spurious correlation between Vitamin D consumption and COVID-19 related infection and mortality.

2.3. Complexity of biological systems

The second reason why correlations are limited is due to the inherent complexity of biological systems, as highlighted in the Introduction of this Thesis. Metabolic dysfunction involves a complex interplay of various biological processes, such as ectopic fat deposition and hormonal disturbances that interact in myriad ways. Such complexity, if modelled using traditional approaches such as correlations, results in an inevitably large set of tangled

statistical relationships that are not only difficult to interpret but to derive clinically meaningful inferences from. For example, by pulling out just four example correlations between BMI, VAT, SAT, and SMI from the long list presented in the *Obesity* publication, we can highlight the following three limitations:

1) The first limitation is related to that of *causality*, where in the example of BMI – VAT, it is not explicit whether BMI causes the change in VAT, or vice versa. Of course, one's BMI is constituted by one's height and weight, meaning that in this example BMI would be conditionally dependent on VAT (Fig 2.4).
2) The second limitation is related to that of confounding, for we know that the BMI (specifically, one's weight) is influenced by more than just VAT. The correlation BMI – VAT therefore does not make explicit the effect of SAT and SMI also on BMI (Fig 2.5).
3) The third limitation is related to lack of clinical information embodied in the correlation. For example, as highlighted in the introduction of this Thesis, those with MHO can present with high measures of SAT, yet low measures of VAT, which is not captured through the use of correlations.

Table 2.1 Example correlations taken from the *Obesity* publication.

Biomarker interaction	**Rho**	**p value**
BMI - VAT	0.52	<0.001
BMI – SAT	0.69	<0.001
BMI – SMI	0.34	<0.001
SAT - VAT	0.27	<0.001



Figure 2.3. Moving from association (left) to explicit direction of conditional dependence (right).

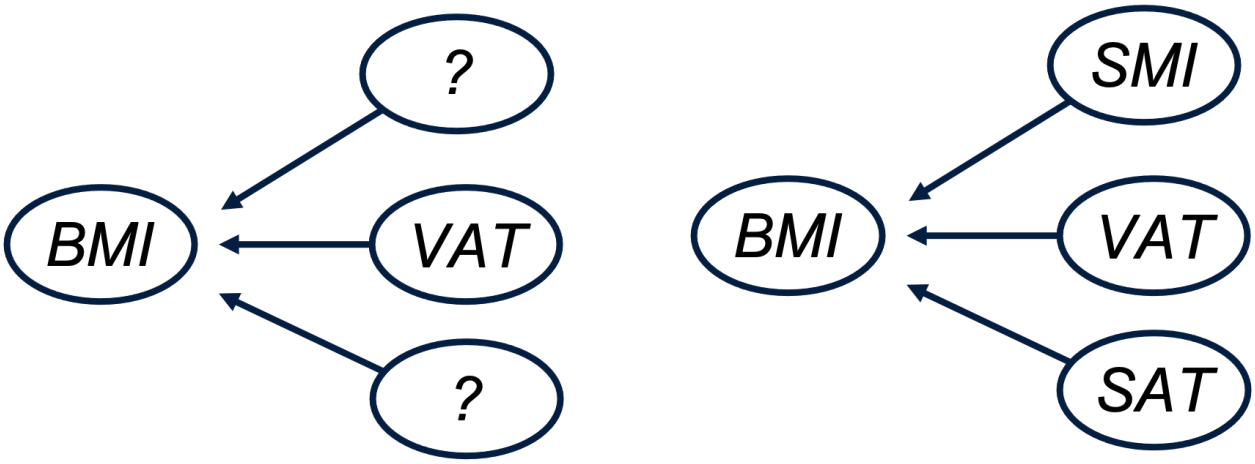


Figure 2.4. Moving from a simple 1-1 interaction between two biomarkers (left), to explicitly encoding multiple conditional dependencies (right).

Lastly, I return to the example 'Case study 2' presented in Chapter 1:

Example case study 2

60-year-old female living with obesity with a 10-year history of T2D and 5-year history of concomitant fatty liver disease. Recent diagnostics reveal evidence of non-alcoholic steatohepatitis (NASH). The patient was initiated on Metformin for 8 years prior but failed to control blood glucose and has reported side effects of nausea and gastrointestinal (GI) upset.

The treatment decision is whether to move the patient onto combinatorial therapy (Metformin plus Semaglutide) or keep the patient on Metformin only. Common side effects of Semaglutide also include nausea and GI upset and the clinician must weigh up the cost/benefit of clinical inertia versus treatment intensification to manage glycaemia and NASH against the likelihood of worsening the patient's side effects.

Here, the clinician must pose an inherently counterfactual question about their patient: given their current health status, what would the potential outcomes be if their patient was receiving combinatorial therapy? Traditional statistical approaches, such as those used in the *Obesity* publication, are simply not equipped to address such questions. As illustrated in Pearl's Ladder of Causality in Fig. 2.6, traditional statistics such as correlations or pairwise comparisons are located on the bottom 'association' rung of the ladder of *seeing* and *observing*. Rather, the question posed in 'Case study 2' is located at the top rung of the ladder of 'counterfactuals', the activity of *imagining* and *understanding*. The difference can sometimes be subtle and is worth expanding here. With association, we are asking the

question of 'if I see *x*, what does that tell me about *y*?' – or 'what does a symptom tell me about a disease?'. In such questions, we are tied to the observable data presented to us, the *seeing* of the data. With counterfactuals, we are instead posing the question of 'What happens to *x*, if I had changed *y*?' - or 'would I still have a headache, had I taken the aspirin?'. In order to answer such questions posed by clinicians we need to employ techniques that can talk their language. As detailed in the following Chapter, Bayesian-networks are one such technique, and have been gaining popularity for unveiling the complexity of biological systems and novel therapeutic targets.

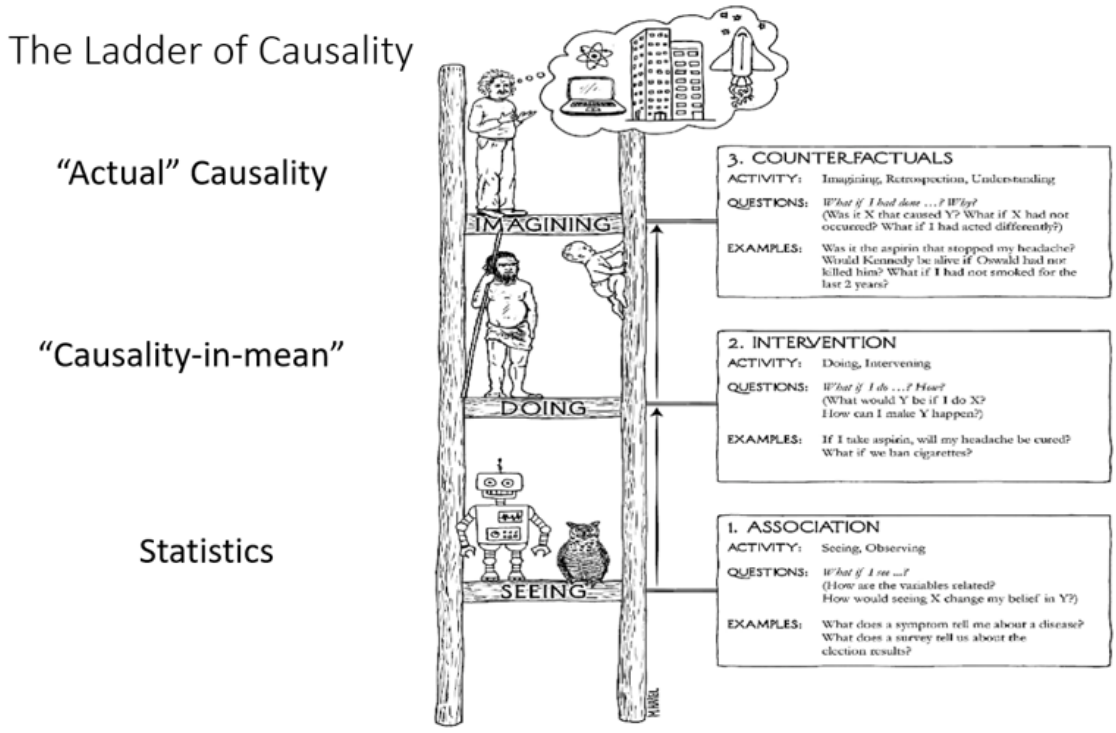


Figure 2.5. Pearl's ladder of causality, from 1) association to 3) counterfactuals, taken from Pearl & Mackenzie (2018) *The Book of Why [72].*

2.4. Foundations of Bayesian-networks

Returning to Pearl's ladder of causality illustrated in Fig. 2.6, we see that a main component of counterfactuals is the ability to *understand*, much like a clinician seeks to understand their patients' needs. That is, they must understand the set of premises and rules that governs their clinical decision making and the outcomes thereof, for example: "IF my patient has obesity and type-2 diabetes THEN they are at risk of developing liver disease UNLESS they are prescribed a given medication or lifestyle change". Such process of enumeration is a type of 'formal reasoning' and cannot be sufficiently expressed through frequentist approaches alone. Since the earliest days of AI there has been a succession of increasingly powerful methods to enable computers to "reason". As early as 1960 programming languages were developed that processed symbolic as opposed to numerical data, with innovations such as the "resolution principle" developed by John Alan Robinson in the 1970s demonstrating that algorithms could be used to reason and formulate plans.

As systems scaled in their ambition, a range of techniques were devised, one of which was Expert Systems whose popularity grew rapidly 1975-1985. If provided with sufficient domain knowledge, expert systems could match the performance of humans in specific and narrow tasks, such as the 1971 expert system 'DENDRAL' which was developed for identifying organic molecules from mass spectrometry. In 1975 Marvin Minsky presented his development of 'Frames', a form of semantic network where objects were represented with relations (edges) to other objects, though Charles Peirce (1909) is credited with developing the first semantic network and devised existential graphs where the graphical notation of nodes and edges were established. In 1975, Wood's [73] paper titled 'What's in a link' also captured the attention of the AI community and the need for precise semantics to formalise knowledge

representation - the foundations of graphical modelling were therefore laid. However, a fundamental problem with such systems were that they were hard to control, a dilemma confronted by Judea Pearl in 1982 with the development of message-passing techniques that allowed tractable inferences [74]. Such seminal work laid the foundations for the main developments within expert systems, specifically:

1. The incorporation of uncertainty, for example by modelling information probabilistically (this was referred to as “reasoning under uncertainty”); and
2. The development of a method to control which inferences are made. This was a feature of the graphical representation of information developed by Judea Pearl and which was called Bayesian networks (BNs).
3. The further development of BNs to model causality. This is based on the introduction of the “DO-operator”, which provides a formal mathematical basis for reasoning about randomised clinical trials. Specifically, if Y is a treatment and X is a subject, then the randomised controlled trial compares two probabilities: $p(X|do(Y))$ vs $p(X|do(not\ Y))$.

Though there exist other flavours of graphical modelling, such as systems mapping as used in the FORESIGHT report to build a graphical model of obesity system map [12], they (unlike Bayesian-networks) do not allow for counterfactual estimation or the probability of an event, given evidence. The first expert system incorporating BNs dates back to 1983 with the system ‘CONVINCE’, with early applications in medicine including the ‘MUNIN’ [75] and ‘PATHFINDER’ [76] networks that applied BNs to model neuromuscular disorders and tumour pathology, respectively. In the following Chapter, I describe the semantics behind Bayesian-networks,

the methodology for probabilistic inference and provide an overview of the current applications of Bayesian-networks in medicine. I then close by highlighting the gap in the literature that the work contained in this Thesis addresses.

3

BAYESIAN-NETWORKS AND PROBABILISTIC INFERENCE

As detailed in the previous Chapter, I conjectured that Bayesian-networks offer a promising tool for modelling the complexity of biological systems and 'speaking' the language of clinicians. This ability stems from the three main components of a Bayesian-network: 1) the network structure that explicitly encodes the direction of conditional dependence; 2) the parameterisation of the structure that permits probabilistic inference; and 3) the localisation and control of information flow, as denoted by the network structure. This Chapter summarises the methodology of Bayesian-networks, their utility for studying metabolic disease and current applications of Bayesian-networks in the field of medicine/healthcare. The structure of a Bayesian-network and the information it models is a natural place to commence this Chapter.

3.1. Structure of Bayesian-networks (BN)

BN are a class of graphical models that encode probabilistic relationships in the form of a directed acyclic graph (DAG). Formally, a DAG is expressed as $G = (X, E)$, where $X = (x_1, \dots x_n)$ denotes the variables within the network, and where E is a set of directed edges relating pairs of variables in X. The directionality of an edge captures the flow of *information* between two variables, where in the case of an edge from A to B, the value of B is conditionally

dependent on the value of A. Concretely, the structure of a DAG is comprised of three types of connections: *chains, forks, and colliders*. Together these allow the user to conveniently (and visually) detect interdependencies within the data. The structure of a DAG is non-trivial, for it solves the "mutual admiration" problem that arises in the case of reinforcement cycles, whereby confidence in A gives an increase confidence in B, which leads to increased confidence in A, and so forth. Such problem is resolved in case of Bayesian-networks for their structure are acyclic.

Chain connections

A chain connection is comprised of an unbroken direct path between any number of variables, for example between three variables (*A*, *B* and *C*) as illustrated in Fig. 3.1a.

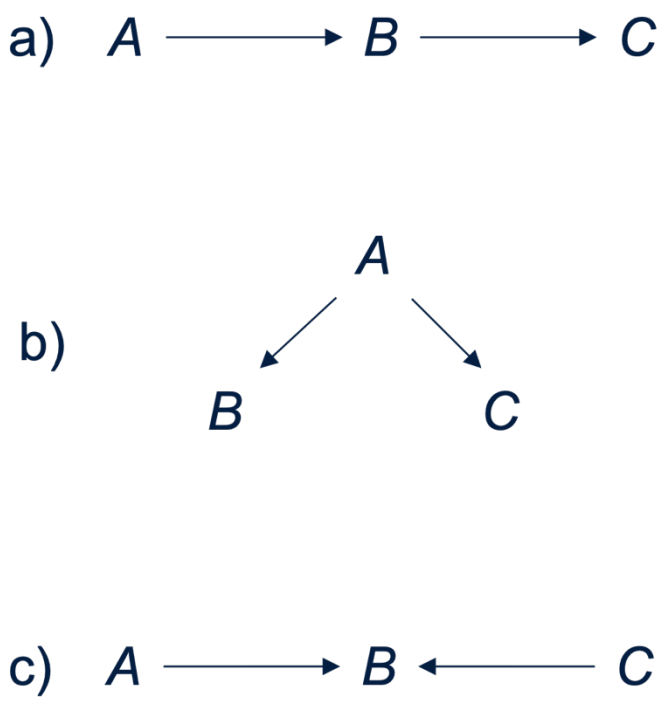


Figure 3.1. a) example chain connection between three variables: *A*, *B* and *C*; b) example fork connection; c) example collider connection.

Firstly, the presence of an edge between two variables renders them dependent, with the directionality of the edge illustrating the direction of conditional dependence. That is, the conditionally dependent variable (in the present example, *B* or C) depends on the preceding variable (in the present example, *A* or *B*) for its value, and that changing the value of *X* will in some manner change the value of *B.* In the example of Fig. 3.1a, we can also ascertain that the value of *C* is dependent on the value of *A*, for the value of *C* is dependent on the value of *B,* and the value of *B* is dependent on the value of *A.* However, it is also useful to examine the three types of connections described here through the study of conditional *independence*, for if we ascertain that A and C are conditionally *independent* given B, then knowing information about A tells us nothing about C, and knowing information about C tells us nothing about A. Of course, this is the similar to the study of dependence (which, exploring the opposite tells us that two variables are dependent), but it is sometimes useful to explore the structure of a Bayesian-network in this way, such as in the case of Markov Blankets. The following three rules from Judea Pearl explore conditional independence in this way. Conditional independence in chain connections is formally defined by Pearl's Rule 1 [77]:

Rule 1 (Conditional Independence in Chains) *Two variables X and Z are conditionally independent given Y if there is only one unidirectional path between X and Z, and that Y is any set of variables that intercepts that path.*

Fork connections

A fork connection is comprised of two variables (*B* and *C*) that are conditionally dependent on a single, common variable (*A),* but where no edge exists between *B* and *C,* see Fig. 3.1b. A

common example of a fork connection is illustrated in Fig 3.2; whereby *X* denotes the city temperature (in Fahrenheit) on a particular day, and *Y* and *Z* the number of violent crimes and local ice cream sales that day, respectively. While the number of ice cream sales and violent crimes in a particular city may increase or decrease dependent on the temperature of that given day, they themselves are not dependent upon one another (in this example, we assume the absence of organised crime units who target local ice cream parlours given the temperature of that day). Another example of a spurious correlation arising from a fork connection is the chocolate consumption and Nobel prize winners' correlation introduced in the previous Chapter.

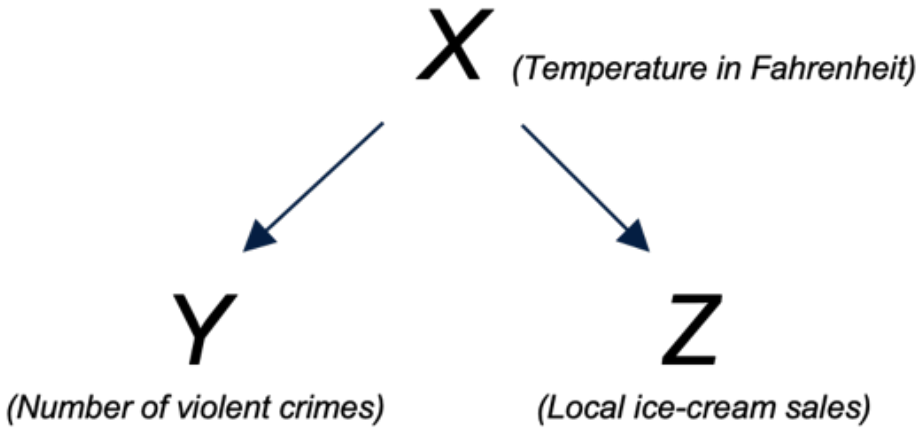


Figure 3.2. Example fork connection, whereby the value of X is a common cause of both Z and Y.

As discussed in Chapter 3, when two conditionally independent variables (in the present example, number of serious crimes and ice-cream sales) appear correlated, despite no plausible causal connection, there is what is referred to as a *'spurious correlation'*. As illustrated by Pearl's Rule 2, conditional independence in fork connections only holds if:

Rule 2 (Conditional Independence in Forks) *If a variable X is a common cause of variables Y and Z, and there is only one path between Y and Z, then Y and Z are independent, conditional on X.*

Collider connections

A collider connection occurs when a single variable (*B*) receives edges from two (or more) variables (*A* and *C*), rendering *B* as a common effect of its two causes, *A* and *C* (see Fig 3.1c). Such a connection raises unique challenges, most notably that when conditioning on a collider variable (*B*), we in turn produce a dependence between the parents of *B* that would otherwise not have existed.

To expand on this further, we take the Pearl et al. [77] description of collider dependence form '*Causal Inference and Statistics'*. If we take our three variables, *A*, *B* and *C,* where *B* is the collider variable and *A* and *C* are independent variables, the following logic applies: where if we learn that the value of *A* = 1, we learn nothing about the value of *C*, for *A* and *C* are independent. However, if we start by learning that the value of *B* = 2, then knowing that the value of *A* = 1 immediately reveals that the value of *C* must also be 1 (here we are under the assumption that both contributions of A and C to B are additive). Thus, when conditioning on *B* by learning that *B* = 2, *A* and *C* are dependent, given *B*. Conditional independence in collider connections is formalised in Pearl's Rule 3:

Rule 3 (Conditional Independence in Colliders) *If Y is a collision between X and Z, and there is only one path between Y and Z, then X and Z are unconditionally independent, but dependent conditionally on Y and any descendants of Y.*

Due to their acyclic nature, under certain assumptions BNs can provide a convenient visual representation of *causal* relationships. When applying BNs to represent biological systems, network structure discovery becomes a vital task, for such systems are often inherently *causal* in nature. The task of learning and parametrising the structure of a BN can be illustrated as maximising the following, as described by [78] :

$$\underbrace{P\left(B\mid D\right)=P\left(\mathrm{G},\Theta\mid D\right)}_{\textit{learning}}\quad=\quad\underbrace{P\left(G\mid D\right)}_{\textit{structure learning}}\quad\cdot\quad\underbrace{P(\Theta\mid G,D)}_{\textit{parameter learning}}$$

where D is a sample taken from **X**; B = (G, Θ) is our BN with a DAG structure G and parameter set Θ. Structure learning is defined as finding the DAG structure G that best encodes the dependence structure inherent in the data, while parameter learning relates to the estimation of the parameters Θ given G.

3.2. Deriving the structure of a Bayesian-network

Constructing a BN can be completed through either applying domain knowledge to manually insert and direct edges between pairs of variables, or via automated algorithms, the latter

being favourable when domain knowledge on the underlying structure is limited or unknown. Constructing a BN through the manual approach is explored in further detail in Chapter 4, where I run through the initial development of a BN in collaboration with medical experts. What follows is an overview of the automated approaches to BN structure learning.

The task of discovering our network structure G using automated discovery algorithms can be expressed as follows, where given a select data set $D = \{x_1, \dots x_n\}$, we return G^* that maximises the following:

$$G^* = \arg \max_{G \in G^n} f\,(G : D)$$

where $f(G{:}\,D)$ is a scoring criterion that ranks candidate G's respective to their 'goodness of fit' to D. Such a task has proven to be NP-hard [79], meaning that its complexity is at least as hard as the hardest problems in nondeterministic polynomial time, with [80] showing how the number of candidate Gs increases exponentially with the number of variables, as demonstrated through the following equation:

$$f(n) = \sum_{i=1}^{n} (-1)^{i+1} \frac{n!}{(n-1)!\,i\,!}\, f(n-i)$$

where n denotes the number of network variables and $f(n)$ the set of candidate G composed from n variables. Furthermore, the difficulty of returning G^* is exacerbated in the case of *equivalence classes* (Θ), where $\Theta = \{G_1, G_2, \dots, G_N\}$ denotes a set of statistically indistinguishable G that encode identical independence relationships despite different edge

orientations [81]. According to Verma and Pearl's 1991 theorem [82], $\Theta = \{G_1, G_2\}$ *iff* G_1 and G_2 contain the same skeletons (same set of edges regardless of orientation) and the same v-structures, where v-structures are the same structure as the 'fork' connections described previously.

To identify the network structure for a given data set (G from D), structure discovery algorithms commonly employ either score and search (SS), constraint-based (CI), or hybrid (SS and CI) techniques.

Score-and-search algorithms

To select an *optimal* model (G*) from a set of possible candidates, SS algorithms employ scoring criteria that rank a model based on its 'goodness of fit' to the given data set, that is, they seek to select a model that maximises a given likelihood function value. Common scoring criteria include the Bayesian Information Criterion (BIC) [83] and Akaike's Information Criterion (AIC) [84], and are used to guide the search space among a set of finite models. For example, the BIC and AIC are formally defined as follows:

$$BIC = k\ln(n) - 2\ln(\hat{L})$$

$$AIC = 2k - 2\ln(\hat{L})$$

where $\hat{L}$ denotes the maximised value of the likelihood function for the model M, where M: $\hat{L} = p(x \mid \hat{\theta}, M)$; $\hat{\theta}$ are the parameter values that maximises the likelihood function and x denotes the observed data; n denotes the sample size of x; and k denotes the number of

parameters estimated by the model. As such, SS techniques use scoring criterion such as the BIC to maximise $\hat{L}$ and move within in the search space to return $G*$. For example, the Greedy search Hill-Climbing algorithm [85] selects the *local edge function* (edge insertion, removal, or reversal) that elicits the greatest improvement in network score as determined by criterion such as the BIC and AIC. The search terminates and the DAG is returned once no local edge function can be found that improves the network score.

To avoid overfitting, when adding model parameters to increase the maximum likelihood function, both the BIC and AIC include a penalised log-likelihood function (PLLF), where the log-likelihood is the log probability of D given G. PLLF's apply a penalty term relative to the complexity of the network, with the AIC employing a penalty term – $2k$ – less than that of the BIC – $\ln(n)\, k$ – and therefore favouring more complex networks.

Constraint-based algorithms

Constraint-based (CI) algorithms typically share a common three-phase process originating from Verma and Pearl's Inductive Causation algorithm [86]. The first (optional) phase learns the Markov blanket of each network variable (x_i), defined as the minimal set of variables conditioned on x_i that render all other network variables independent of x_i. Next, the DAG skeleton (edges, without orientation) is returned through tests of conditional independence that return the parents and children of a target variable [87]. CI metrics such as the X^2 and Fisher's Z test seek to accept or reject the hypothesis that two nodes are conditionally independent, expressed as:

$$P\ (x_1, x_2 |\ x_z) = P\ (x_1 |\ x_z)\ P\ (x_2 \mid x_z)$$

where x_1 and x_2 denote network variables and x_z a set of variables. Lastly, network edges are orientated as detailed in [88], however some edges can remain undirected as both directions encode equivalent decompositions of the global distribution. Such networks are called *completely partially directed acyclic graphs* (CPDAGs) and can belong to a set of equivalence classes (Θ).

Hybrid-search algorithms

Due to the NP-hard nature of discovering G^*, hybrid structure learning algorithms that combine the benefits of SS and CI techniques have gained popularity. Hybrid techniques such as the Max-Min Hill Climbing (MMHC) algorithm [89] constrain the search space of candidate DAGs by firstly running tests of conditional independence to return a network skeleton (edges, without orientation) as per the CI techniques described above. Next, local SS techniques such as greedy hill climbing orientate network edges by selecting the function (edge insertion, reversal, or removal) that elicits the greatest increase in network score (such as the BIC or AIC). Notably, The MMHC algorithm is constrained to only consider inserting an edge if identified during the initial skeleton learning step, improving computational time, efficiency, and accuracy [89]. A brief overview of some additional automated structure discovery algorithms are detailed below in Table 3.1. Consequently, each scoring criterion and automated search algorithm may return a different *optimal* network, and it's the responsibility of the user to select a methodology suitable for the task at hand. For example, if the user is seeking to return a more parsimonious network from a large dataset, applying the MMHC algorithm with the BIC scoring criterion should be selected, for the MMHC is a constrained algorithm and the BIC applies a higher penalty term to each parameter in the network.

Table 3.1. Overview of example automated network construction algorithms.

Score and search algorithms
Greedy search Hill-Climbing (HC): A greedy search hill-climb selects the local edge function (edge insertion, removal, or reversal) that elicits the greatest improvement in network score as determined by criterion such as the BIC and AIC. The search terminates and the DAG is returned once no local edge function can be found that improves the network score [85].
Tabu search (TS): Similar to HC, Tabu search selects the local edge function that elicits the greatest improvement in network score. However, TS employs a guided local search procedure that rejects any move within the search space that has already visited. This is achieved through maintaining a 'tabu list' of visited spaces to avoid a locally optimal, yet globally sub-optimal, network score [90].
Constraint-based algorithms
PC (stable) (PC): The PC stable algorithm is a modification of the PC algorithm detailed by [91]. This modification removes all or part of the order dependence, a complication of the original PC algorithm where the network output has shown to be dependent on the order in which the variables are introduced. The PC stable modification has demonstrated superior performance when applied to higher dimensionality datasets [92].
Fast Incremental Association (Fast.IAMB): The Fast Incremental Association Markov Blanket algorithm is a modification of the original Incremental Association Markov Blanket algorithm detailed by [87]. Fast.IAMB employs a heuristic to recover the Markov-blanket of a target variable faster and more reliably, without adversely affecting the discovered Markov blanket [93].
Hybrid algorithms
Max-Min Hill-Climbing (MMHC): The Max-Min Hill-Climbing algorithm first returns the network skeleton (edges, without orientation) through the local discovery Max-Min Parents and Children (MMPC) algorithm. Next, network edges are orientated using a greedy Bayesian-scoring hill-climbing search, where the search is constrained to only consider adding an edge if returned by the MMPC algorithm in the first instance [89].
Hybrid Parents and Children (H2PC): Like the MMHC algorithm, the Hybrid Parents and Children algorithm firstly learns the network skeleton through tests of conditional independence and performs a Bayesian-scoring greedy hill-climbing search to orientate network edges. The difference between the H2PC and MMHC algorithms is in the CI approach for returning the network skeleton, where H2PC instead applies an ensemble method combining the *Data-Efficient*

Parents and Children Superset, Data-Efficient Spouses Superset and Interleaved Incremental Association Parents and Children to return the parents and children of a target variable [94].

3.3 Probability distributions & parameterization of Bayesian-networks

Each variable within the BN is fitted (parameterised) with data, that is, each variable has a corresponding conditional probability table in the case of discretised variables (for example, obesity status), or conditional probability distribution for continuous variables (for example, body mass index). Bayesian-networks can contain variables that are discretised only, continuous only, or both continuous and discretised (labelled 'hybrid' networks) sources of data.

Concretely, a variable – x_1- is conditionally independent of its non-descendants *given* its parents within the network, denoted $P(x_1 | Pa\ (x_1))$, with the value of x_1 conditionally dependent on that of its parent (Pa) node(s). Such conditional dependencies are factorised to form a joint probability distribution across the entire network via the chain rule:

$$P\ (x_1, x_2, \dots, x_n) = P\ (x_1)\ P\ (x_2|x_1) \dots P\ (x_n|x_1 x_2 \dots\ x_n{-}_1)$$

Such factorisation also allows a compact representation of the joint probability distribution of a specific variable, for x_1 and x_2 will not appear in each other's conditioning set if deemed conditionally independent given G, denoted as:

$$x_1 \perp\!\!\!\perp_{\boldsymbol{G}} x_2 \mid x_z \Leftrightarrow x_1 \perp\!\!\!\perp_{\boldsymbol{P}} x_2 \mid x_z,$$

where P denotes the joint probability distribution and x_z a set of variables. Under such scenarios, it is said that G is 'a perfect map' of P and any conditional dependency can be recovered as follows:

$$P(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} P(x_i | Pa(x_i))$$

Conditional probability tables (CPT), learned from the data we fit to our network, illustrate the conditional probability of each value state of that variable. For example, in Fig. 3.3 the value states of 'Obesity status' are Non-obese (0.4 = 40%) and Obese (0.6 = 60%), where the sum of each probability contained in the CPT must equal 1, denoting absolute probability (100%). In the case of continuous networks, each variable has a local distribution which is represented as a linear regression model, where each value state contains a regression coefficient associated with the 'parents' of each variable.

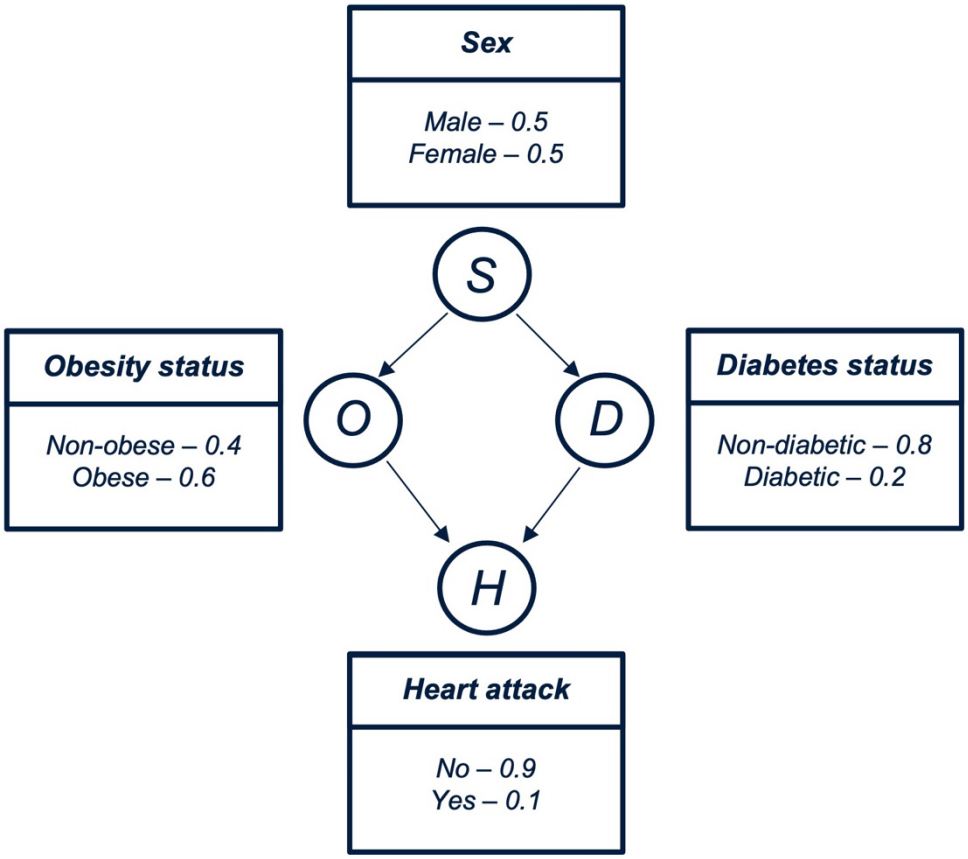


Figure 3.3 Example simple discrete Bayesian-network with corresponding conditional probability tables.

For transparency, much of the following information on network parameterisation is drawn from the work of [95]. BNs are parameterised using both the network structure, i.e.., the conditional independencies made explicit by the DAG, and the dataset used in the construction of the network. The techniques used for fitting the data are dependent on whether the network contains discrete, continuous or hybrid data types. In the case of discrete networks, the variables can be parameterised using a maximum likelihood estimate (MLE) or Bayesian posterior estimator.

The MLE approach applies a likelihood function L to select the parameters $\hat{\theta}$ in Θ that 'best' fits our data, where given a data set D, we choose parameters $\widehat{\boldsymbol{\theta}}$ that satisfy:

$$L\left(\widehat{\boldsymbol{\theta}} : D\right) = \max_{\boldsymbol{\theta}\in\Theta} L\left(\boldsymbol{\theta} : D\right).$$

Using such approach, the MLE takes a *point estimate* from the distribution of values derived from the data, and while a common approach, can sometimes be limited and oversimplistic. For example, say we flip a coin 10 times and observe that it lands on heads 3 times, would it be reasonable to ascertain that the parameter θ is 0.3? There is nothing to the contrary to say this isn't a fair assumption. Now what about if we repeated our experiment 10,000 times, noting that we observe the coin landing on head 3,000 times? In such a case, though we may be more likely to accept such θ given the large data set and therefore a greater confidence in θ, our *prior knowledge* might inform us that such result is atypical and potentially the result of a trick coin (biased towards landing on tails). The MLE does not allow us to distinguish between such scenarios – whether a trial has been conducted 10 or 10,000 times – nor incorporate our prior domain knowledge into parameter estimation (in this example, our subjective belief about the coin).

Alternatively, the Bayesian-posterior estimator does not attempt to return a single point estimate and instead builds a posterior probability distribution by incorporating the prior probability and likelihood of θ through Bayes' rule, as in Fig. 3.4.

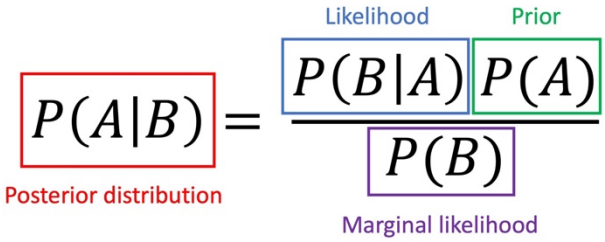


Figure 3.4. Representation of Bayes' Theorem.

Prior probabilities allow us to incorporate our subjective beliefs about θ's values and quantify the probability we assign to them. For example, if we suspect a trick coin or biased die is more likely to land on a specific value, we may assign that value a higher *prior probability*, even though we haven't observed the data yet. We then, using the Bayesian approach, incorporate these prior probabilities about the parameters θ with the true observations using probabilistic reasoning, as follows:

$$P(\boldsymbol{\theta} \mid D) = \frac{P\ (D \mid \boldsymbol{\theta})\ P\ (\boldsymbol{\theta})}{P\ (D)}$$

where $P(D \mid \boldsymbol{\theta})$ is the likelihood function (derived from the observations); $P(\boldsymbol{\theta})$ is the prior probability distribution over the possible values of Θ; and $P(D)$ is the marginal likelihood of the data as calculated by the integration of the likelihood over all possible parameter values:

$$P(D) = \int_{\Theta} P(D \mid \boldsymbol{\theta})\ P\ (\boldsymbol{\theta}) d\boldsymbol{\theta}$$

As such, the Bayesian approach returns the marginal posterior probability for each variable from the global distribution of the BN. However, as is the case with the work contained in this Thesis, we can also apply uniform priors for parameterising the network, that is, priors with a uniform distribution $P(\theta) = 1$ for all $\theta \in [0,1]$. But how is this then different from the MLE approach? The main philosophical difference relates to how we use the posterior distribution, where instead of selecting a single value for θ as in the MLE approach, we instead use the entirety of the posterior distribution for predicting the marginalised probability. Indeed, with a uniform prior distribution, the posterior distribution can simply be viewed as the normalised likelihood function, which has demonstrated superior predictive power and smoothness compared to MLE [95].

0.2. **Estimating treatment outcomes and sampling from the Bayesian-network**

Conditional dependence (or causality) encoded within Bayesian-networks is especially critical in the context of biological systems, for each variable is conditionally independent of its non-descendants given its parents. This allows the user to apply probabilistic inference and pose counterfactual 'what if' questions to aid clinical decision-making such as those presented in Chapter 1 – 'what would happen to the probability of value *i* in variable Z (Z^i), given that I change the probability of value *X*?'. Returning to the third main development in expert systems detailed in Chapter 2, such intervention relates to Pearl's "DO-operator" methodology, which provides a formal mathematical basis for reasoning about randomised clinical trials (RCTs). Specifically, if Y is a treatment and X is a subject, then the RCT compares two probabilities: P(X|do(Y)) vs P(X|do(not Y)). In the example of Fig. 3.3, such a question could take the form of: 'what is the difference in the probability that my patient will experience a heart attack, *given that* they are obese or non-obese?'. We pose and estimate

such events by *intervening* on the network, where the user fixes the value of a specific variable(s), denoted as the 'evidence', and estimates the probability of an event, *given* the evidence.

Through posing counterfactual questions, we are comparing between the potential outcomes of doing, or not doing, a specific action (in the previous example, moving a patient from obese to non-obese). This is mathematically modelled by comparing the outcome of the following:

P (Heart attack = 'Yes | Obesity status = 'non-obese') and P (Heart attack = 'Yes | Obesity status = 'Obese')

where 'P' denotes probability and '|' denotes 'given'.

The probability of the selected 'event' can be estimated using approximate inference algorithms that random instantiations – or samples – based on Monte Carlo sampling [96]. Approximate sampling algorithms such as Logic Sampling (LS) and Likelihood Weighting (LW) generate random samples using a forward propagation method as denoted by the graphical structure. Returning to the example network in Fig. 3.3, we have expanded this to include the full joint probability distribution as illustrated in Fig. 3.5.

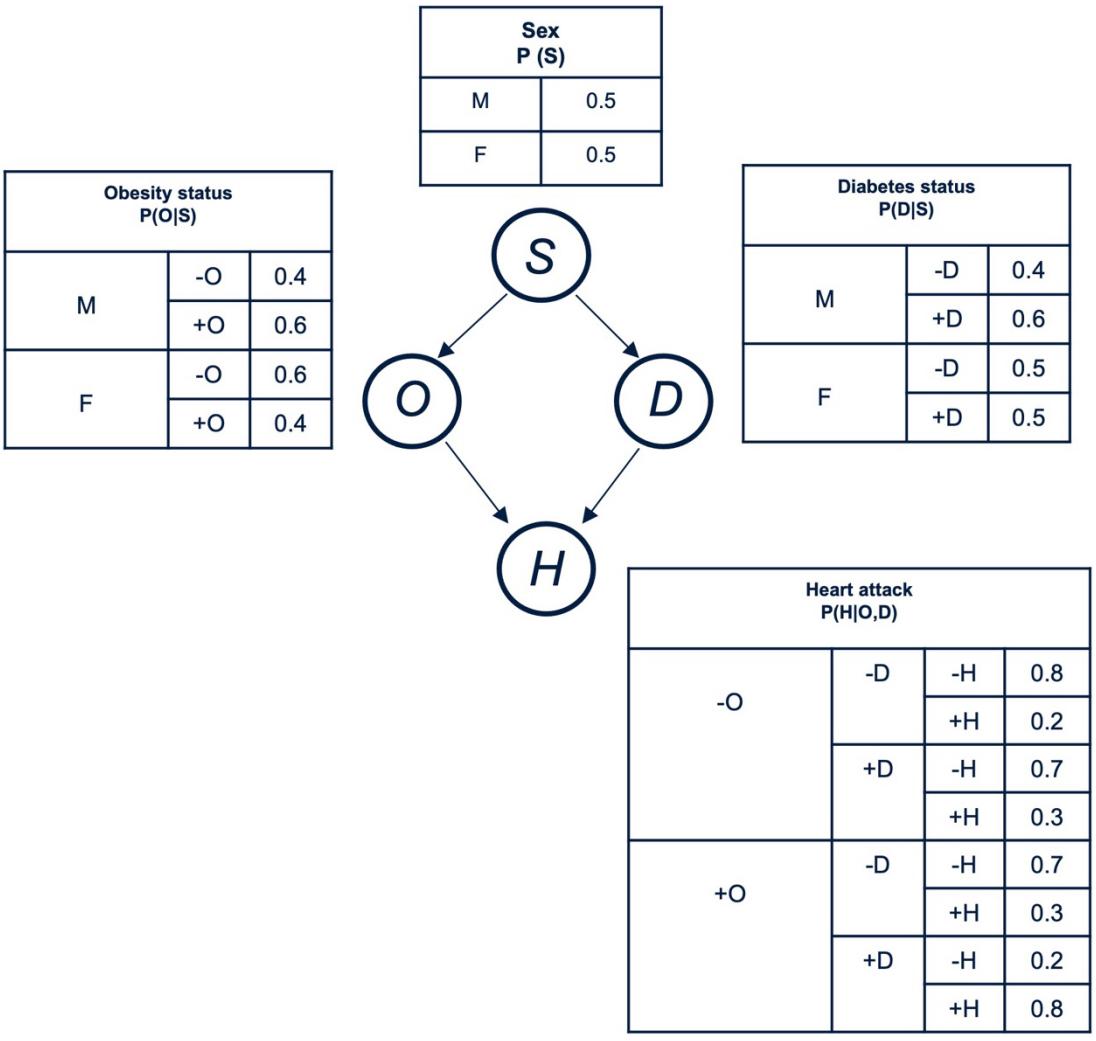


Figure 3.5. Full joint probability distribution of example Bayesian-network with conditional probability tables. M = male, F = female; -0 = non-obese, +O = obese; -D = non-diabetic, +D = diabetic; -H = no heart attack, +H confirmed heart attack and '|' denotes 'given'.

In the case of prior sampling, or any sampling algorithm for that matter, we compute the probability query by counting the actual samples generated by the sampling procedure. This is because, given we have N total samples, where $N_{PS}(x_1, \dots x_n)$ denotes the number of times our query $x_1, \dots x_n$ occurs in our sample N_{PS}, we expect this number as a fraction of N to converge to its expected value as per sampling probability:

$$\lim_{n\to\infty} \frac{N_{PS}(x_1, \dots x_n)}{N} = S_{PS}(x_1, \dots x_n) = P(x_1, \dots x_n),$$

where S_{PS} is our sampling procedure (prior sampling). The probability of an event can therefore be expressed as a normalised fraction of $N_{PS}(x_1, \dots x_n)$ bounded between 0-1. For example, using the joint probability distribution illustrated in Fig. 3.5., we may be interested in estimating $P(Heart\ attack = +H)$. Table 3.2 illustrates an example distribution after running 10 random samples.

Table 3.2. Example distribution from 10 random sample instantiations.

	Variable			
Sample	**Sex**	**Obesity status**	**Diabetes status**	**Heart attack**
1	M	+O	+D	+H
2	M	+O	+D	-H
3	F	-O	-D	-H
4	M	-O	-D	-H
5	M	-O	+D	+H
6	M	+O	-D	+H
7	F	+O	-D	-H
8	F	+O	+D	-H
9	F	-O	-D	-H
10	M	+O	+D	-H

To answer this query, we then retrospectively analyse the distribution to select which samples match our evidence ($+H$). Such process returns a sample with a probability as follows:

$$S_{PS}\,(x_1 \dots x_n) = \prod_{i=1}^{n} P\big(x_i \big| Parents(X_i)\big) = P(x_1 \dots x_n),$$

whereby S_{PS} is the sampling procedure (prior sampling), x_i is the network variable and $(x_1 \dots x_n)$ is the generated sample. In our example, samples 1, 5 and 6 are consistent with our query ($+H$), therefore meaning $\hat{P}(Heart\ attack = +H) = 0.3$. Going one step further, we may also be interested in estimating conditional probabilities, that is, the probability that someone will experience a heart attack ($+H$), given they are obese ($+O$) and diabetic ($+D$), expressed as $P(+H \mid +O, +D)$. Returning to the distribution in Table 3.2., samples 1, 2, 8 and 10 are all consistent with our query ($+O$ and $+D$), and of these four samples, +H occurs once (sample 1), meaning our probability of $\hat{P}(+H|+O, +D) = 0.25$.

However, this approach can be computationally expensive and inefficient, for it requires generating an entire distribution and then picking out the samples that match our evidence. Instead, *rejection sampling* follows the same sampling process as prior sampling, however, will terminate if the sample does not match the evidence. For example, in the distribution shown in Table 3.2., the sample would terminate at the point of -H in sample 2 and -O in sample 3. The sample count from the generated distribution is then returned as a marginalised probability value bounded 0-1 as per prior sampling. However, rejection sampling can also be inefficient if the prior probability of the specified evidence is low, for the

entire generated sample will be discarded if different from the evidence, regardless of how far through the network the sample has reached. This can also lead to inaccuracies when trying to estimate outcomes with a low prior probability, as is often the case in medical research, where low-probability evidence is usually the rule rather than the exception.

To overcome this limitation of rejecting sampling, *the Likelihood Weighting algorithm* – a type of importance sampling – does not reject samples, but instead fixes the evidence variables to the query value (in the previous example: H+, D+ and O+) and assigns each generated sample a weight. After assigning an initial weight of 1.0 to the sample, the evidence variables are then accorded an individual weight that reflects the probability of that sample based on the likelihood. Such weights are derived from the original joint probability distribution, as illustrated below in Fig. 3.6., while non-evidence variables are sampled according to the conditional distribution.

As an example, returning to the BN in Fig. 3.5., we will simplify this query to calculate the probability of Heart attack = +H, given Obesity status = +O. Given we generate the following three random samples with +O and +H as our fixed evidence variables, we would derive the following weighted samples as illustrated in Table 3.3:

Table 3.3. Example distribution with corresponding weights from three random samples.

	Variable				
Sample	**Sex**	**Obesity status**	**Diabetes status**	**Heart attack**	ω
1	M	+O	-D	+H	0.18
2	F	+O	+D	+H	0.32
3	F	+O	-D	+H	0.12

The weighted sample is calculated from the product of the individual variable weights derived from the original joint probability distribution shown in Fig 3.5. As an example, 'sample 1' is weighted as follows and is also illustrated in Fig. 3.6:

$$\omega = 1.0\ (initial\ weighted\ sample)\ \text{x}\ 0.6\ (evidence)\ \text{x}\ 0.3\ (evidence) = 0.18$$

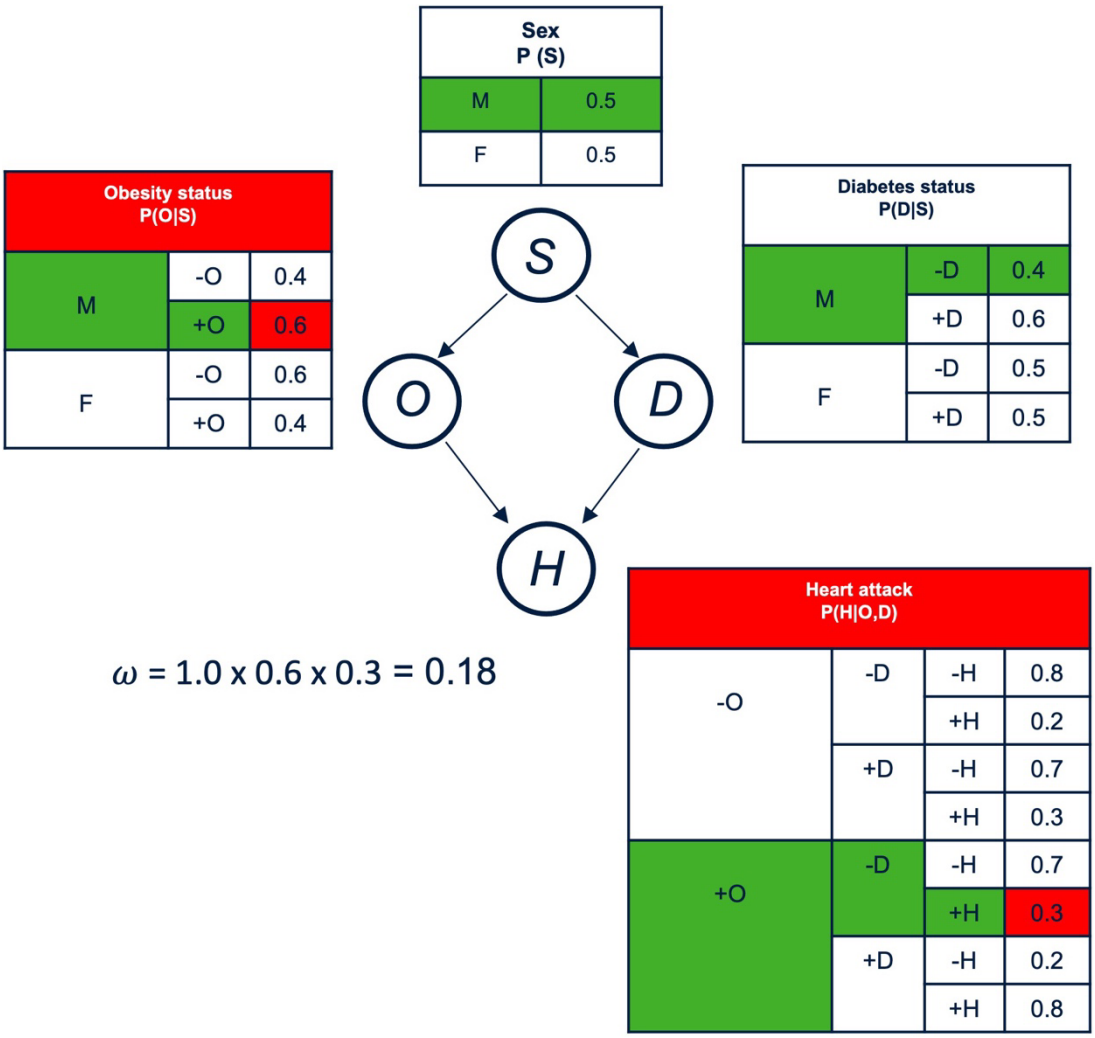


Figure 3.6. Full conditional probability distribution with highlighted (green) sample and highlighted (red) evidence variables and corresponding weights from 'sample 1'. M = male, F = female; -0 = non-obese, +O = obese; -D = non-diabetic, +D = diabetic; -H = no heart attack, +H confirmed heart attack and '|' denotes 'given'.

The weighted sample distribution is then used to correct the original sample distribution, whereby we firstly calculate the product of the (unweighted) probabilities of the sampled variables, given their parents, as follows:

$$S_{WS}(\boldsymbol{z}, \boldsymbol{e}) = \prod_{i=1}^{l} P\big(z_i \big| Parents(z_i)\big),$$

Whereby $\boldsymbol{z}$ and $\boldsymbol{e}$ denote our sampled and evidence variables, respectively. We then calculate the product of the weighted samples:

$$\omega(z, e) = \prod_{i=1}^{m} P(e_i | Parents(E_i))$$

The original sample distribution is then corrected for by multiplying by the weighted evidence variable distribution, returning a weighted normalised probability value:

$$S_{WS}(z, e) \cdot \omega(z, e) = \prod_{i=1}^{l} P\big(z_i \big| Parents(z_i)\big) \prod_{i=1}^{m} P\big(e_i \big| Parents(e_i)\big)$$

$$= P(z, e)$$

The LW algorithm process is formally illustrated as follows, as taken from [97]:

function LIKELIHOOD-WEIGHTING (X, e, bn, N) **returns** an estimate of $P(X|\boldsymbol{e})$
inputs: X, the query variable
$\boldsymbol{e}$, observed values for variables $\boldsymbol{E}$
bn, a Bayesian-network specifying joint distribution $\boldsymbol{P}(X_1, \ldots X_n)$
N, the total number of samples to be generated
local variables: $\boldsymbol{W}$, a vector of weighted counts for each value of X, initially zero

for *j* = 1 to N **do**
$\boldsymbol{x}$, $\omega \leftarrow$ WEIGHTED-SAMPLE$(bn, \boldsymbol{e})$
$\boldsymbol{W}[x] \leftarrow \boldsymbol{W}[x] + w$ where x is the value of X in $\boldsymbol{x}$
return NORMALISE$(\boldsymbol{W})$

function WEIGHTED-SAMPLE$(bn, \boldsymbol{e})$ returns an event and a weight
$\omega \leftarrow 1$; $\boldsymbol{x} \leftarrow$ an event with n elements initialised from $\boldsymbol{e}$
foreach variable X_i in $X_1, \ldots, X_n$ **do**
if X_i is an evidence variable with value x_i in $\boldsymbol{e}$
then $\omega \leftarrow \omega\, P(X_i = x_i \mid parents(X_i))$
else $\boldsymbol{x}[i] \leftarrow$ a random sample from $P(X_i \mid parents(X_i))$
return $\boldsymbol{x}$, ω

The likelihood-weighting algorithm for inference in Bayesian networks. In WEIGHTED-SAMPLE, each nonevidence variable is sampled according to the conditional distribution given the values already sampled for the variable's parents, while a weight is accumulated based on the likelihood for each evidence variable.

3.6. Existing applications of Bayesian-networks for studying biological systems

The aforementioned features render Bayesian-networks especially pertinent to biological systems, which are not only causal (not simply associational) in nature, but naturally stochastic and difficult to model while capturing *uncertainty*. The ability to update CPTs to reflect new evidence or to explore a specific phenotype provides a flexibility to Bayesian-networks that allows them to be customised to the individual patient, thus promoting the delivery of personalised care. Furthermore, Bayesian-networks are particularly apt for

incorporating different sources of data into a single model, for example index measures such as the body mass index (kg/m^2) MRI-derived quantitative biomarkers such liver proton density fat fraction (%), or binary variables such as sex (male/female). The ability to include such a plethora of data, together with the ability to explore individual patient phenotypes, has led to the application of Bayesian-networks across a spectrum of medical fields. From using Bayesian-networks to explore the effect of biochemical abnormalities on acute kidney injury in patients with gastrointestinal cancer, to predicting cancer treatment outcomes, I now highlight some example applications.

As noted previously, Bayesian-networks can estimate potential treatment outcomes in a way that mimics the objective of randomised controlled trials. To illustrate this, I highlight the work of Vundavilli et al. [98] who applied Bayesian-networks to gene expression data to explore the effect of different drug therapies on breast cancer tumour cell death. Applying clinical domain knowledge, the authors firstly identified which genes conferred the greatest effect on the breast cancer pathway and the corresponding drug therapies known to target such pathways. This highlights a key strength of Bayesian-networks – the ability to apply clinical domain knowledge in constructing the network, resulting in what is referred to as a 'semantic' network. A semantic network is a graph that encodes our knowledge about a system and how concepts (or variables) in such system are related to each other through labelled nodes and edges. Semantic networks often encode our expert prior knowledge about a system, as in the present example. The authors restricted the maximum number of drugs per combinational therapy to two, given the harmful side-effects of such therapies, and set out to find which combination of drugs maximised cell death. Using Bayesian inference to

'switch' on and off each drug therapy and then measure the corresponding apoptosis ratio (a marker of cell death), the authors discovered that Cryptotanshinone (known to target the STAT3 and mTOR genes) taken in isolation or in combination with other therapies resulted in the greatest enhanced apoptosis ratio. This finding was experimentally validated by treating breast cancer cell lines with Cryptotanshinone, where the results not only agreed closely with the computational predictions but were vastly more effective compared to the existing standard therapy. BNs revealed a superior therapeutic target for breast cancer that was not the current standard of care.

I now illustrate how Bayesian-networks have been applied to predict patient outcomes following invasive procedures, firstly highlighting the work of Bradley et al. [99] who applied Bayesian-networks to predict patient outcomes following pancreatic resection surgery. Pooling data from 77 studies (n = 31,214), the authors identified 25 variables (for example, age, tumour size and tumour location) for construction of the Bayesian-network that was approved by a panel of pancreatic surgeons – and hence, is another example of a 'semantic' network. Individual patient data (n = 251) was then inserted into the network and post-surgical predictions of survival were compared against the patients' survival data. Despite containing missing data points, the Bayesian-network model achieved an AUC of 0.7 for predicting 12 months or less survival time post pancreatic resection. Although the authors did not perform any comparative analysis against other ML prediction models, similar work using a feedforward neural network reported an AUC of 0.65 [100]. Another example, which involves a comparison against other prediction models, is the work of Cao et al. [101] who applied Bayesian-networks to predict specific health-related quality of life (HRQoL) scores and comorbidities following bariatric surgery. In this paper, the authors compared the

performance of a Bayesian-network against convolution neural network (CNN) and multivariate logistic regression (MLR) models. 35 data variables taken from 6542 patients were used for building the models. Overall, the Bayesian-network demonstrated superior predictive performance than the CNN and MLR models, with a lower mean squares error for predicting HRQoL scores and higher AUCs for prediction of 5-year comorbidities such as hypertension and diabetes. Additional applications of Bayesian-networks in medicine including predicting T2D onset risk from electronic health records [102], identifying T2D patients prone to high medical costs [103] and for estimating primary coronary heart disease risk from measures such as obesity, age, and exercise status [104].

In addition to predicting treatment effects, Bayesian-networks can be leveraged to identify specific risk factors and mechanistic drivers of disease. For example, Li et al. [105] applied Bayesian-networks to model biochemical abnormalities and acute kidney injury (AKI) in patients with gastrointestinal cancer. Using probabilistic inference to 'switch' on and off biochemical abnormalities, the authors reported a 51.9% probability of AKI in patients with anaemia and hyponatremia, which was reduced to 13.9% when removing such abnormalities. Next, I highlight the work of Tian et al. [106] who applied Bayesian-networks to model obesity, metabolic health status (for example: cholesterol levels, blood glucose and blood pressure) and cardiovascular disease (CVD) risk in a Chinese population cohort (n=6276). Exploring individual obesity phenotypes, the authors revealed a 7% and 10% probability of high CVD risk in metabolically *healthy* normal weight and obese individuals, respectively. However, such probability was increased to 21% and 34% in their metabolically *unhealthy* normal weight and obese counterparts, respectively. Lastly, Xu et al. [107] applied Bayesian-networks to model the interrelationships between health behaviours such as BMI, physical activity

status, circulating biomarkers and quality of life scores in breast cancer survivors. Firstly, the authors identified how Bayesian-networks can be applied to identify pathways of disease, revealing that elevated BMI was a direct driver of arthritis and subsequent impaired quality of life. Secondly, through intervening on specific variables such as obesity status, the authors concluded that moving from the obese to overweight category resulted in an estimated 26% and 50% reduction in insulin and c-reactive protein, respectively. Furthermore, reducing both obesity status and the level of sleep impairment was predicted to cause a 26.7-point higher physical quality of life score on average.

The aforementioned usages of Bayesian-networks highlight their ability to go beyond simply classifying disease, but for identifying the drivers and risk factors behind such diseases and for estimating clinically meaningful counterfactual outcomes. A key theme underpinning the papers reviewed here is the use of domain clinical knowledge in the construction and application of Bayesian-networks. For example, [99] built their Bayesian-network in collaboration with a panel of surgeons, and [98] used domain knowledge to intervene on therapies known to affect specific genes. Additional examples of collaborating with medical experts for the construction of BNs, and the inferences made thereof, include the works of [108] and [109] who applied BNs to model glucose control in gestational diabetes and to manage serious pathology in lower back pain, respectively. The effectiveness of working with medical experts is highlighted by [110], who showed that a semantic BN built in collaboration with medical experts outperformed an automated, data-driven network for differentiating rheumatoid arthritis from similar inflammatory disorders. This capability not only confers a sense of ownership for the user, but also a transparency to Bayesian-networks that is particularly appealing to clinicians for it affords the ability to answer “why”. This is in contrast

to the 'black box' AI approaches that have faced resistance in medical adoption for several reasons, but specifically because "physicians are reluctant to trust and adopt something they do not fully understand" [111].

As the aforementioned studies demonstrate, Bayesian-networks have been applied to explore a wide variety of clinical applications. However, at the time I began my thesis research, Bayesian-networks had not been applied to address the limitations of the BMI through studying fat deposition and body composition. Nor had any work been done to explore imaging-derived measures of liver health to study the complex interplay between fat deposition, circulating biomarkers and liver disease. I conjectured that such work would allow clinicians to dive deeper into the phenotype of metabolically healthy and unhealthy obesity (as discussed in Chapter 1) and the distinct risk factors driving disease. Furthermore, the posing of counterfactual queries in such a context would allow the effect of therapeutic targets to be estimated, opening new lines of enquiry into the management of patients with liver disease and T2D. This Thesis is not concerned with the development of new Bayesian-network techniques such as inferential algorithms, rather, the novelty of this Thesis stems from demonstrating example use cases on how Bayesian-networks can be applied to inform clinical decision making, specifically through applying sampling techniques to clinical biomarkers in order to estimate counterfactual outcomes.

The following Chapters demonstrate the utility of Bayesian-networks to study MRI-derived quantitative measures of liver and body composition in combination with patient demographics, circulating biomarkers and anthropometrics. These projects were conducted in collaboration with world-leading experts into metabolic disease, namely Professor Dan

Cuthbertson (Liverpool University) and Professor Jean-Pierre Despres (Laval University, Quebec), to steer the clinical inferences made from the networks and the construction of such network structures. The contents of the following Chapters have been published in peer-reviewed journals and presented at international conferences.

4

MODELLING FAT DEPOSITION AND HYPERGLYCAEMIA: SEMANTIC NETWORKS AND INSIGHTS FROM BAYESIAN-NETWORKS

4.1. Introduction

As I noted in Chapter 1, in the United States and United Kingdom an estimated 30 and 4.3 million people, respectively, are diagnosed with type-2 diabetes (T2D), with disease prevalence projected to double by 2030 [14]. People with overweight or obesity are three and seven times more likely, respectively, to develop T2D compared to those considered as 'normal' body weight [16]. However, as I also highlighted in Chapter 1, the BMI is intrinsically limited for studying disease phenotypes for it does not consider body fat distribution or ectopic fat accumulation. Rather, the development of T2D includes a complex interplay of variables such as measures of fat distribution and circulating biomarkers that interact in myriad ways. Unlocking such interactions and their association with T2D would certainly help to open new lines of enquiry for exploring high-risk phenotypes and as a result, personalised medicine.

The previous Chapter emphasised how the embodiment of domain clinical knowledge within the Bayesian-network (BN) – deriving what is referred to as a 'semantic network' – is a key facet for deriving clinically meaningful inferences. While [102], [103], [105], [107], [112]

employed discovery algorithms in their work, only [112] report on correcting network structures in collaboration with medical experts. This trend is concerning, for while automated algorithms seek to return an *optimal* network, they don't possess knowledge of underlying biological processes. Such concern is exacerbated when applying BNs to represent biological systems, for the structure of the network explicitly encodes the direction of *causality* between variables.

The purpose of this Chapter therefore is to walk through the conception of a BN, including deriving the semantic structure, selecting a programming language in which to digitise such structure, and the application of probabilistic reasoning to pose and derive inferences. The main objective was to build a semantic BN in collaboration with medical experts using data from the *Obesity* publication. This would allow for 1) exploration of the utility of BNs for unpicking the complexity of fat deposition and diabetic-status hyperglycaemia; and 2) a comparison of BNs against the traditional statistical techniques used in the *Obesity* publication. The work contained in this Chapter was completed in collaboration with Professor Dan Cuthbertson from the University of Liverpool, and Professor Jean-Pierre Despres from the University of Quebec, both leading Endocrinologists and metabolic disease specialists. The work contained in this Chapter was accepted for oral presentation at the 2023 European Association for the Study of Diabetes.

4.2. Methodology

4.2.1. Data collection & preparation

The first task of this project was to derive a semantic structure from the clinical variables contained within the dataset. In this project, I applied the same BMI, age and sex matched UK Biobank dataset containing people with and without T2D, as per the *Obesity* publication. Data from 351 participants (171 with self-reported T2D, 180 with no reported diabetes) were collected retrospectively from the UK Biobank (UKBB) online resource (Application 9914).

MRI-derived measures of subcutaneous fat (SAT) (cm^2), VAT (cm^2), skeletal muscle index (SMI) (cm^2/m^2), liver fat (%) and liver cT1 (ms) were combined with additional biomarkers including glycated haemoglobin (HbA1c), age (yrs), sex (M/F), BMI (kg/m^2), waist circumference (WC) (cm), low density lipoprotein (mmol/l), alanine transaminase (IU/L), aspartate aminotransferase (IU/L) and triglycerides (mmol/L), hypertension (yes/no), total cholesterol (mmol/L), triglycerides (mmol/L), bilirubin (umol/L), gamma glutamyl transferase (U/L).

These clinical variables were chosen in collaboration with Prof Cuthbertson and Prof Despres due to their use within primary diabetes care clinics, clinical trials, or relevance for studying the mechanisms of metabolic dysfunction. For example, HbA1c provides a measure of the level of glucose circulating within the blood and is used for the diagnosis of T2D (NICE guideline NG28), while as discussed previously, the BMI is used for risk stratifying patients based on their obesity status. AST and ALT are a type of enzyme secreted by the liver that can be measured from blood plasma and are commonly used for assessing liver health. Elevated measures of AST and ALT are associated with liver dysfunction and are used in common scoring criteria such as the FIB-4 for the screening of high-risk patients to assess biopsy

requirements. However, as highlighted in the Introduction and *Obesity* publication, circulating biomarkers are limited in their diagnostic utility and should be combined with non-invasive imaging for a direct assessment of liver health. Such limitations are overcome in this Thesis by the inclusion of MR-derived measures of liver PDFF and cT1. The current project also included measures of VAT, SAT, and SMI to provide a quantitative measure of fat distribution and body composition, overcoming the aforementioned limitations of the BMI and to explore the associations with metabolic health.

4.2.2. Deriving the Bayesian-network structure

As detailed in the previous Chapter, the structure of a Bayesian-network can be derived using either automated, human or a combination of both automated and human approaches. Structure discovery algorithms are used to select an *optimal* model (G^*) from a set of possible candidates (G), where given a select data set D = {v^1, v^2, …, v^n), we return G* that maximises the following:

$$G^* = \arg \max_{G \in G^n} f\,(G : D)$$

where $f\,(G : D)$ is a scoring criterion that ranks candidate G's respective to their 'goodness of fit' to D.

To explore the potential of automated algorithms for novel structure learning, I compared the network structures returned by automated search algorithms against a 'semantic' network constructed in collaboration with medical experts. This semantic network was contrived by working with Prof Cuthbertson and Prof Despres to review relevant literature and subsequently insert and orientate edges to relate pairs of variables deemed to be *causally*

connected. For example, the edge from liver PDFF (fat) to liver cT1 was inserted due to the well-established causal connection whereby hepatocyte fat accumulation leads to cellular degradation of the hepatocytes and an influx of inflammation into the extracellular space (as detected by cT1) [113]. Other edges are perhaps less direct, but nevertheless still *causal* in nature. For example, the edge from Sex to VAT and SAT was inserted due to the well cited literature showing a sex specific distribution of body fat, whereby men typically accumulate greater VAT and women greater SAT due to hormonal factors such as levels of testosterone [114]. See Table 4.1 for some example edge connections and corresponding supporting literature.

Table 4.1 Example edge connections and supporting literature inserted into the semantic network.

Edge connection	Supporting literature
Liver PDFF → liver cT1	Hepatocyte fat accumulation leads to cellular degradation and an influx of inflammation within the extracellular space (as measured by cT1) [115].
SMI → HbA1c	Skeletal muscle plays a critical role in postprandial glucose disposal through insulin-stimulated recruitment of the glucose transporter protein 4 (GLUT-4), allowing for uptake of glucose into the skeletal muscle cells (myocytes) [116], [117].
Sex → VAT/SAT	Literature supports a sex-dependence on fat deposition, whereby men typically accumulate greater VAT and women

	greater SAT due to hormonal factors such as levels of testosterone [114].
Liver cT1 → AST/ALT	AST and ALT are enzymes found in the liver and are elevated when the liver is damaged. Although AST/ALT are not caused by the value of cT1 – an MR metric – cT1 is a tightly correlated marker of liver pathology. As such, an increase in liver cT1 (liver disease activity) will be associated with corresponding changes in measure of AST and ALT.
VAT/SAT/SMI → BMI	The BMI is constituted by one's height and weight, which is comprised of their body composition.

Once the structure had been established, the next task was to digitise the structure in order to perform the subsequence probabilistic inference. To do so, a software platform needed to be selected that provided the capabilities of end-to-end network construction and inference. Platforms considered included those written in either the R or Python programming languages and included the bnlearn I, dbnleaI(R), bnsIct I, abIR), dbnr (R), CausalNex (Python) and Hugin packages. After reviewing and assessing each package in turn, the 'bnlearn' package within the R-studio platform was selected due to its rich capabilities (for example, the ability to illustrate networks using the GraphViz package) and user-friendly means of constructing the networks using both automated and manual means and performing probabilistic inferences.

4.2.3. Comparison of automated and expert-led network construction

All structure learning algorithms, manual network construction and statistical analysis were therefore completed using the 'bnlearn' package in R Studio. Six automated network structures were obtained through the following six structure learning –lgorithms: SS - Greedy Search hill climbing (HC) & Tabu search (TS); CI – PC stable (PC) & Incremental Association (IAMB); hybrid – Max-min hill climbing (MMHC) & hybrid parents and children (H2PC). Each SS and hybrid algorithm was completed twice, firstly adopting the BIC and secondly the AIC scoring criterion for model selection, resulting in a total of 10 automated network structures. See Appendix for an overview of the selected structure learning algorithms.

4.2.4. Network comparison metrics

The automated network structures were then scored against the semantic network using the Structural Hamming Distance (whereby a lower score is considered superior), Topological network score (whereby a higher score is considered superior) and Global network score (whereby a higher score is considered superior), as proposed by [118].

1) *Structural Hamming Distance (SHD):* The SHD between is determined by the number of operations (edge insertion, deletion, or reversal) required to make two network structures match. Such operations are uniformly penalised, that is, the SHD increases by +1 regardless of whether an edge insertion, removal or reversal is required.
2) *Topological score:* The topological score is the normalised, combined weighted measure of the number of edges missing from the automated network (missing

edges), number of edges returned in the automated network that are not present in the gold standard network (added edges), and the number of edges (regardless of orientation) present in both the automated and gold-standard network (correct edges).

$$\frac{\sum Correct\ edges}{\sum Correct\ edges + \sum Missing\ edges + \sum Added\ edges}$$

3) *Global score:* The global score is the normalised, combined weighted measure of the number of edges missing from the automated network (missing edges), number of edges returned in the automated network that are not present in the gold standard network (additional edges), number of edges with incorrect orientation present in the automated network (wrongly oriented edges) and the number of correctly orientated edges in the automated network (correctly oriented edges).

$$\frac{\sum Correctly\ orientated\ edges}{\sum Correctly\ orientated\ edges + \sum Wrongly\ orientated\ edges + \sum Additional\ edges + \sum Missing\ edges}$$

4.2.5. Results from network comparison

Table 4.2 Comparison metrics of automated network structures against semantic network.

Algorithm	SHD	Topological score	Global score	BIC score	AIC score
HC (BIC)	22	0.45	0.28	-11569	-11460
HC (AIC)	37	0.36	0.23	-11622	-11425
TS (BIC)	22	0.45	0.28	-11569	-11460
TS (AIC)	37	0.36	0.23	-11622	-11425
MMHC (BIC)	15	0.60	0.26	-11686	-11595
MMHC (AIC)	15	0.60	0.26	-11686	-11595
H2PC (BIC)	20	0.33	0.10	-11777	-11703
H2PC (AIC)	20	0.33	0.10	-11777	-11703
PC.stable	13	0.60	0.33	-	-
Fast.Iamb	23	0.38	0.08	-	-

SHD = Structural Hamming Distance, BIC = Bayesian Information Criterion, AIC = Akaike's Information Criterion. BIC and AIC scores

Overall, this analysis highlighted the limitation of automated approaches for Bayesian-network structure learning. Firstly, when compared to our 'semantic' network, we found a large discrepancy in network score across all structure discovery algorithms and scoring criterion, especially when the performance metric considers edge orientation. We found the PC.stable algorithm achieved the closest match to our 'semantic' network, however the

performance margin compared the other automated networks was trivial and this network included non-directed edges, rendering them unsuitable for biological inference regardless. Furthermore, I observed several non-sensical edges within the automated networks, such as edges from 'Liver PDFF' directed to the variables 'Sex' or 'Age', which of course, is biologically implausible. Automated structure learning has proven to be NP-hard [79], with [80] showing that the number of candidate networks increases exponentially with the number of variables. Furthermore, optimisation of this task is a highly non-convex problem, meaning there are possibly several local minima/maxima found within the optimisation search space and the algorithm will likely struggle to return an 'optimal' network. Therefore, the automated search algorithms cannot be expected to deliver results that are consistent with clinical knowledge, for of course, they do not enshrine any such knowledge.

Consequently, for the current project and those contained within this Thesis, the decision was made to take a hybrid approach to constructing the BN structure. This includes deriving a BN structure using the Max-Min Hill Climbing algorithm [89] to form a basic network skeleton, and then adjusting this network in collaboration with Prof Cuthbertson and Prof Jean-Pierre by removing, reversing, or inserting edges based on the prior network, resulting in the final 'semantic' BN, see Fig. 4.1

4.3. Application of probabilistic reasoning

The next step following the digitisation of our BN structure was to perform the probabilistic reasoning in order to derive clinically meaningful inferences from the network. Again, this was completed in collaboration with Prof Cuthbertson and Prof Despres who guided the questions

that would not only be clinically meaningful but also aid their clinical decision-making. As detailed in the 'Bayesian networks and probabilistic inference' Chapter, probabilistic inference is performed by posing conditional probabilities whereby the probability of a specific 'event' is estimated, given an 'evidence list' that contains the fixed values of network variables. In this project, I applied probabilistic reasoning to estimate the probability of diabetic status hyperglycaemia (DHG) (HbA1c >48mmol/L) in specified example phenotypes.

To do so, conditional probability queries were estimated using a likelihood weighting technique, an importance sampling technique that utilizes Monte-Carlo approximation. In the case of continuous variables, the value of the variable is sampled from a uniform distribution between a fixed interval. For example, 'Low' and 'elevated' measures of variables such as VAT and SAT were determined at the 25th and 75th percentile values, respectively, derived from descriptive statistics. This allowed us to explore specific phenotypes, such as those with 'low' SAT but 'elevated' VAT and liver fat, or the individual effect of 'elevated' VAT alone on the probability of hyperglycaemia.

4.4. Results

4.4.1. Descriptive statistics of population

As expected, HbA1c was significantly elevated in those with T2D ($p<0.001$), while we observed no significant differences in BMI ($p=0.25$), SAT ($p=0.47$) and waist circumference (WC) ($p=0.16$) between those with and without T2D. VAT ($p=0.042$), liver fat ($p<0.001$) and liver cT1 ($p<0.001$) were significantly elevated in those with T2D, while SMI ($p=0.018$) was significantly lower in those with T2D.

4.4.2. Estimating the probability of hyperglycemia

The baseline probability of diabetic status hyperglycaemia (DHG) in our population was 24%. We then estimated the probability of DHG after fixing the value of specific variables within the BN, see Table 4.1.

Specifying an example case study with elevated VAT, elevated SAT, liver fat >10% (severe steatosis) and reduced SMI, we observed a 39% probability of DHG. This was reduced by 11% (28%) when specifying 'normal' liver fat (<5.6%), 13% (26%) when specifying 'low' measures of VAT, and 9% (30%) when specifying 'normal' SMI. 'Normal' liver fat (<5.6%) and SMI together caused a 19% (20%) reduction in the probability of DHG. Interestingly, we saw a 13% increase (51%) in the probability of DHG when specifying 'low' measures of SAT, suggesting that a high VAT, low SAT phenotype is the most 'at risk' of diabetic status DHG.

Table 4.1 Probability of DHG after fixing the value (evidence) of specific network variables.

Evidence	Probability of hyperglycaemia (%)
Obesity status	
Normal weight (<25kg/m^2)	23%
Overweigh (25-30 kg/m^2)	23%
Obese (30-40 kg/m^2)	26%
Severe obesity (>40 kg/m^2)	31%
Body composition	
Low VAT	22%
Elevated VAT	26%
Low SAT	22%
Elevated SAT	24%
Low SMI	27%
Elevated SMI	19%
Liver markers	
Normal liver fat (<5.6%)	19%
Moderate steatosis (5.6-10%)	23%
Severe steatosis (>10%)	29%

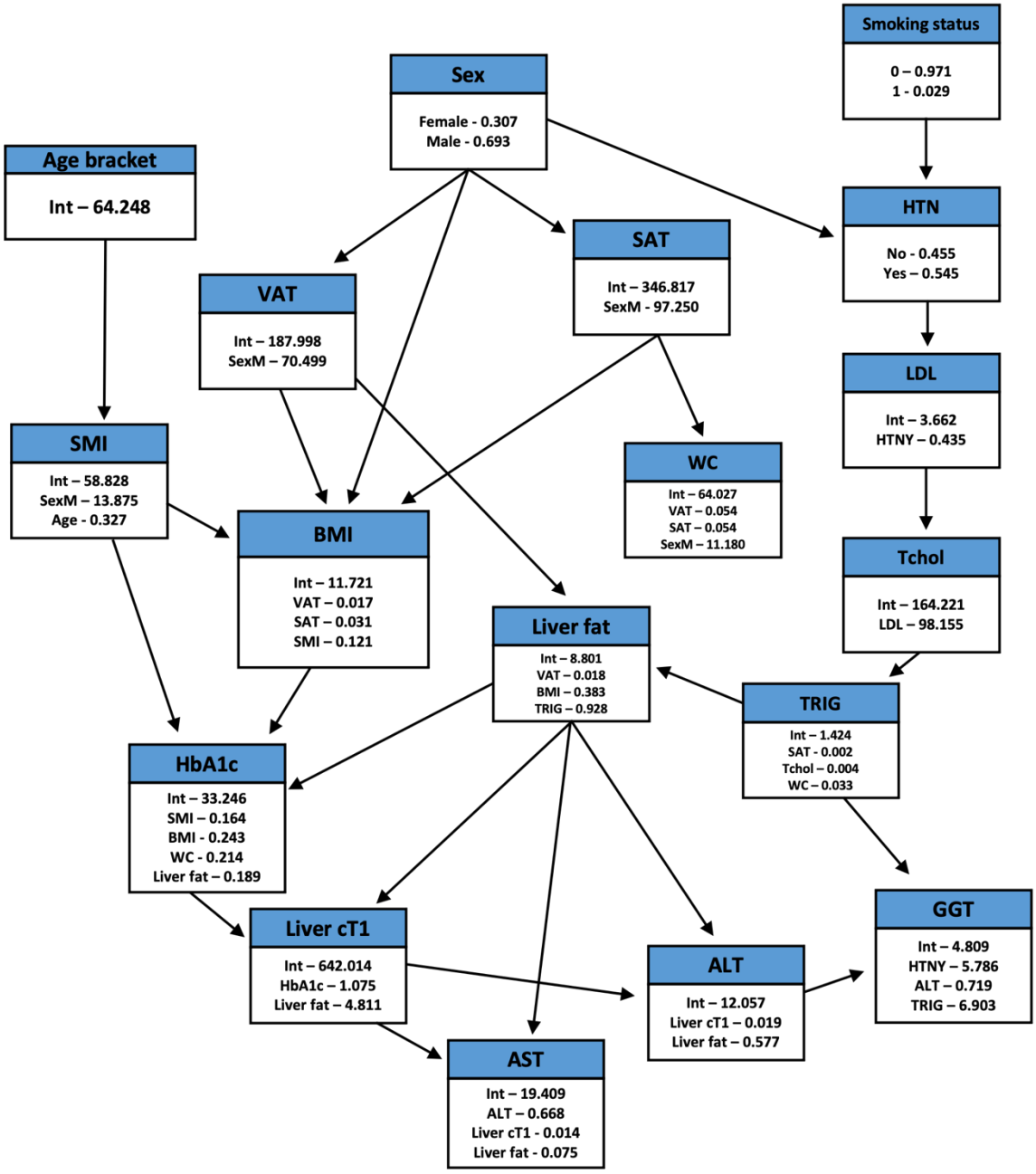


Figure. 4.1 Final 'semantic' Bayesian-network structure following manual adjustment. Values within each variable denote the coefficient value, given the parents of that specific variable. L_PDFF = liver fat; L_cT1 = liver cT1; GGT = gamma GT; HTN = hypertension; TRIG = triglycerides; SMI = skeletal muscle index; VAT = visceral adipose tissue; SAT = subcutaneous adipose tissue; Tchol = total cholesterol; WC = waist circumference; LDL = low-density lipoprotein; AST = aspartate transaminase; ALT = alanine transaminase; BMI = body mass index; Int = intercept.

4.5. Discussion

In this first development of a BN and experiment, we applied BNs and probabilistic reasoning to explore the direct effect of fat deposition on the probability of diabetic status DHG. Overall, our work immediately highlighted that the development of T2D is a complex interplay of fat deposition in multiple, as opposite to single, sites, notably as VAT and within the liver. Such an approach can help to unravel the pathophysiology of disease and in the process, potential therapeutic strategies for disease prevention and treatment. In isolation, specifying elevated measures of VAT or SAT did not elicit discernible changes in the probability of DHG (+4% and +2%, respectively). However, specifying a phenotype classified by elevated VAT, liver fat >10% and low SAT resulted in a 51% probability of DHG, consistent with reports suggesting that SAT may serve as a 'protective' mechanism against dyslipidemia and disturbed glucose metabolism [119].

This finding is also notable for it emphasises how the weight of a person is constituted (that is, their *body composition*) needs to be considered for studying metabolic disease risk. This is because such a person (given they present with low SMI and SAT), may be classified as normal weight despite being at a significant risk of T2D. In this example patient, the clinical care may be different to the usual advice for those with overweight or obesity, where a 5-10% weight loss is considered a primary treatment target (NICE guideline NG28). For example, following a Mediterranean diet has shown to elicit favourable changes in body composition and metabolic health without requiring such significant weight loss [120], and therefore may be encouraged for this specific patient. However, it is important to highlight the unbalanced groupings in this analysis, notably 'Sex' and 'Smoking status', whereby there are considerably

fewer females and smokers in the cohort. Such unbalancing may lead to inaccuracies when performing statistical inferences, notably probabilistic inference, for the model will be biased and drawing on far fewer females and smokers when performing random sampling. The results here should therefore be treated with caution and instead an example of how Bayesian-networks can be applied.

As discussed in the Introduction, this finding also shows the utility of BNs for exploring the topic of metabolically healthy and unhealthy obesity, with the results here again indicating that the deposition of fat in ectopic sites is a primary differentiating factor between these two groups. This is highlighted through the observation that the probability of DHG increased by 10% when specifying elevated fat (>10% PDFF), again showing the risk association between concomitant NAFLD and T2D. Such finding is supported by the 'Twin Cycle Hypothesis' proposed by Prof Roy Taylor and colleagues (see Fig. 4.2) which places elevated fat deposition as a central driver in the pathophysiology of T2D [121].

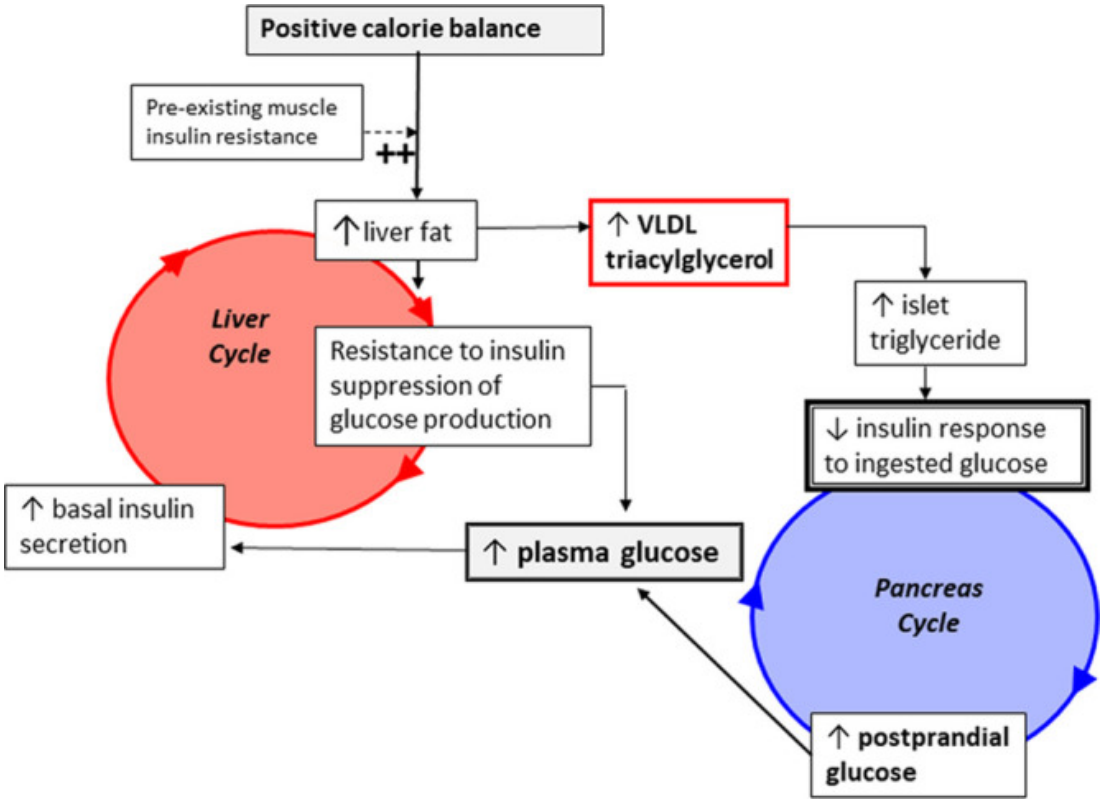


Figure 4.2 Illustration of the twin-cycle hypothesis taken from Taylor et al. [121]. Elevated plasma glucose is driven by the liver and pancreas cycles, whereby elevated fat deposition drives de-novo lipogenesis and dysregulated insulin secretion.

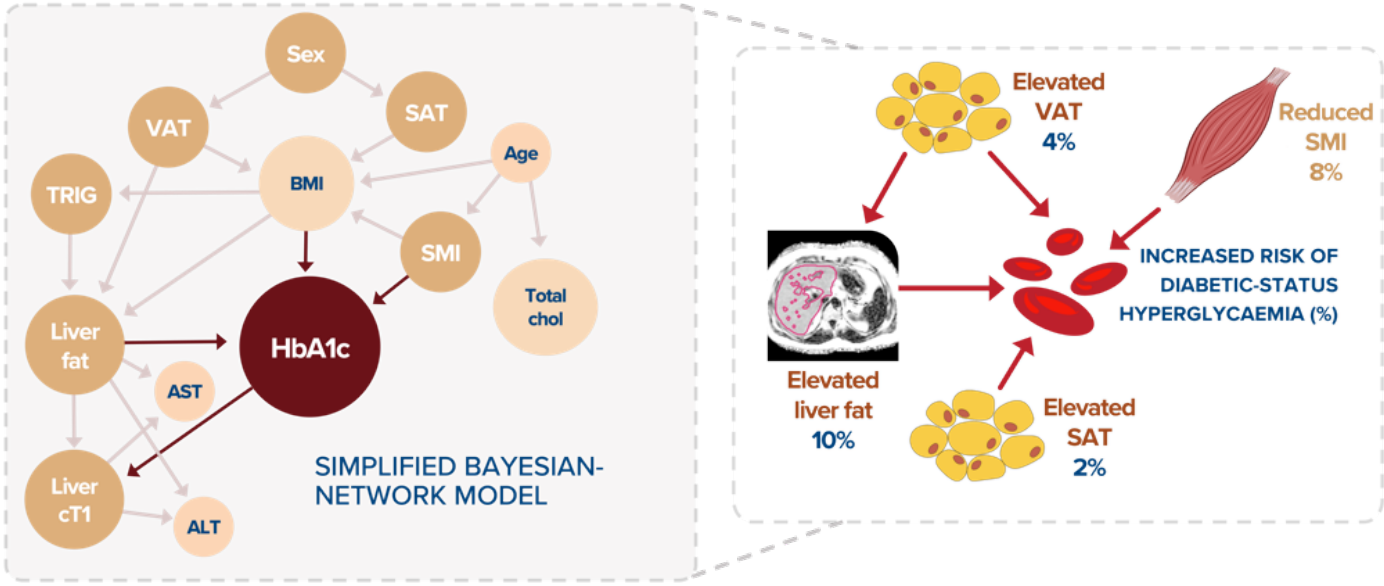


Figure 4.3 Illustration showing the simplified Bayesian-network structure (left) and corresponding probability increases of DHG (right). TRIG = triglycerides; AST = aspartate transaminase; ALT = alanine aminotransferase; chol = cholesterol.

4.6. Conclusion

This Chapter exemplifies the complexity of developing a BN for clinical use, specifically the need to relate causal links to the literature and how BNs can provide a rich understanding of the interplay between fat distribution and hyperglycemia. We walked through the inception of a semantic BN for modelling a multi-organ disease. This included collaborating with medical experts for the end-to-end project; including manually constructing a network structure, digitizing such structure using a programming platform and then working with Prof Cuthbertson and Prof Despres to steer the conditional probability queries to derive clinically meaningful inferences. Notably, this project highlighted that a high VAT/low SAT phenotype resulted in the greatest probability of hyperglycemia, which though consistent with literature, wasn't revealed in the descriptive statistics of body composition within the *Obesity* publication. This is not to say that such traditional descriptive statistical approaches are inadequate, rather the two techniques are simply posing different questions of the data. The inferences posed by Bayesian-networks allow for specific phenotypes to be explored and the intricacies of metabolic risk discovered.

Returning to the 'Case study 2' described in the Introduction of this Thesis, we can now see how the application of Bayesian-networks can enable the necessary reasoning for assisting

with such clinical decision-making. While this Chapter did not address such questions directly, namely the effect of different treatment regimens on the probability of nausea (as per case study 2), it serves as a proof of concept that Bayesian-networks are capable of posing and answering such queries. For example, with the necessary data we could estimate the probability of nausea and GI after fixing the treatment to either Metformin only or Metformin plus Semaglutide, while also estimating the effect on NASH and hyperglycaemia risk using the same methodology applied in this Chapter. We expand on the application of Bayesian-networks for modelling treatment outcomes in greater detail in the 'Future work' Chapter of this Thesis. Rather, in the next Chapter, we turn to a different application of Bayesian-networks by exploring the prevalence of obesity within the UK and use Bayesian-networks to model the potential consequences of projected increases in obesity prevalence on a common comorbidity – non-alcoholic liver disease.

5

APPLYING BAYESIAN-NETWORKS TO STUDY THE IMPACT OF PROJECTED INCREASES IN OBESITY PREVALENCE ON INCIDENT LIVER DISEASE IN THE UK

5.1. Introduction

In the previous Chapter we applied BNs to explore the distinct phenotypes driving diabetic-status hyperglycaemia at an individual case study level. But what about disease forecasting and epidemiological modelling through posing counterfactuals at a population level? After all, the question of estimating the prevalence of disease *x*, given an increase in *y,* is an inherently counterfactual question.

As highlighted in the Introduction, obesity is a complex multi-system disease and a growing global public health challenge. Obesity is significantly associated with a greater incidence of T2D, NAFLD and NASH. Furthermore, we also highlighted the stark projections made by Cancer Research UK in which, if current trends continue, by 2040 71% of the adult population in England (36 million people) will be living with overweight or obesity [11]. This equates to an 11% increase in the prevalence of overweight and obesity from 64% in 2019 that threatens to engulf an already-overstretched public health system. Such concern is highlighted through the astonishing estimated 385% relative increase in obesity-related costs from £2 billion to £9.7 billion between 2001 and 2050 [5].

But what do such projected increases in obesity mean for one of the most common obesity-related co-morbidities, non-alcoholic liver disease? Evidence suggests that currently one in three UK adults have NAFLD, of which one in five will go on to develop NASH, the most common indication for liver transplantation in the UK [122]. More than 12% of liver transplantations in the UK are due to obesity-related liver disease, with the healthcare costs of NAFLD-related hepatocellular carcinoma estimated at up to £10.5 billion [123]. Furthermore, the per-person cost of diagnosed NASH in the UK is estimated at £9,450, corresponding to a total economic cost between £2.3 billion (low prevalence scenario) and £4.2 billion (high-prevalence scenario) [28]. A major challenge is that early-stage NASH is typically asymptomatic, with an estimated 79.8% of prevalent NASH in the UK currently going undiagnosed [28].

It has been difficult to model obesity and to develop clinician decision support systems because it is multi-faceted, involving a complex interplay between many variables including age, sex and fat deposition within the liver and as visceral adiposity. To address such complexity and explore the potential impact of projected obesity prevalence on liver disease, this Chapter illustrates how BNs can be applied to model relationships such as that between liver fat and the BMI. Specifically, we apply BNs to determine the impact of the projected increases in obesity prevalence on metabolic health, in particular metabolic-associated liver disease. We have based current estimates of the prevalence of liver disease from MRI-techniques that provide quantitative measures of liver fat and fibroinflammation with results taken from the imaging dataset acquired in the UK Biobank (UKBB).

Similar to the work contained in the previous Chapter, we included validated MRI measurements at each voxel of proton density fat fraction (PDFF) and liver tissue inflammation and fibrosis (cT1) [39]. Using the current prevalence of liver steatosis (high PDFF) and fibroinflammation (high cT1) according to the prevalence of overweight and obesity in the UKBB, we have been able to model the projected changes in NAFLD from estimates of future obesity prevalence. This work was accepted for an oral presentation at the 2023 International Conference of Health and Informatics and poster presentation at the 2023 European Association for the Study of Liver Disease.

Project citation

Waddell, T., Namburete, A., Duckworth, P., Cuthbertson, D., McCracken, C. and Brady, J.M., 2023. The impact of projected increases in obesity prevalence on incident liver disease in the UK: insights from Bayesian-network modelling. *Journal of Hepatology*, *78*, pp.S677-S678.

5.2. Methods

5.2.1. Data collection and preparation

Participant data were collected retrospectively from the UKBB online resource (Application 9914). As well as liver cT1 (ms); and liver PDFF (%), we collected data on: sex (male/female); age (years); smoking status (current/previous/never); hypertension status (yes/no); reported high cholesterol (yes/no); previous myocardial infarction (yes/no), BMI (kg/m^2); waist circumference (cm); HbA1c (mmol/L); Low Density Lipids (LDL-c) (mmol/L); HDL-c (mmol/L); and other frequently used enzymes namely ALT (alanine transaminase) [IU/L]; AST (aspartate aminotransferase) [IU/L]; triglycerides (mmol/L). Each continuously varying variable was

discretized based on pre-defined clinical thresholds. For example, HbA1c was divided into: 'normal' (HbA1c <42mmol/mol); 'pre-diabetes' (42-48mmol/mol); and 'diabetes' (HbA1c >=48mmol/mol). The presence of NASH was based on liver PDFF >5.6% and liver cT1 >800ms. As noted previously, such thresholds were chosen for they have consistently demonstrated high diagnostic accuracy for stratifying patients with NASH and classifying between NAFLD and NASH [47].

5.2.2 Bayesian-network construction and inference

We adopted the same methodology for BN construction and inference as detailed in the previous Chapter. This included working in collaboration with Prof Cuthbertson to build a semantic network and then digitizing the network using the 'bnlearn' package within the R software platform. This network was then visualized using the 'graphviz' functionality within bnlearn. Again, an initial skeleton network was derived using only the MMHC algorithm to establish a base network in which to work with. This network was then adjusted by removing or reversing non-sensical edges and inserting edges based on the prior, semantic network, see Fig 5.1. In total, 11 changes were made to the initial network, including four edge removals, four edge reversal and three edge insertions. For example, the edge from 'H_chol' to 'Age' was reversed, for of course, one's cholesterol level cannot affect their age. Of course, the application of experts' prior knowledge when correcting such network structure can also raise limitations in this analysis, for the experts themselves may be biased or choose to insert/correct edges dependent on flawed medical literature that is not reflective of the true underlying biology. This may be due to out-of-date literature that has since been corrected,

or a general bias of the clinician from their personal experiences within the clinics or with patients.

Unlike the previous Chapter, however, when performing the probabilistic inference, we did not set an event to estimate, given the fixed values set within an evidence list. Rather, we fixed the probability values of the variable 'Obesity' in line with the 2040 projections and observed the 'ripple' effects throughout the rest of the network, observing the resultant changes in network variables consequently. This analysis is described in greater detail below under 'estimating consequences of increased obesity prevalence'. Descriptive statistics, showing median and interquartile ranges [IQR], were calculated to summarize population characteristics.

5.3. Results

We first identified a total of 41,994 participants from the UKBB online resource. We then removed participants whose associated data did not include values for the variables listed above (for example, those with missing liver cT1 data). This resulted in 27,002 participants, who were included for statistical analysis and BN construction. The final BN structure following manual adjustment is shown in Fig. 5.1.

5.3.1. Population characteristics

Based on the BMI, the prevalence of normal weight, overweight, obesity, and severe obesity in the analysed population was 41.3%, 41.1%, 16.8% and 0.8%, respectively. Of these, 12% and 10%, respectively, had liver fat values between 5.6-10% (mild steatosis) and greater than 10% (severe steatosis). 4.4% met the criteria for NASH (PDFF >5.6% & cT1 >800ms).

Table 5.1. Descriptive statistics of population cohort with complete data. Data is median [IQR].

Variable	Whole population (n = 27,002)
Liver markers	
Liver fat (%)	3 [2-5]
Liver cT1 (ms)	694 [662-730]
Demographics	
Sex (% M)	48%
Age (yrs)	56 [49-61]
Anthropometrics	
BMI (kg/m^2)	25 [23-28]
% Normal weight	41%
% Overweight	41%
% Obesity	16%
% Severe obesity	<1%
Waist circumference (cm)	88 [78-96]
Circulating biomarkers	
HbA1c (mmol/L)	34 [32-37]
HDL (mmol/L)	1.4 [1.2-1.7]
Triglycerides (mmol/L	1.3 [0.9-2]
Medical history	
Hypertension (% yes)	33%
High cholesterol (% yes)	29%
Myocardial infarction (% yes)	2.4%

5.3.2. Estimating consequences of increased obesity prevalence

Given the baseline probabilities of being normal weight, overweight, obese, and severely obese, we can estimate the consequences of increased obesity prevalence on liver health and metabolic markers. To this end, we fixed the probabilities of being normal weight, overweight, obese, and severely obese to 29%, 27%, 37% and 7%, respectively, in line with 2040 projections. The updated conditional probability tables, given the specified evidence, were then investigated to measure the resultant changes (if any) in the remaining network variable states, see Table 5.2. It is important to note here, however, that by only fixing the variable 'Obesity' only, we are not exploring the influence of other components such as smoking and ethnicity which are known to influence health outcomes. Future work could explore such outcomes, for example the influence of reducing smoking or alleviating socioeconomic deprivation on the health outcomes explored here.

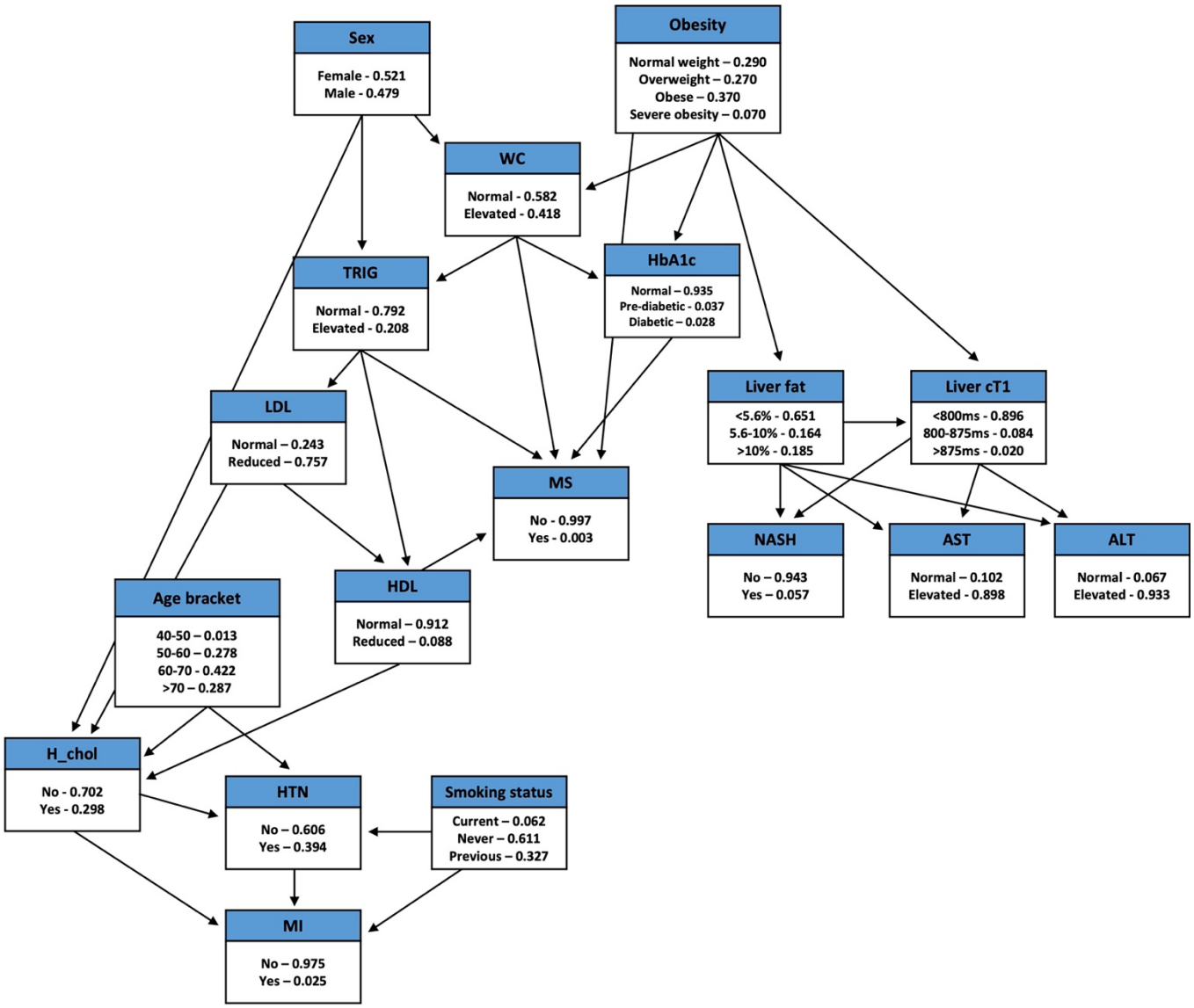


Figure 5.1. Final semantic Bayesian-network structure following intervention on 'Obesity'. WC = waist circumference; LDL = low density lipoprotein; HDL = high density lipoprotein; H_chol = high cholesterol; HTN = hypertension; MI = myocardial infarction; MS = metabolic syndrome.

Table 5.2. Baseline and estimated probabilities of variable states with corresponding change in probability and population number after intervening on the variable 'Obesity'.

Network variable (value)	**Baseline probability (%)**	**Probability following network intervention (%)**	**Change in probability (+/- %)**	**Corresponding increment in participant number (+/- n)**
Liver markers				
Liver fat (5.6-10%)	12%	17%	+5%	+1350
Liver fat (>10%)	10%	18%	+8%	+2160
NASH (yes)	2%	5%	+3%	+800
Circulating biomarkers				
HbA1c status (pre-diabetes)	2%	3%	+1%	+235
HbA1c status (diabetes)	1%	2%	+1%	+235
HDL (reduced)	8%	8%	-	-
Triglycerides (elevated)	18%	20%	+2%	+470
Misc. health parameters				
Hypertension (yes)	33%	39%	+6%	+1620

Metabolic syndrome (yes)	2%	2%	-	-
High cholesterol (yes)	29%	29%	-	-
Waist circumference (elevated)	24%	41%	+17%	+3240
Myocardial infarction (yes)	2%	2%	-0	-

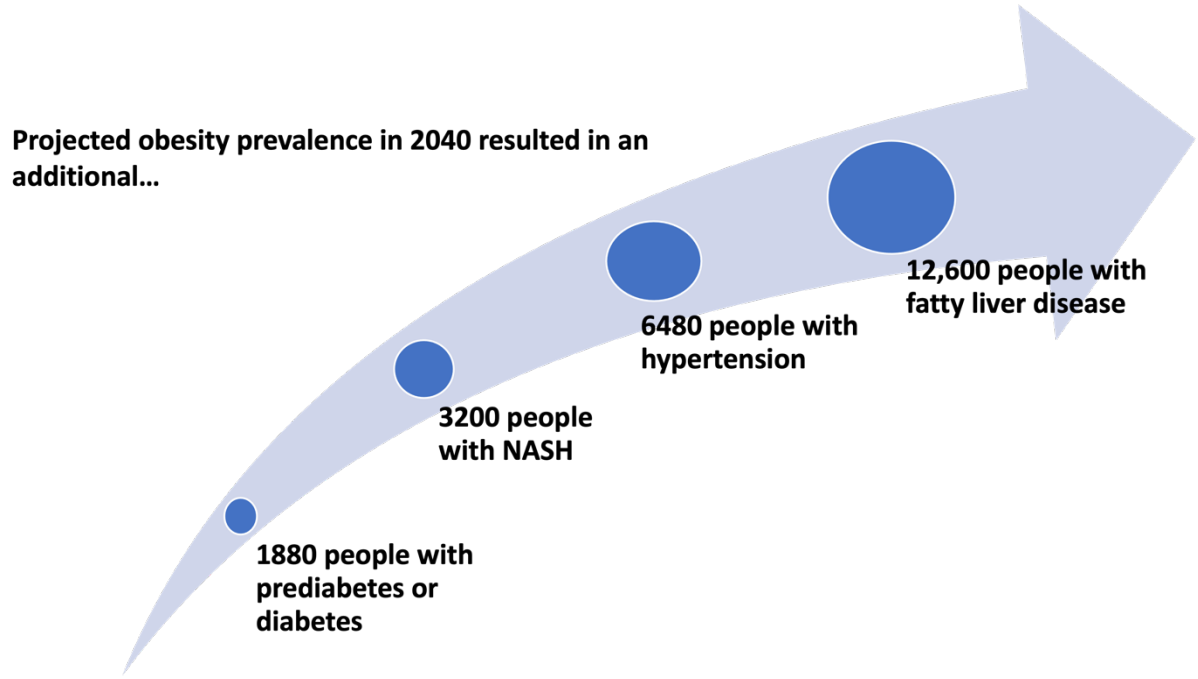


Figure 5.2. Illustration of the estimated key consequences of increased obesity on liver and metabolic health parameters, per ~108k people.

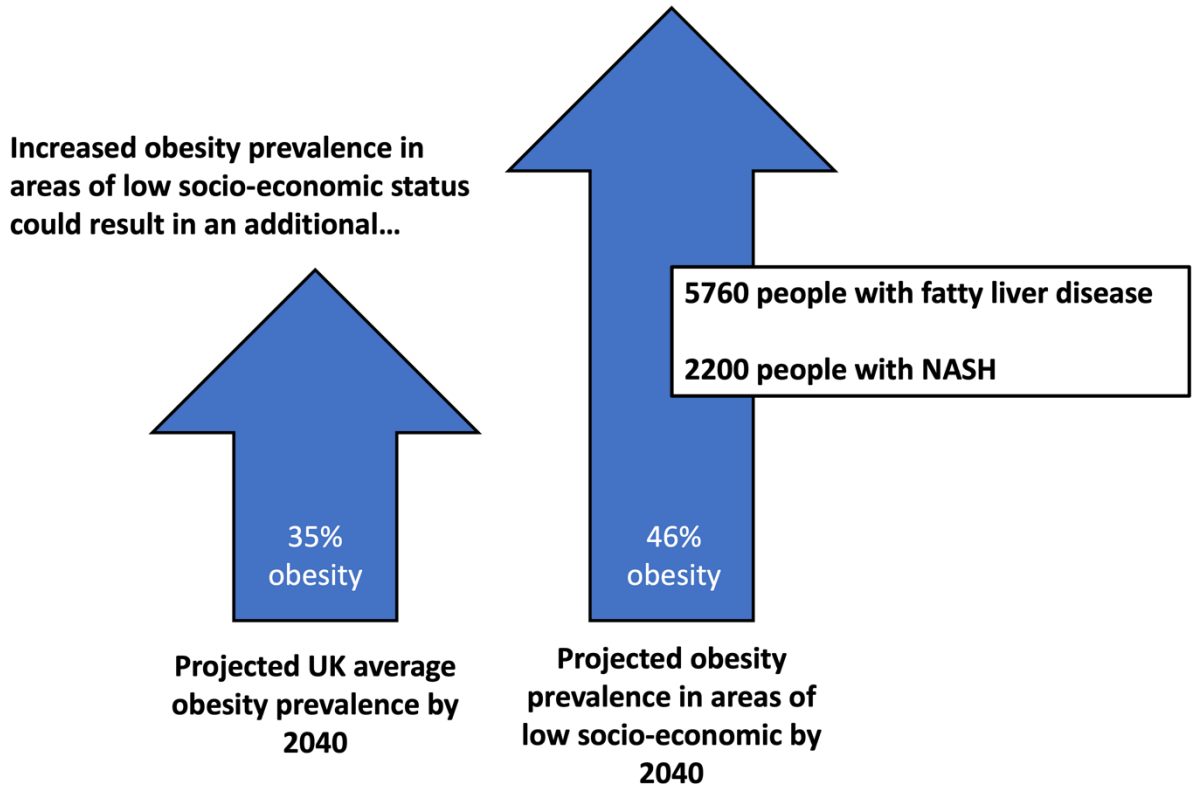


Figure 5.3. Illustration of the estimated impact of increased obesity prevalence in areas of low socio-economic status on liver disease, per ~108k people.

5.4. Discussion

The prevalence of people living with overweight and obesity in the UK continues to increase rapidly, posing massive challenges to healthcare systems. Applying 2040 projections of the prevalence of obesity to a large UK-based population cohort, we demonstrate how BNs can be applied to estimate the resultant effect on measures of liver and metabolic health. We show probabilistic inference estimates an alarming rise in the rates of liver disease with a 5%, 8% and 3% increase in the probabilities of moderate steatosis (5.6-10%), severe steatosis (>10%) and NASH, respectively. This projects to an additional 12,600 people with fatty liver disease and 3200 people with NASH per ~108k people, still just a small fraction of the UK population of ~67 million people.

The economic implication of this change is significant. With the economic costs of diagnosed NASH estimated at £9,450 per person [28], this equates to an additional £~7.5 million in direct NASH-related costs for this sub-population of 27,002 people, which is <1% of the entire UK population. However, this ignores the fact that fatty liver disease and early-stage NASH are typically asymptomatic, with an estimated 79.8% of the prevalent UK NASH population being undiagnosed in 2018. While a diagnosis of NASH increases the level of healthcare support (and therefore cost) for the given individual, the true direct and indirect costs associated with liver disease in the entire population are likely considerably higher.

We note that our analysis was conducted on a population that has much lower levels of obesity than in the general population (16% obesity vs UK population estimate = 29%). As such, our population could be considered 'low risk' and the estimations represent an underestimate. In geographical regions of socio-economic deprivation, with a typically higher prevalence of obesity and liver disease [124], the projected increase in obesity is likely to have a disproportionate impact on public health. To illustrate, in 2019 the obesity rate in the most deprived areas of England were already estimated to be 35%, the projected rate for the UK average in 2040, and is expected to increase to 46% by 2040 [11]. In our population, specifying a 46% probability of obesity resulted in an additional 4590 cases (+17% probability) with fatty liver disease and 1350 cases (+5% probability) with NASH, equating to an additional 1440 and 550 cases, respectively, above our previous UK average estimations.

Furthermore, our analysis projects a 6% (1620 cases) and 2% (470 cases) increase in the probability of hypertension and HbA1c above pre-diabetes levels, respectively. While we observed no discernible changes in the probability of myocardial infarction, both

hypertension and diabetes are significant risk factors in the development of cardiovascular complications and other microvascular and macrovascular complications [125]. Such outcomes, i.e. the lack of changes in the probability of MD and MS in response to increasing obesity, is surprising, and not in line with current literature as described in the Introduction of this Thesis. Therefore, these outcomes should be viewed with caution as a potential artefact due to the low incidence of these clinical events within the cohort. Rather, the analysis shown here should be seen as an example of how Bayesian-networks can be applied to model clinical biomarkers and outcomes, as opposed to providing definitive results on the projection of disease. The socio-economic consequences and costs of increasing levels of obesity are increasingly challenging. It is important to quantify those personal, public, and economic costs as accurately as possible, across different socio-economic strata.

5.5. Conclusion

This Chapter illustrates how graphical modelling and probabilistic reasoning, based on Bayesian networks, provides a novel method for estimating the personal, public, and economic costs associated with increasing obesity and liver disease in the UK. This Chapter provides a stark reminder of the increased and disproportionate burden of liver disease brought forth by the growing obesity challenge.

The objective of this project was to explore the application of Bayesian-networks for modelling clinical data and estimating counterfactual outcomes from a large population-based cohort. However, throughout this Thesis we have highlighted the limitations of the BMI and our current definition of 'obesity'. Furthermore, this Chapter did not explore the role of fat deposition and the specific mechanisms driving one of the severe consequences of

obesity explored here - NASH. By unraveling such mechanisms, we can begin to identify high-risk phenotypes and explore novel therapeutic options to assist clinicians with managing their patients. The following Chapter addresses these shortcomings by applying Bayesian-networks to model MR-derived body composition and liver health in a cohort of participants with and without concurrent T2D.

6

APPLYING BAYESIAN-NETWORKS TO STUDY CONCOMITANT NON-ALCOHOLIC STEATOHEPATITIS IN PEOPLE WITH TYPE-2 DIABETES

6.1. Introduction

As detailed in the Introduction, the global prevalence of both obesity and type-2 diabetes (T2D) has risen substantially in the past 50 years [7], with a projected 25% and 50% increase in T2D expected by 2030 and 2045, respectively [126]. We also highlighted the common co-prevalence of T2D and liver disease, notably how people with NAFLD and concomitant T2D have significantly worse liver-related outcomes including higher rates of advanced fibrosis, cirrhosis and liver-related cancers compared to those with NALFD only [127]–[130].

20-year projections of the economic and clinical burden of NASH/NAFLD estimate that co-prevalent NASH and T2D will account for 65,000 liver transplants, 812,000 liver-related deaths and 1.37 million cardiovascular-related deaths, totalling $55.8 billion in healthcare costs [131]. Early detection of those patients with T2D at high-risk of NASH is therefore of considerable importance and would enable early access to personalised care/medicines and improved clinical and disease outcomes. People with T2D represent a clinical population, and relative to those without T2D, are characterised by a distinct body composition profile with

significantly higher volumes of visceral adipose tissue (VAT), increased liver fat deposition and fibroinflammation and reduced skeletal muscle mass [20], [132].

By unraveling the mechanisms driving NASH risk we can begin to identify high-risk phenotypes and assist clinicians with managing their patients. In this Chapter, we use MRI-derived measures of body composition and liver health using PDFF and liver cT1 to identify participants with NAFLD and NASH and to overcome the intrinsic limitation of the BMI noted throughout this Thesis by exploring regional fat distribution. We describe a use case on the application of Bayesian-networks for studying the associations between MRI-derived measures of body composition and liver health. We highlight how Bayesian-networks have the potential to provide a comprehensive assessment of the high-risk metabolic phenotypes associated with co-prevalent T2D and NASH.

We also show how BNs can be used to identify potential therapeutic targets for alleviating NASH risk. The work contained in this Chapter was published in 2023 within the journal *Frontiers in Endocrinology.*

Paper citation

Waddell, T., Namburete, A., Duckworth, P., Fichera, A., Telford, A., Thomaides-Brears, H., Cuthbertson, D. and Brady, M., Poor glycaemic control and ectopic fat deposition mediates the increased risk of non-alcoholic steatohepatitis in high-risk populations with type-2 diabetes: insights from Bayesian-network modelling. *Frontiers in Endocrinology*, *14*, p.412.

6.2. Project methodology

6.2.1. Data collection & preparation

We investigated two datasets: the first included 221 participants with T2D pooled from the MODIFY study (NCT04114682) [133] and the UKBB online resource under application 9914. MODIFY recruited participants with T2D from primary or secondary care settings from three sites across the UK. The second cohort included 135 participants without diabetes of any kind (non-T2D) drawn from the UKBB. All participants underwent an abdominal MR scan that included multi-parametric imaging of the liver and pancreas.

6.2.2. Clinical details and MRI image acquisition and analysis

Body composition was examined from a 2D MR slice positioned at the third lumbar (L3) vertebra and measurements of VAT, SAT and skeletal muscle were based on manual delineations by trained analysts. The L3 section is typically selected because it has been shown to strongly correlate with whole-body skeletal muscle distribution and accurately estimate total subcutaneous (SAT) and visceral adipose tissue (VAT) volumes [49]–[51]. Measures of skeletal muscle were then indexed to the participant's height to produce a measure of skeletal muscle index (SMI) (cm^2/m^2). All continuous variables were discretized based on pre-defined clinical thresholds or by splitting the data into 'low' and 'high' value groups determined at the 75th percentile value, see Appendix. For example, HbA1c within the T2D cohort was discretised by splitting the cohort into 'low' (HbA1c <62mmol/mol) and 'high' (HbA1c >=62mmol/mol) value groups. Presence of NASH was again classified based on liver PDFF >5.6% and liver cT1 >800ms.

6.2.3. Bayesian-network construction

BN construction included the following variables: liver fat (PDFF %), liver cT1 (ms), pancreatic fat (PDFF %) visceral adipose tissue (cm^2), subcutaneous adipose tissue (SAT) (cm^2), skeletal muscle (SMI) (cm^2), sex (1[male]/0[female]), BMI (kg/m^2), age (yrs), HbA1c (mmol/mol), AST (IU/L), ALT (IU/L), smoking status (0[non-smoker]/1[current smoker]/2[past smoker]). As with previous Chapters, BN construction and probabilistic inference were completed using the 'bnlearn' package and visualised using 'graphviz' within the R software platform (version 3.6.1). Automated network structures, derived from the MMHC algorithm, were adjusted by removing or reversing nonsensical edges and inserting edges based on domain knowledge gleaned from medical literature and the prior semantic network, see Fig. 6.1. Again, to investigate the probability of NASH in specific clinical characteristics, we conducted conditional probability queries using the likelihood weighting algorithm.

6.2.4. Statistical analysis

All statistical analyses used the R software platform (version 3.6.1). Descriptive statistics, showing median [inter quartile range], are reported to summarise population characteristics. Wilcoxon and X^2 tests were used to make between group comparisons.

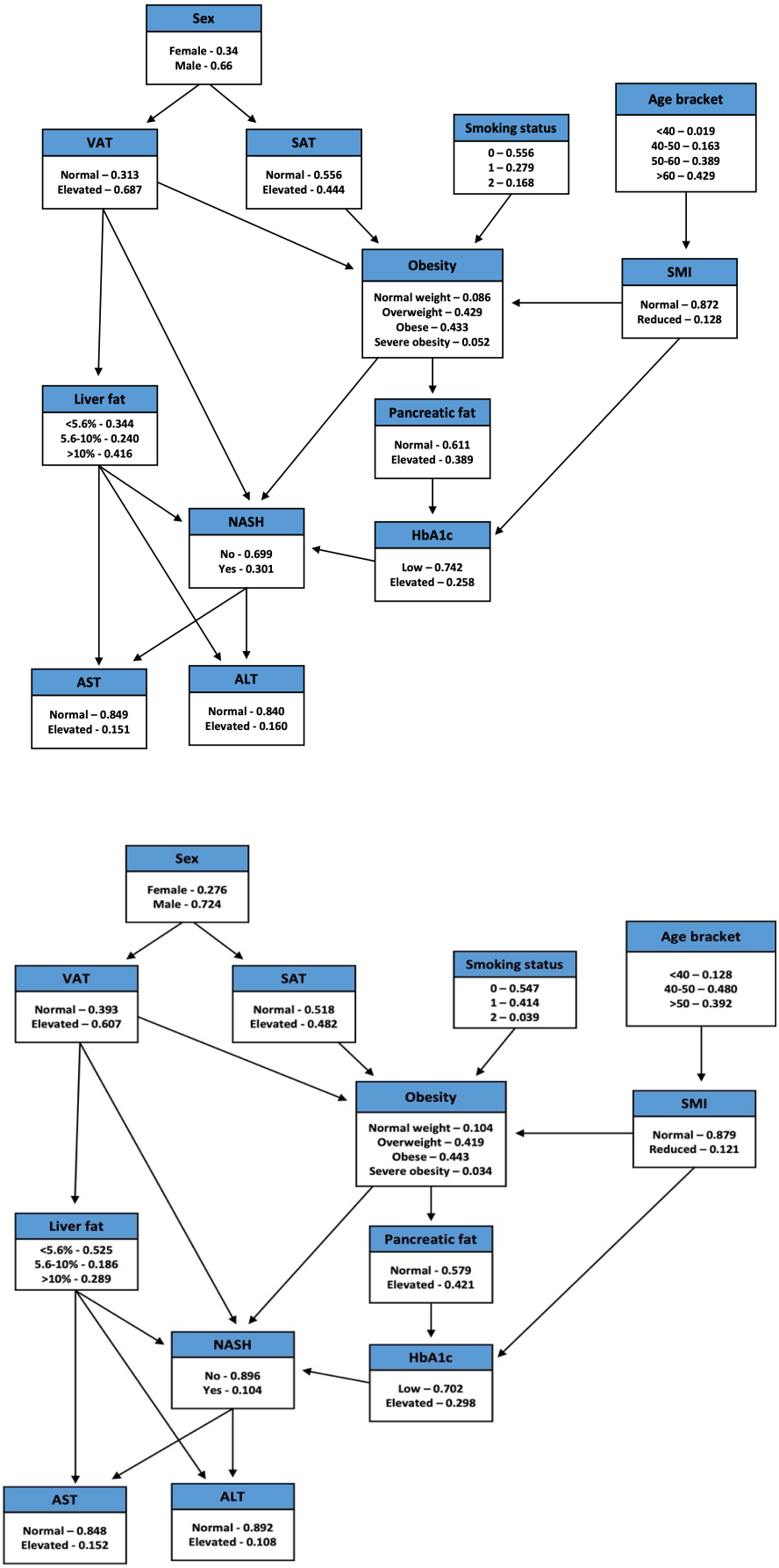


Figure 6.1. Bayesian-network structures from the T2D (top) and non-T2D (bottom) cohorts.

6.3. Results

Descriptive characteristics of participants

The baseline clinical, biochemical, and imaging characteristics are shown in Table 6.1. The descriptive statistics show there were no significant differences in age, sex (% male participants) and BMI between the T2D and NT2D groups. However, within both the T2D and non-T2D cohorts there were several sub-group imbalances, for example a greater percentage of male participants and poor balancing between the age bracket discretization. Therefore, the results from this analysis should be treated with caution and merely to demonstrate the potential application of Bayesian-networks. Furthermore, in the T2D cohort, the variable 'Age' was only discretised into three groups (not four as in non-T2D), meaning a direct comparison between the two networks with regards to the influence of age on clinical outcomes should be treated with caution. Despite similar liver biochemistry, participants from the T2D cohort (and thus higher HbA1c (<0.001)) had significantly elevated liver cT1 (<0.001), liver PDFF (<0.001), VAT/SAT ratio (p=0.034) and greater prevalence of NASH (<0.001).

Table 6.1. Baseline characteristics of T2D and non-T2D cohorts. Data is median [IQR]. Bold denotes significance.

Characteristic	**T2D (n = 221)**	**Non-T2D (n = 135)**	**p-value**
Clinical data			
Age (yrs)	57 [52-64]	57 [53-63]	0.3
Sex (n of male participants [%])	146 (66%)	98 (73%)	0.2
BMI (kg/m^2)	30 [27-34]	29 [26-33]	0.3

HbA1c (mmol/mol)	50 [41-62]	36 [33-38]	**<0.001**
AST (IU/L)	24 [19-29]	27 [23-32]	**<0.001**
ALT (IU/L)	27 [20-36]	26 [19-35]	0.5
MRI data			
Liver cT1 (ms)	751 [695-827]	709 [672-752]	**<0.001**
Liver fat (%)	8 [4-14]	5 [3-11]	**<0.001**
Pancreatic fat (%)	5 [3-9]	5 [3-8]	0.6
VAT (cm^2)	238 [173-307]	215 [149-279]	0.065
SAT (cm^2)	249 [180-322]	271 [181-366]	0.3
VAT/SAT ratio	0.90 [0.58-1.33]	0.79 [0.57-1.15]	0.034
SMI (cm^2/m^2)	49 [42-56]	51 [43-56]	0.6
NASH (n of participants [%])	69 (31%)	13 (9%)	**<0.001**

Estimation of NASH using probabilistic reasoning

The baseline probabilities of NASH in the T2D and non-T2D cohorts were 30% and 10%, respectively. 'Intervening' on a variable within the BN means that a specific value is assigned to it. For example, the variable 'HbA1c' may be assigned the values '0' ('low' value) or '1' ('high' value). The effect of amending level of glucose regulation is then propagated throughout the network, where the variables(s) conditionally dependent on the intervened variable(s) are updated to reflect the specified evidence.

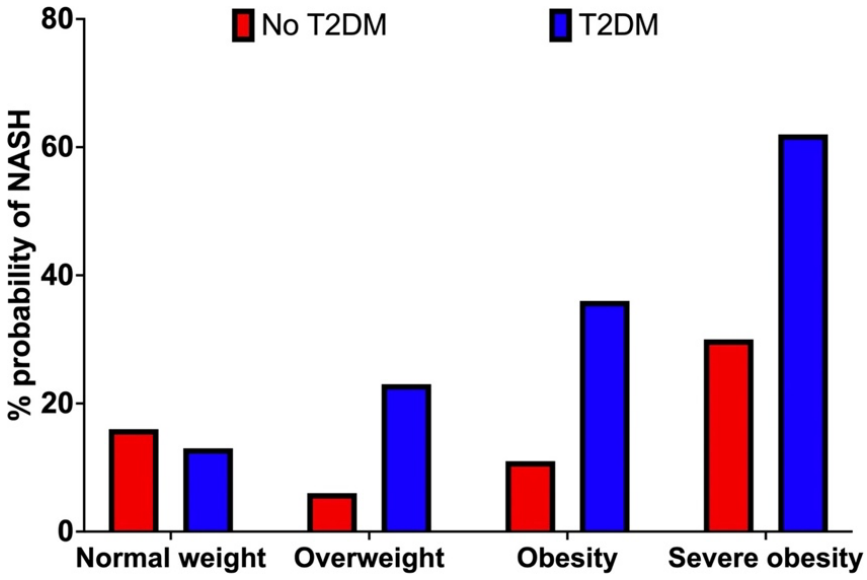


Figure 6.2. Probability of NASH (%) by obesity category (BMI) in the non-T2D (red) and T2D (blue) cohorts.

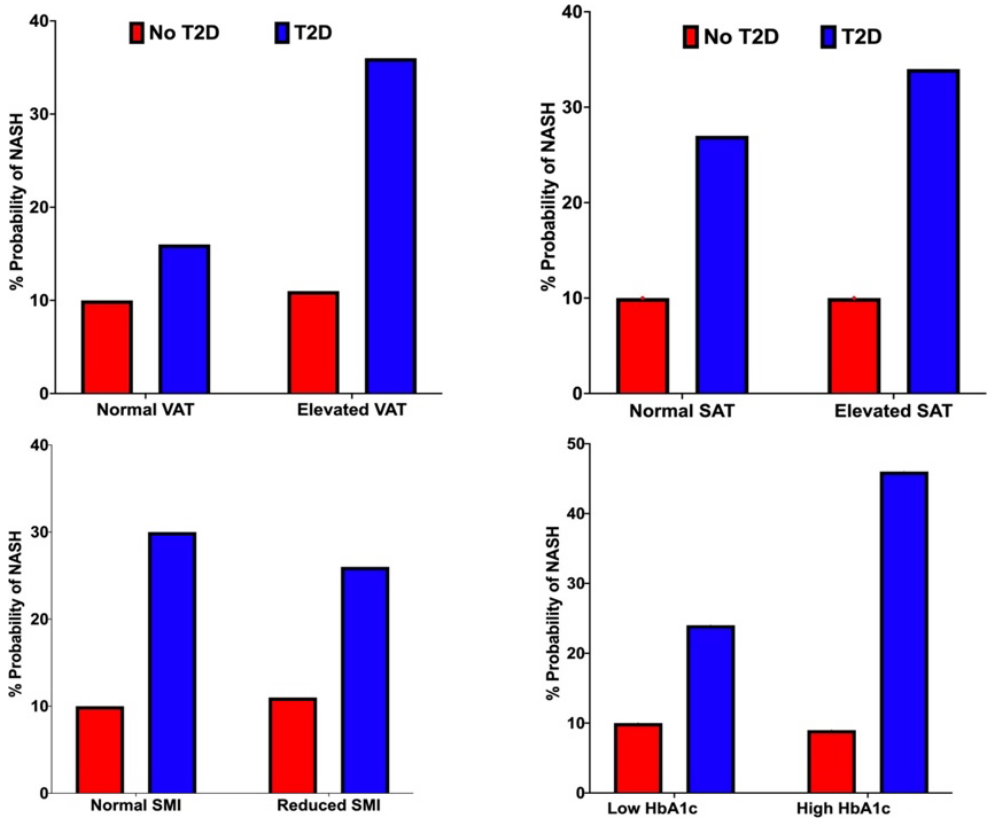


Figure 6.3. Probability of NASH (%) under VAT (top left), SAT (top right), SMI (bottom left) and HbA1c status (bottom right) conditions in the non-T2D (red) and T2D (blue) cohorts.

T2D cohort

Impact of body mass index: We first explored the effect of obesity status on NASH risk, observing a 13%, 23%, 36% and 62% probability of NASH under normal weight (<25kg/m^2), overweight (25-30kg/m^2), obesity (30-40kg/m^2), and severe obesity (>40kg/m^2) settings, respectively. This equated to a 23% greater probability of NASH when comparing 'obesity' *vs.* 'normal weight'. (Table 6.2 & Fig. 6.2).

Impact of HbA1c: Increasing HbA1c status from the low (HbA1c <62 mmol/mol) to high value group (HbA1c >=62mmol/mol) increased the probability of NASH by 22% (24% vs. 46%). (Table 6.2 & Fig. 6.3).

Impact of adipose tissue volumes (VAT and SAT): We observed a 20% and 7% increase in the probability of NASH when specifying elevated vs. normal measures of VAT and SAT, respectively. (Table 6.2 & Fig. 6.3).

Impact of skeletal muscle area: Reduced SMI decreased the probability of NASH by 4% (30% vs 26%). (Table 6.2 & Fig 6.3).

Impact of liver and pancreatic fat: Increasing liver fat from 5.6-10% to greater than 10% increased the probability of NASH by 31% (25% vs 56%). Elevated pancreatic fat increased the probability of NASH by 1% (29% vs 30%). (Table 6.2 & Fig. 6.3).

Combination of risk factors: BNs enable a user (e.g., clinician) to specify an individual phenotype, enabling personalised 'what if' analysis. For example, a high-risk phenotype with obesity, elevated VAT, elevated liver fat (>10%) and 'high' HbA1c had an 86% probability of NASH. In this phenotype, reducing HbA1c to the 'low' value reduced the probability of NASH by 29% (67%).

Cohort without T2D

We observed a 16%, 6%, 11% and 30% probability of NASH under normal weight, overweight, obese, and severe obesity settings. We observed a 1% increase in the probability of NASH when specifying elevated vs normal measures of VAT. Increasing liver fat from 5.6-10% to >10% increased the probability of NASH by 23% (7% vs. 30%). Increasing HbA1c from the 'low' <38mmol/mol) to 'high' (>38mmol/mol) value reduced the probability of NASH by 1% (10% vs 9%). (Table 6.2 & Figs. 6.2 and 6.3).

Exploring the same high-risk phenotype specified within the T2D cohort, we observed a 30% probability of NASH. Reducing HbA1c lowered the probability of NASH by 3% (27%).

Table 6.2. Probability of NASH (%) under different variable settings in the T2D and non-T2D cohorts.

Biomarker	Cohort	
	T2D (n = 221)	**Non-T2D (n = 135)**
Obesity status (BMI)		
Normal weight (<25kg/m^2)	13%	16%
Overweight (25-30kg/m^2)	23%	6%
Obese (30-40kg/m^2)	36%	11%
Severe obesity (>40kg/m^2)	62%	30%
Body composition & fat deposition		
Normal VAT	16%	10%
Elevated VAT	36%	11%
Normal SAT	27%	10%
Elevated SAT	34%	10%
Normal SMI	30%	10%
Reduced SMI	26%	11%
Liver fat 5.6-10%	25%	7%
Liver fat >10%	56%	30%
Normal pancreatic fat	29%	10%
Elevated pancreatic fat	30%	10%
HbA1c range		
Low (<62mmol/mol [T2D]; <38mmol/mol [non-T2D])	86%	30%
High (>=62mmol/mol [MODIFY]; >=38mml/mol [non-T2D])	67%	27%

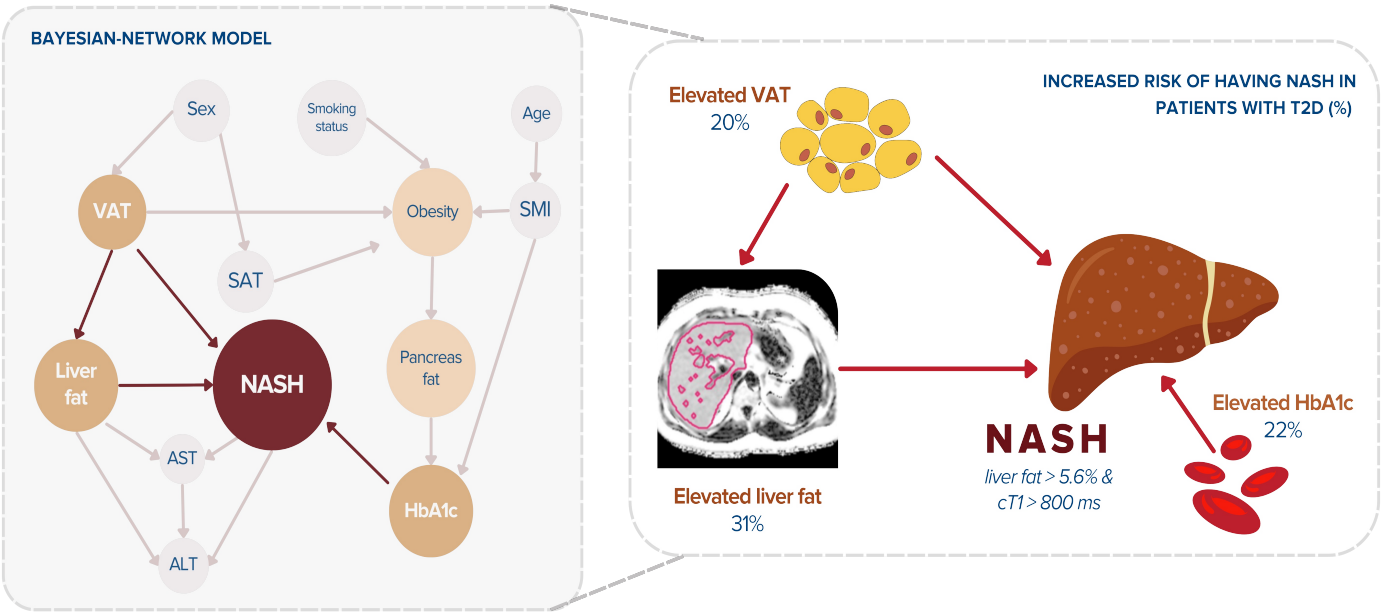


Figure 6.4. Illustration of main project findings.

6.4. Discussion

In this Chapter, we applied a BN approach to demonstrate the mediating influence of T2D, particularly determining the impact of poor glycaemic control and higher ectopic fat deposition. We focused on the impact of liver and visceral fat depots on the increased risk of NASH in people with T2D and those in the general population. Firstly, we found that obesity *(versus* normal BMI) conferred a 23% increase (within the T2D cohort) and 5% decrease (in the cohort without T2D) in the probability of NASH. However, this Thesis is concerned with moving beyond the limitations of BMI, such as the assumption that all individuals will have a similar relative proportion of the different biological tissues.

This Chapter overcomes this limitation by applying BNs to study the association between MR-derived measures of body composition (using regional adipose tissue volumes, VAT and SAT

and organ specific fat measurements) and NASH risk, highlighting the independent contribution of fat deposition within visceral and subcutaneous sites. Specifically, within the T2D cohort the probability increase of NASH was more than double under elevated VAT (+20%) than elevated SAT (+7%) conditions. Such a finding is similar to that of [134] who, despite applying different statistical techniques, also observed that higher measures of VAT relative to SAT was predictive of advanced liver fibrosis in people with NAFLD. VAT area is independently associated with NASH and correlates significantly with histology confirmed NAFLD with significant fibrosis [135]. Proposed mechanisms behind the association of elevated VAT and NASH include lipotoxicity and an overexpression of proinflammatory cytokines that promote inflammation and fibrosis within the hepatocytes [136]. Such overexpression of cytokines has been linked to fat deposition within visceral, but not subcutaneous, adipose sites [137].

In participants with T2D, our analysis revealed a 22% reduction in the probability of NASH when lowering HbA1c from the 'high' (>62mmol/mol) to 'low' (<62mmol/mol) range. Furthermore, NASH risk was reduced by 29% when lowering HbA1c in an example high risk phenotype. Importantly, this was achieved without specifying changes to obesity status, liver fat content or body composition. Chronically elevated blood glucose (i.e., glucotoxicity) has been linked to NASH development via its effects on increasing TCA cycle activity and synthesis of Acyl CoA that promote de novo lipogenesis and oxidative stress [138]. At the time of writing, no FDA-approved pharmacological medications are available for the treatment of NASH. To this end, this Chapter highlights the potential of glucose lowering therapies for mitigating NASH risk, particularly in high-risk metabolically unhealthy populations or those with overweight/obesity. Interestingly, newer T2D therapies such as GLP1-receptor agonists

(liraglutide, semaglutide), novel dual and triple peptides (e.g., tirzepatide) and SGLT2-inhibitors have been shown to decrease levels of HbA1c, ALT and lower liver fat (measured by MRI-PDFF) in people with T2D and areas of fibrosis, ALT/AST, and hepatic lipid content in murine NASH models [139].

Notably, and despite similar measures of BMI and body composition (see Table 1), we found significantly greater measures of liver cT1, liver fat and prevalence of NASH within the T2D cohort. This highlights the clinical need to screen people with T2D for concomitant liver disease and NASH, where critically, we observed similar measures of ALT and significantly greater AST in the non-T2D cohort. Screening for liver disease needs to adopt a multi-modality approach that extends beyond the reliance on circulating biomarkers alone.

6.5. Conclusion

In conclusion, this Chapter applied Bayesian-networks and probabilistic reasoning to identify populations at a high risk of NASH. We emphasise the role of elevated VAT, liver fat and obesity status in driving the probability of NASH, and these effects are most significant in people with T2D highlighting the importance of prevention and good glycaemic management as a potential therapeutic target for addressing the epidemic of NAFLD/NASH. Limitations of these inferences ultimately derive from the fact that NASH is a complex and multifactorial disease that involves numerous mechanisms that are not investigated here. However, the BN methodology used in this Chapter is intrinsically extensible and can be expanded to include additional variables and explore different disease indications, as the following Chapter illustrates.

In 2019, metabolic-related diseases and obesity were placed firmly in the spotlight with the emergence of the Coronavirus (COVID-19) pandemic. People with obesity and T2D were significantly more likely to experience severe COVID-19 related outcomes such as hospitalisation and use of mechanical ventilators compared to those without. However, what is it specifically about *obesity* which is driving such association? In order to address such questions and understand the driving mechanisms behind COVID-19 related hospitalisation, in the next Chapter I demonstrate a different application of Bayesian-networks for modelling multi-organ pathologies, this time exploring the relationships between fat deposition and hospitalisation risk following acute COVID-19.

7

APPLYING BAYESIAN-NETWORKS TO STUDY HOSPITALISATION RISK FOLLOWING ACUTE COVID-19

7.1. Introduction

In 2020, the SARS-CoV-2 virus spread rapidly across the globe. COVID-19 is caused by the SARS-CoV-2 virus and can lead to acute complications and chronic ill health, called Long-Covid. As of 6th July 2021, over 4.9 million people in the United Kingdom have been infected with SARS-CoV-2, resulting in 473,039 hospital admissions and 128,231 COVID-related deaths [140]. Hospitals and primary care facilities have been pushed to breaking point and the full extent of the public health impact will likely be felt for many years to come.

Patients suffering from COVID-19 present with a heterogenous display of symptoms including fatigue, respiratory failure, and neurological complications. An estimated 50% of hospitalised patients go on to experience Long-COVID, however Long-COVID sufferers are themselves an heterogenous population, also consisting of those not hospitalised during the acute stages of infection [141]. Several risk factors for adverse COVID-related outcomes have emerged, including elderly age, obesity, male sex and pre-existing NAFLD, all reporting to significantly increase the risk of hospitalisation following COVID-19 [142]. Indeed, the Department of Health for England announced its new strategy for obesity which includes a particular

emphasis on the risks associated with COVID-19, noting that 'excess weight is one of the few modifiable factors for COVID-19' [143].

However, though obesity has been repeatedly highlighted as a significant risk factor for adverse outcomes following COVID-19, this Thesis is concerned with addressing the limitations of the BMI through exploring fat deposition and the unique phenotypes at high-risk of metabolic-associated diseases. For example, independent of BMI, an elevated visceral adipose tissue (VAT)/ subcutaneous adipose tissue (SAT) ratio is a significant predictor of poorer COVID-19 prognosis [144], [145].

Evidently, both T2D and COVID-19 share a common clinical dilemma: both are associated with multi-organ dysfunction and affect individuals differently. There is also a common overlap in risk factors, namely obesity, which significantly increases both the risk of COVID-19 hospitalisation and T2D onset. Furthermore, COVID-19 outcomes in those with pre-existing T2D has received considerable attention. T2D is amongst the major comorbidities for adverse COVID-19 related outcomes, leading to an increased risk of mortality, hospital stay and admission to intensive care wards [146]. Postulated causal links between T2D and adverse COVID-19 outcomes include an overexpression of Angiotensin-converting enzyme 2 in the lungs, the receptor associated with COVID-19 viral entry, and pre-existing chronic inflammation resulting in an extreme systemic immune response (or 'cytokine storm') that is strongly associated with severity of COVID-19 [147].

Given the similarity in clinical presentation and healthcare challenges between T2D and COVID- 19, this Chapter applies the same probabilistic Bayesian-network approach to study the association between fat deposition and hospitalisation risk in participants with acute

COVID-19. While BNs have been applied to model COVID-19 risk with great success [148]–[150], they have yet to be applied for studying body fat distribution and risk of COVID-19 hospitalisation. The aim of this Chapter was to unravel the association between 'obesity' and hospitalisation risk following COVID-19 by applying Bayesian-networks to model fat deposition and body composition in a cohort of participants from the COVERSCAN study with confirmed COVID-19.

We describe how MR-derived-imaging phenotypes, assessing volumes of different fat depots including liver fat and body composition, combined with BN modelling to capture *conditional* dependencies, enables phenotypic characterisation of the highest risk of metabolic phenotypes to be hospitalised following COVID-19. We also compare the predictive performance of BNs against traditional classification algorithms. This project was published in 2023 within the journal *Frontiers in Bioinformatics*.

Paper citation

Waddell, T., Namburete, A.I.L., Duckworth, P., Eichert, N., Thomaides-Brears, H., Cuthbertson, D.J., Despres, J.P. and Brady, M., 2023. Bayesian networks and imaging-derived phenotypes highlight the role of fat deposition in COVID-19 hospitalisation risk. *Frontiers in Bioinformatics*, *3*, p.1163430.

7.2. Project methodology

7.2.1. Data collection & preparation

Anthropometric, demographic, and imaging data were extracted from the COVERSCAN study (NCT04369807) that recruited participants with confirmed COVID-19 between April 2020-October 2021. Participants with laboratory confirmed SARS-CoV-2 infection (as tested by

nasopharyngeal or oropharyngeal swab), positive antibody test or a strong clinical indication of SARS-CoV-2 as confirmed by two independent clinicians were eligible for recruitment. Exclusion criteria included either three or more episodes of coughing in the past 24 hours, symptoms of an active respiratory infection or any contraindications to MRI such as metallic implanted devices and claustrophobia. All participants underwent an abdominal MRI assessment at two UK research imaging sites (Perspectum Ltd, Oxford or Mayo Clinic Healthcare, London) that included acquiring MR-images of the liver and body composition. In total, data from 466 participants had the necessary imaging data available at the time of analysis and were collected, then after removing missing data entries (for example, missing liver PDFF values), 459 remained (395 non-hospitalised, 64 hospitalised). It is important, therefore, to highlight the low sample size as a limitation of this analysis. Specifically, low sample sizes increase the risk of an inadequately powered analysis and risk of type -2 error, an increased likelihood of unbalanced groups, and an increased biased when discretising the groups for classification and inferential statistics. The results in this analysis, therefore, should be viewed with caution and more as an example as to how Bayesian-networks can be applied. See Appendix for a full cohort flow diagram.

7.2.2. Bayesian-network construction

The following variables were selected for constructing the BN: smoking status (never smoked, current smoker, past smoker]), hospitalisation status (1[hospitalised]/0[non-hospitalised]), liver fat (proton density fat fraction [PDFF] %), visceral adipose tissue (cm^2), subcutaneous adipose tissue (SAT) (cm^2), skeletal muscle (cm^2), sex (male/female), BMI (kg/m^2) and age (yrs). Similar to previous work, body composition was examined from a 2D MR slice positioned

at the third lumbar (L3) vertebra and measures of VAT, SAT and skeletal muscle were based on manual delineations by trained analysts. Measures of skeletal muscle were then indexed to the participants height to produce a measure of skeletal muscle index (SMI) (cm^2/m^2). All continuous variables were discretised based on pre-defined clinical thresholds, for example BMI was discretised into obesity categories: normal weight (BMI <25kg/m^2), overweight (BMI 25-30kg/m^2) and obese (BMI >30kg/m^2). See Appendix for a full overview of discretisation thresholds.

As with the previous Chapters, a semantic network was constructed in collaboration with Prof Cuthbertson and Prof Despres which was then digitised using the 'bnlearn' package and visualised using 'graphviz' within the R software platform (version 3.6.1). The MMHC automated structure learning algorithm provided the initial network skeleton. This network was then adjusted by removing or reversing nonsensical edges and inserting edges in collaboration with Prof Cuthbertson and Prof Despres based on the established, semantic network, see Fig. 7.1.

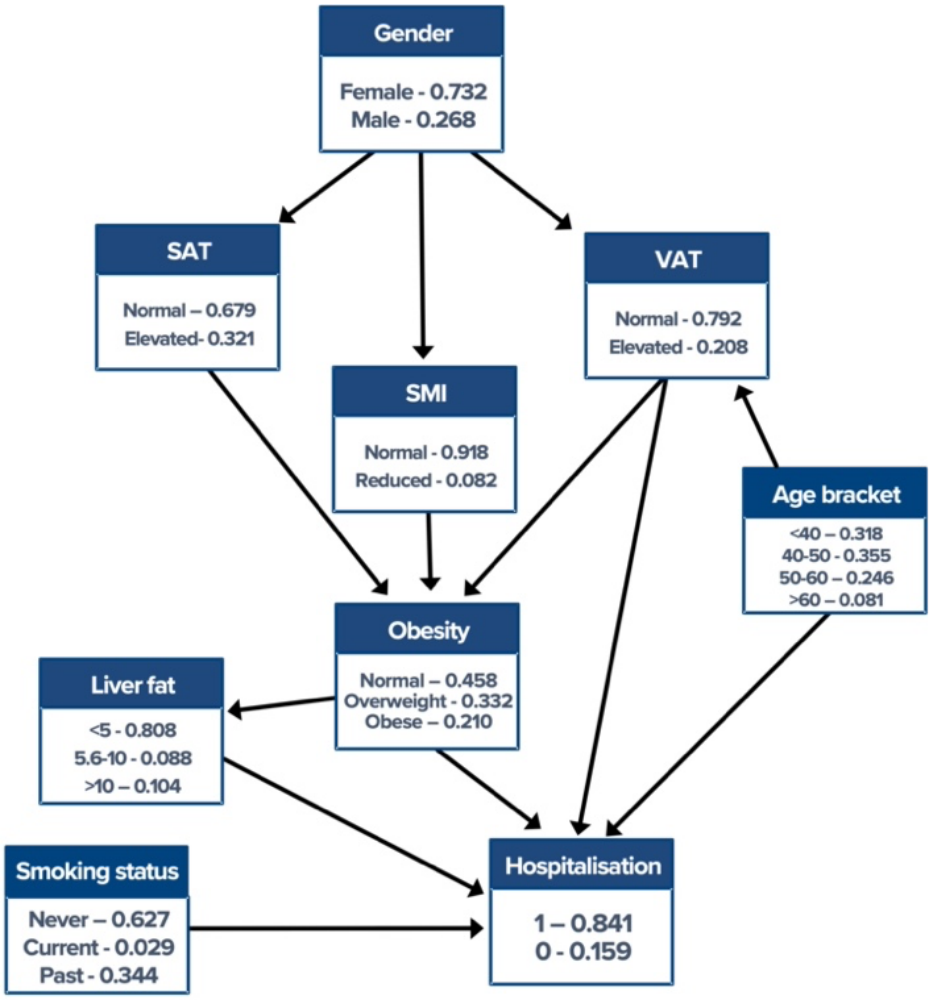


Figure 7.1. Final semantic Bayesian-network structure.

7.2.3. Probabilistic inference

In this Chapter we estimated the probability of hospitalisation, given an 'evidence list' containing the fixed values of network variables such as liver fat or VAT. For example, the probability of hospitalisation was estimated after fixing the values of 'Liver fat' to normal liver fat (<5%), mild steatosis (5-10%) and severe steatosis (>10%), allowing the direct effect of elevated liver fat on hospitalisation risk to be inferred. Conditional probability queries were

performed using the 'cpquery' function in bnlearn and estimated using the likelihood weighting algorithm.

7.2.4. Prediction of hospitalisation status

The following Statistical and Machine Learning (ML) classification algorithms were used as a comparison for predicting hospitalisation status: logistic regression, Naïve Bayes, and Decision Tree. These models were selected for they are commonly applied to classification problems and are representative of the gold-standard techniques for classifying clinical outcomes, such as acute hospitalisation. These were implemented within R using the 'brglm2', 'klaR' and 'tree' packages, respectively. The dataset was split into a train (internal) (n=46 [6 hospitalised, 40 non-hospitalised]) and test (n=413 [58 hospitalised, 355 non-hospitalised]) validation cohorts. Area under the receiver operating characteristic curves (AUC) are reported to measure model performance.

7.2.5. Statistical analysis

All statistical analyses used the R software platform (version 3.6.1). Descriptive statistics, showing median [inter quartile range], are reported to summarise population characteristics. Adopting a significance threshold of $p<0.05$, Wilcoxon and X^2 tests revealed hospitalised participants were significantly older ($p<0.001$) and had significantly elevated measures of BMI ($p=0.042$), liver fat ($p=0.0028$) and VAT ($p<0.001$). See Table 7.1.

7.3. Results

Population characteristics

Table 7.1. Population characteristics between hospitalised and non-hospitalised participants. Data is median [IQR], significant p-values in bold.

Participant variable	Hospitalised (n =64)	Non-hospitalised (n = 395)	p-value
Age (yrs)	50 [43-57]	43 [37-51]	**0.00035**
Sex (% M)	35%	33%	0.8
BMI (kg/m^2)	27 [23-32]	25 [22-28]	**0.042**
Liver fat (% PDFF)	3 [2.2-6.7]	2.4 [1.9-3.9]	**0.0028**
VAT (cm^2)	120 [66-203]	85 [47-131]	**0.00077**
SAT (cm^2)	237 [141-315]	205 [136-305]	0.19
SMI (cm^2/m^2)	45 [39-49]	43 [38-49]	0.31

Model prediction

ROC curves indicating model performance are shown in Fig. 7.2. Overall, performance on the validation cohort was highest within the BN model (AUC [95% CI]) (0.84 [0.79-0.89]), followed by logistic regression (0.68 [0.60-0.77]), Naïve Bayes (0.66 [0.58-0.74) and Decision Tree (0.57 [0.51-0.64]).

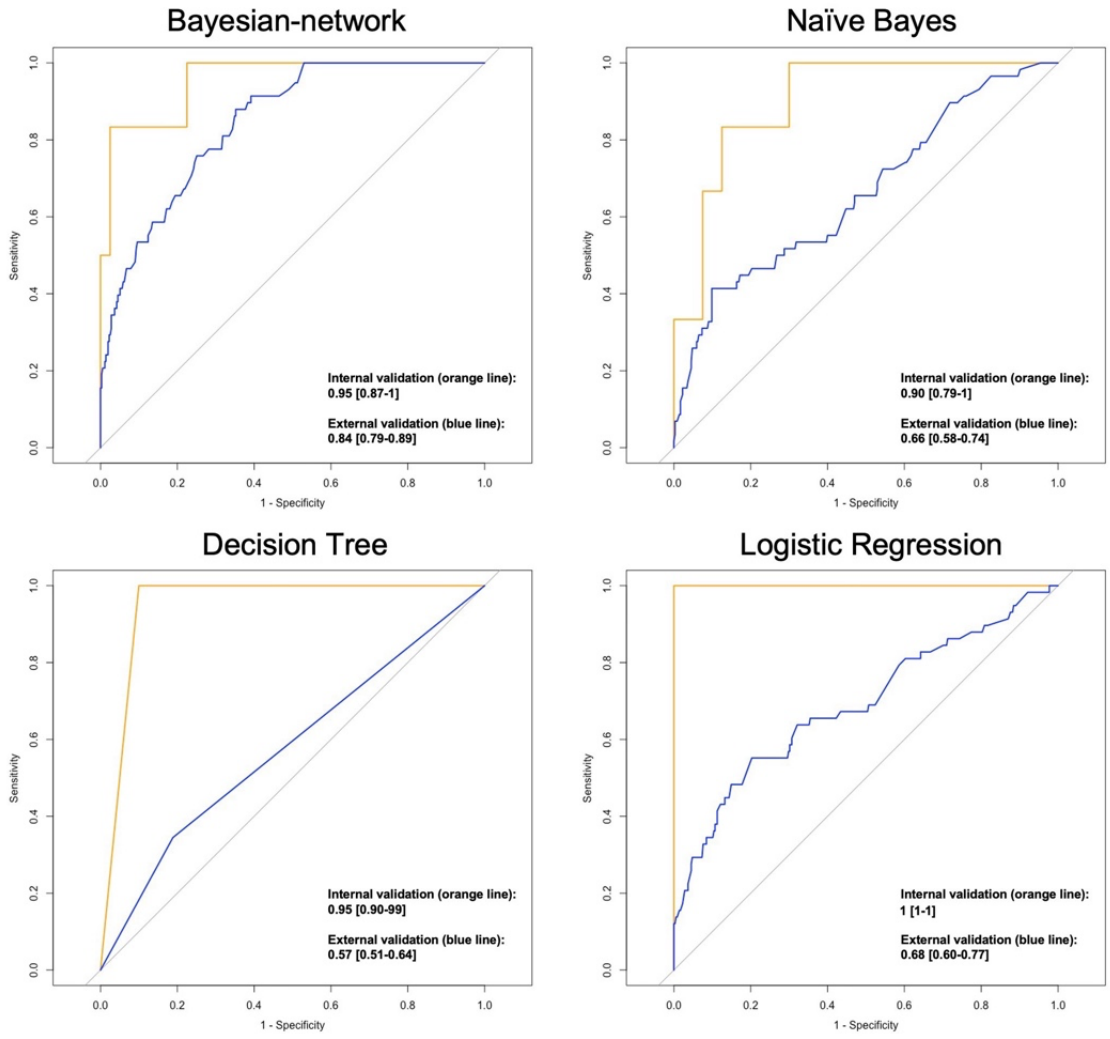


Figure 7.2. ROC plots indicating the performance of classification models in predicting the hospitalisation status AUC [95% CI]. Bayesian-network: internal (training) validation (orange line) 0.95 [0.87-1], external validation (blue line) 0.84 [0.79-0.89]; Naïve Bayes: internal validation (orange line) 0.90 [0.79-1], external validation (blue line) 0.66 [0.58-0.74]; decision tree: internal validation (orange line) 0.95 [0.90-0.99], external validation (blue line) 0.57 [0.51-0.64]; logistic regression: internal validation 1 [1-1], external validation (blue line) 0.68 [0.60-0.77].

Estimation of hospitalisation using probabilistic reasoning

The baseline probability of hospitalisation from acute COVID-19 was 15%. We first intervened on the variable 'Obesity' only, observing a 11%, 13% and 29% probability of hospitalisation when the value was set respectively to normal weight, overweight and obese. This equates to a 18% greater probability of hospitalisation when 'Obesity' was set to obese compared to normal weight.

Table 7.2. Probability of hospitalisation (%) and probability change (+/- %) across BMI-categories.

	Liver fat <5% & non-elevated VAT (baseline)	**Liver fat 5-10%**	**Liver fat >10%**	**Elevated VAT**	**Elevated VAT & Liver fat >10%**
Normal weight	10%	26% (+16%)	39% (+29%)	27% (+17%)	45% (+35%)
Overweight	8%	2% (-6%)	9% (+1%)	31% (+23%)	41% (+33%)
Obesity	14%	55% (+41%)	35 (+21%)	25% (+11%)	31% (+17%)

To ascertain what component of body composition was driving the elevated obesity risk, we examined the probability of living with obesity given varying measures of body composition. We observed that elevated VAT doubled the probability of living with obesity compared to elevated SAT (50% vs 24%), identifying elevated VAT as the primary determinant of obesity. We then intervened on the variables 'Liver fat' and 'VAT' without fixing the value of 'Obesity'. Elevated VAT resulted in a 31% probability of hospitalisation, this was reduced by 20% when specifying non-elevated VAT, which elicited an 11% probability of hospitalisation. We

observed a 'Liver fat' value of <5% (normal), 5-10% (mild steatosis) and >10% (severe steatosis) resulted respectively in a 12%, 25% and 30% probability of hospitalisation.

We then examined the influence of VAT and Liver fat on hospitalisation risk across each BMI-based category. We first estimated the probability of hospitalisation with normal measures of liver fat (<5%) and VAT, denoted here as our 'baseline' measurement. Next, the value of the variables 'VAT' and 'Liver fat' was individually fixed to ascertain the direct effect of each variable state on hospitalisation risk, see Table 7.2.

7.4. Discussion

Obesity is a significant risk factor of hospitalisation following acute COVID-19; however, use of the BMI as a measure of body fat distribution is intrinsically limited. In this Chapter we illustrate how BN modelling and probabilistic reasoning enables the associations between 'obesity' and hospitalisation risk to be unravelled. However, it is important to highlight the unbalanced groupings in this analysis, notably Sex, whereby there are considerably fewer males in the cohort. Such unbalancing may lead to inaccuracies when performing statistical inferences, notably probabilistic inference, for the model will be biased and drawing on far fewer males when performing random sampling. Furthermore, it is important to again highlight the low sample size as a limitation of this analysis, leading to an increased risk of low power and risk of type -2 error and biased when discretising the groups for classification and inferential statistics. The results in this analysis, therefore, should be viewed with caution and instead seen as an example of how Bayesian-networks can be applied to model biological

systems, as opposed to definitive answers regarding the endotype of diseases. Future work could apply a larger and more comprehensive dataset to this network to reduce the limitations of imbalanced groups and low sample size, in order to derive inferences that could be used with greater confidence in guiding public health decisions.

We first demonstrate an 18% greater probability of hospitalisation in people with obesity *vs.* being of normal weight, with the primary determinant of obesity being elevated VAT. We postulate that the association between obesity and hospitalisation following COVID-19 is primarily driven by VAT deposition rather than SAT. This is consistent with [145], who reported a higher VAT/SAT ratio being significantly predictive of adverse outcomes following COVID-19, independent of BMI status.

Except for the case when 'Obesity' was set to overweight, we demonstrate liver fat measures of 5-10% and >10% consistently increased the probability of hospitalisation by an average of 29% and 25%, respectively. It is important to note that the lack of change within the 'overweight' setting could be an artefact arising from the small dataset and poor sub-group balancing in this analysis. Such results should therefore be seen as a potential artefact and not a novel finding within this cohort. Similar work on data from the UK Biobank by Roca-Fernandez et al. [142] reported a fivefold greater risk of hospitalisation in people with obesity with concurrent NAFLD *vs* those with obesity but without NAFLD. Furthermore, across all obesity categories both elevated VAT and liver fat together caused an average increase in the probability of hospitalisation of 39%. Taken together, these results emphasise the role of ectopic fat deposition in driving hospitalisation risk which is not captured through use of the BMI alone. Postulated mechanisms between elevated VAT and poorer COVID-19 prognosis

include an overexpression of proinflammatory cytokines and increased lipolysis leading to epithelial injury [151]. Furthermore, it is thought that the SARS-CoV-2 virus has a direct, damaging effect on the liver hepatocytes, with viral entry via the angiotensin-converting enzyme 2 (ACE2) receptors resulting in mitochondrial dysfunction, hypoxia, and liver injury [152]. Consequently, not only are those with pre-existing liver disease at a greater risk of adverse COVID-19 related outcomes, but COVID-19 itself is thought to induce liver injury in otherwise healthy persons.

As illustrated in the previous chapters, probabilistic inference can be used to estimate patient outcomes after fixing the values of specific network variables, providing an insight into potential treatment outcomes. For example, we show that even under normal weight settings, reducing liver fat content from >10% to less than 5% reduced the probability of hospitalisation by 29%. Therefore, in this population, therapies known to elicit favourable changes in liver health without requiring significant changes in body weight, such as the Mediterranean diet or exercise training, should be considered as a viable therapeutic option [120]. This highlights the utility of BNs for promoting *personalised* care for patients, for the care of this normal patient would be considerably different for those with overweight or obesity where weight loss is considered the primary treatment target for reducing COVID-19 related risk [153]. A >10% to less than 5% reduction in liver fat has been demonstrated in people with obesity following bariatric surgery [154], while treatment with GLP-1 receptor agonists have shown to reduce liver fat, VAT and associated metabolic markers such as hyperglycaemia [155], [156].

In this chapter we applied BN modelling and probabilistic inference to study the association between body composition, liver fat deposition and hospitalisation risk following COVID-19. We show that elevated VAT and liver fat both increase the probability of hospitalisation and illustrate how BNs can be applied to estimate acute clinical events such as hospitalisation. This chapter exemplifies how BNs can be expanded to explore different clinical indications, pivoting from exploring T2D and NASH to COVID-related outcomes, as per the previous chapters. Crucially, this Chapter demonstrates how the basic structure of the BN developed in chapters 4-6 with the inclusion of body composition and liver biomarkers can be adjusted to include additional data sources and clinical outcomes. This is key, for it highlights the fluidity of BNs and their ability to be moulded to the specific clinical need, so long as the underlying premise of *causality* is maintained, as denoted in the semantic network developed in Chapter 4. As we discuss in the 'Further work' section in the following chapter, BNs can be expanded to not only include additional data types, but also time slices, allowing the modelling of temporal changes in biomarkers.

This Chapter also highlights the importance of the industry collaboration with Perspectum Ltd throughout this DPhil project. The COVID-19 pandemic brought about many clinical challenges and, under such challenging circumstances, Perspectum expanded their imaging suite to cover the lungs and heart in order to capture the full consequences of COVID-19 on multi-organ health. In doing so, Perspectum successfully set up the world's largest imaging study into long COVID (COVERSCAN), the data from which was used in this Chapter. This industry collaboration allowed me to gain quick access to such data and the ability to work with leading authorities in COVID-19 research such as Dr Rajarshi Banerjee in order to complete such work. This highlights how collaborating across academic and industry sectors

can quickly advance medical research and awareness of public health challenges such as the COVID-19 pandemic. In the following Chapter, we review and conclude the contents of this Thesis. We draw upon the main contributions of this DPhil project and detail two notable, promising areas of further work on applying Bayesian-networks for modelling multi-organ pathologies.

8

CONCLUSIONS & FURTHER WORK

As researchers, we now have a wealth of information at our disposal, from MR-derived measures of organ health, circulating biomarkers and anthropometrics, to simple demographics including one's age and sex. But with such richness comes the question of how best to utilise such data, and clinicians are increasingly having to balance this complexity with treatment guidelines and their patients' individual phenotype and needs. Traditional approaches for studying metabolic diseases have been driven by frequentist statistical approaches – the correlations, pairwise comparisons and significance tests that we are all too familiar with. While such techniques certainly have their place in medical research, Chapter 2 laid out how they're intrinsically limited for dealing with the complexity of clinical decision making, which is also inherently *counterfactual* in nature. An alternative approach is the emergence of AI-powered decision support algorithms, which though effective and promising as they are, have been resisted for their 'black box' nature and inability for the clinician to often ask the question, why? This Thesis proposed a third approach for studying the complexity of metabolic disease – the field of Bayesian-networks and probabilistic inference.

To this end, this Thesis specifically explored the relationships between fat deposition, hyperglycaemia, and liver disease; estimated the consequences of projected increases in

obesity prevalence and explored the association between fat deposition and hospitalisation from acute COVID-19. We applied probabilistic inference to estimate counterfactual questions typically posed by clinicians, such as: *'what is the probability that my patient reduces their risk of NASH, given that their blood glucose is controlled?'*. Such questions were explored across four main questions, though a single theme connected each project – can Bayesian-networks help to support clinicians with their increasingly complex decision making?

Though the work contained within this Thesis has been published in peer-reviewed journals and presented at international conferences, thereby contributing to the body of literature, I believe its true contribution comes from bringing together the fields of engineering and medicine. Though so called 'black box' decision-making algorithms are a powerful and positive tool in medical care, such approaches come at the expense of our ability to understand or access the knowledge imbedded within these algorithms. We are often unable to ask the question - why did I get this answer, or why is this person regarded as high-risk, or not? Disregarding an explanation for individual cases, leaders in the field are even calling the move to black box systems as "contravening the profound moral responsibilities of clinicians" [157]. Though this comment is certainly unfair to those applying such algorithms effectively, a commonly cited barrier to the adoption of AI in medicine is interpretability, or lack thereof [158].

Through working with Prof Cuthbertson, Prof Despres, and colleagues at Perspectum, this Thesis has taken a step away from the 'black box' approaches which, as quoted in the Introduction, are resisted because "physicians are reluctant to trust and adopt something

they do not fully understand" [111]. Rather, the work contained in this Thesis was completed within a multi-disciplinary team of clinicians, data scientists and industry leaders, from formulating the semantic networks embodied with clinical knowledge to forming the clinical questions in which to pose the network. Through completing this DPhil, I have come to truly believe that the future of medicine will require the continual merging of such multi-disciplinary teams in order to deliver personalised care. This is especially pertinent with the proliferation of AI-enabled systems in medicine, from the headline grabbing breakthroughs of DeepMind's 'AlphaFold' technology [159], to lung cancer nodule detection used in routine clinical care [160]. Leaving the clinician out of the loop when developing such techniques is likely to create disjoint between AI-development in academic research institutions and its application within primary and secondary care.

This Thesis highlights Bayesian-networks as a promising technique in which to overcome the barrier of interpretability in medicine, for two main reasons. Firstly, is the structure of the Bayesian-network, which in this Thesis work was explicitly encoded with the clinicians' prior expert knowledge – the *causality* of the system. This not only confers the clinician with a sense of ownership of the model, but also allows us to understand, explicitly, that the observed change in x (say, liver cT1) is being caused by a change in y (say, liver PDFF). Secondly, as noted throughout this Thesis, probabilistic inference allows us to ask questions of the network that are counterfactual in nature, mimicking those often found in the clinical care. Together, this allows clinicians to estimate patient outcomes as they would in the primary care setting, but also understand how such estimations are being propagated throughout the network.

8.1. Summary of key research findings and Thesis contributions

Obesity is recognised as a focal risk factor for T2D and non-alcoholic liver disease, however the *distribution* of fat – one's body composition – is the primary determinant of metabolic health. Those with greater VAT are more likely to develop T2D than those of a similar BMI [31], [32], and elevated VAT but not SAT is associated with a significant increase in circulating glucose [35], [36]. Indeed, the first publication of this Thesis revealed those with T2D had significantly elevated measures of liver fat, liver disease activity (as measured by cT1), and reduced measures of skeletal muscle vs BMI-matched non-T2D persons [20]. There also exists a synergistic risk of liver related complications in people with concomitant obesity and T2D, with those with NAFLD and T2D significantly more likely to develop NASH than those with NAFLD only [19], [21]. The central theme of this Thesis, therefore, was to explore the relationship between fat deposition, liver disease and metabolic biomarkers such as HbA1c.

Chapters 4 and 6 applied BNs to model MR-derived measures of body composition, fat deposition and routinely collected health biomarkers such as HbA1c to explore the unique phenotypes associated with diabetic-status hyperglycaemia (DHG) and non-alcoholic liver disease. Specifically, Chapter 4 included the initial construction of the Bayesian-network in collaboration with Prof Cuthbertson and Prof Despres, and the selection of a software package in which to digitise such structure for performing probabilistic inference. Here, we highlighted the limitations of automated structure learning algorithms and how a hybrid approach is necessary for determining the final structure. Such exercise was certainly non-trivial, for it emphasised the how a collaborative team was necessary for the success of this DPhil project.

In chapter 4 we then estimated the probability of DHG, revealing this was driven by fat deposition in multiple, as opposed to single, sites. That is, elevated VAT or SAT in isolation did not elicit discernible changes in the probability of DHG (4% and 2% increase, respectively). However, an example phenotype characterised with fat deposition within the liver and elevated SAT and VAT resulted in a marked increase of DHG (39% probability). This Chapter also highlighted the limitations of the BMI and how a phenotype likely classified as normal weight (due to specifying 'low' SAT and SMI), but with elevated VAT and liver fat, resulted in the highest probability of DHG (52%). This finding emphasises the importance of considering how the *weight* of someone is constituted and allowed us to explore the theme of metabolically healthy/unhealthy obese and non-obese phenotypes. Finally, this Chapter reinforced the strong association between concomitant NAFLD and T2D, revealing elevated liver fat as a driving factor in DHG probability.

To explore this association between T2D, elevated liver fat and non-alcoholic liver disease further, Chapter 6 modelled the effect of fat deposition and elevated blood glucose on the probability of NASH in a cohort of participants with T2D, recruited from UK primary care clinics. In this work, published in the journal *Frontiers in Endocrinology,* we illustrated the role of elevated VAT, liver fat and blood glucose on driving the probability of NASH, each causing a 20%, 31% and 22% increase, respectively. Notably, due to the conditional independence made explicit in the network structure, such increases are independent of obesity status. This Chapter highlighted how management on one's blood glucose levels could be an effective approach for reducing their risk of NASH, an approach supported by the emerging clinical trial evidence reporting favourable liver outcomes (such as reductions in liver fat or cT1) following prescription of (traditionally) T2D medications [161]–[163].

Chapter 5 changed tack to show how Bayesian-networks could be used to estimate disease prevalence at the population, as opposed to individual, level. We applied the 2040 obesity projections from Cancer Research UK to estimate the potential effect of increasing obesity on incident liver disease in a large (though nominally healthy) UKBB cohort. While the network estimated a 5%, 8% and 3% increase in the incidence of mild fatty liver disease, severe fatty liver disease, and NASH, respectively, such estimates were likely a 'best case scenario' and underestimation of the true burden of disease. Furthermore, using Bayesian-networks we were able to estimate the potential economic burden of such increases in liver disease and highlighted the disproportionate impact obesity will have in areas of low socioeconomic status.

Chapter 7 moved away from the study of T2D and addressed the well-documented association between obesity and adverse COVID-19 related outcomes. As this Thesis attests, our measure of 'obesity' is intrinsically limited, and therefore this Chapter illustrated how the existing Bayesian-network could be pivoted to study the effect of fat deposition on the probability of hospitalisation. This Chapter revealed that although moving from 'normal weight' to 'obesity' caused an 18% increase in the probability of hospitalisation, this was primarily driven by elevated fat deposition within VAT, and not SAT, sites. Furthermore, across all obesity classifications, elevated liver fat (>10%) caused a 39% average increase in the probability of hospitalisation, highlighting the role of ectopic fat deposition as an essential consideration in health risk profiling, independent of obesity. Similar to Chapter 4, Chapter 7 highlighted how Bayesian-networks can reveal potential novel treatment targets, specifically for those considered 'normal weight', where the typical advice of 'weight loss' may not

suffice. That is, we revealed that reducing liver fat can drastically reduce the probability of both DG and COVID-19 related hospitalisation, and therefore treatments known to elicit reductions in liver fat without requiring significant weight loss (such as the MedD) should be recommended for these persons.

At the beginning of this Chapter, we stated that a single theme connected each question addressed within this Thesis – can Bayesian-networks help to support clinicians with their increasingly complex decision making? Together, the work here illustrates that Bayesian-networks are a capable tool for doing so. We have used Bayesian-networks to pose questions typically asked by clinicians, such as *'what is the probability that my patient reduces their risk of NASH, given that their blood glucose is controlled?*', which in Chapter 6, was estimated at a 22%. Of course, with any DPhil project, the work contained within this Thesis is not a completion of what can be achieved by applying Bayesian-networks to study metabolic disease. What follows in this Chapter are two promising further areas of work, highlighted during the study of my DPhil, presented through the lens of if this work were to be picked up by two additional students.

8.2 Area of further work 1 – Dynamic Bayesian-networks

The first area of further work relates to dynamic Bayesian-networks (DBNs) – the modelling of data across multiple 'time slices'. Modelling sequential data is beneficial in many areas of science, but especially in the context of clinical decision making whereby patient biomarkers can dynamically change across study visits. Afterall, it is unlikely that a patient with NAFLD or

T2D will only ever be examined once and rather, will require sequential visits in order to manage their condition correctly. The work in this Thesis was completed using single time point (observational) datasets only, which though they can be applied to study phenotypes and counterfactual outcomes, do not encapsulate such sequential patient outcomes.

Not only can Bayesian-networks be expanded to include different sources of data, but also multiple 'time slices' that each correspond to a separate network at a given time interval, see Fig. 8.2. Such time intervals could include study visits before and after a given intervention, for example following enrolment into a pharmacological trial, or simply collected during population screening trials, such as the UK Biobank. The ability to model temporal changes in biomarkers would certainly strengthen the application of Bayesian-networks, specifically by providing a means of validating the estimations inferred from the observational data. For example, [164] applied DBNs to predict the health status and treatment effect in patients with chronic lymphocytic leukaemia, reporting their network was able to predict the probability of survival in line with true survival data collected from large electronic health registries.

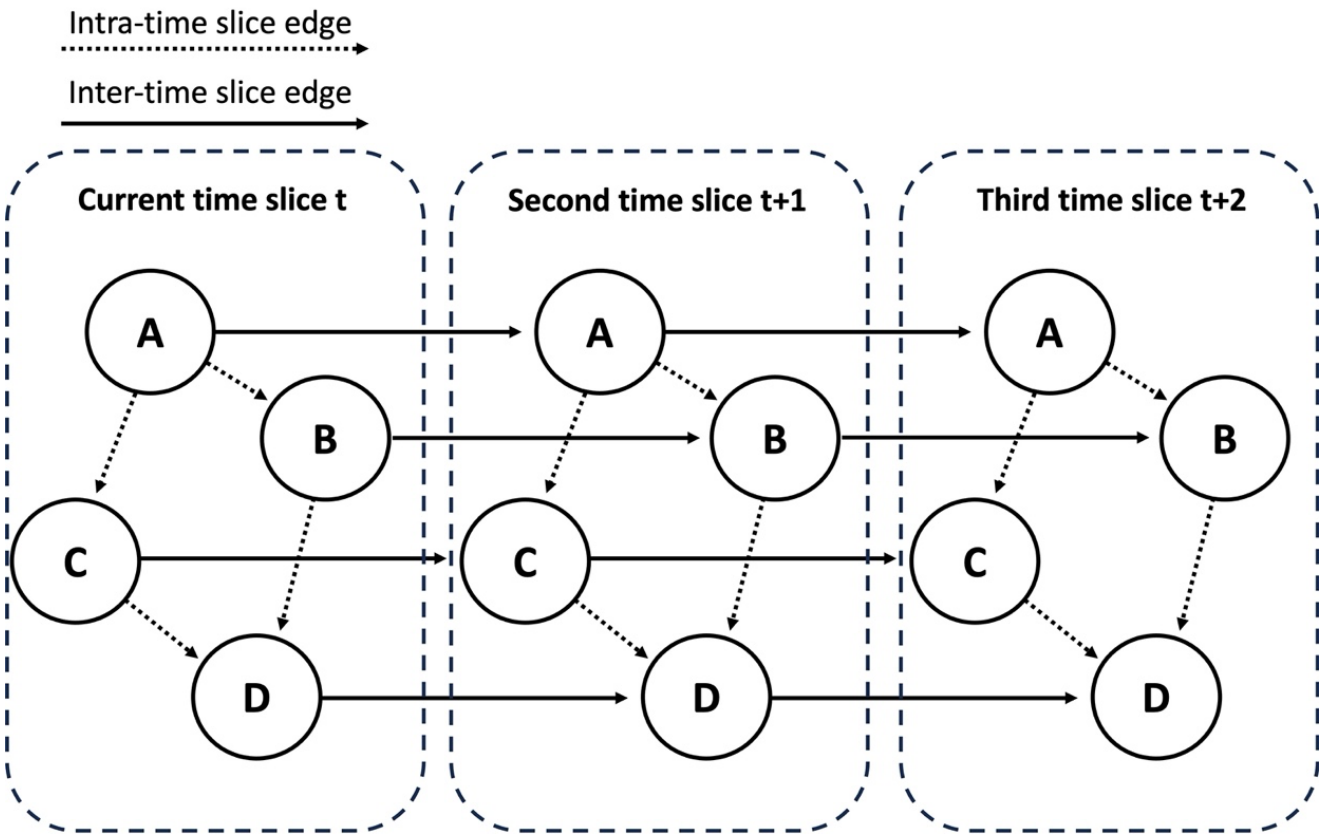


Figure 8.2. Example dynamic Bayesian-network with three time slices. Dashed and solid arrows represent intra and inter time slice edges, respectively.

To explore the utility of DBNs for modelling imaging biomarkers and metabolic health outcomes, we conducted a short project in a cohort of participants with NAFLD. Data were collected retrospectively from a randomised controlled trial in which 22 participants were randomised to receive either a GLP-1RA therapy (n = 11) or a commercial weight-loss programme (n = 11) (Lifestyle treatment).

8.2.1. Participant data

Participants underwent three study visits; firstly, at baseline (visit 1), 12-weeks (visit 2), and 24-weeks (visit 3); basic anthropometrics and abdominal MRI were collected at each study visit. Participants received either GLP-1RA or lifestyle treatment between study visit 1 and 2 and underwent a 'washout' period until visit 3. In the current analysis, measures of sex (male/female), age (years), weight (kg), BMI (kg/m^2), liver PDFF (%) and liver cT1 (ms) were collected for construction of the Bayesian-network and are labelled as the biomarker followed by the study visit, for example BMI [1] denotes BMI at study visit 1. Descriptive statistics, reporting mean ± standard deviation, were calculated to measure changes across time points between the study cohorts, see Table 8.1.

8.2.2. Bayesian-network construction

BN construction was conducted using the bnlearn package within the R software platform (version 3.6.1) and visualized using the 'graphviz' package. The network structure in this project, and the inferences made upon it, was manually constructed by inserting and orienting edges in collaboration with medical experts. For example, treatment with GLP1-RAs is known to elicit significant reductions of liver fat, and therefore an edge from 'Treatment' to 'Liver PDFF' was inserted. To retain the temporal sequence of study visits, an edge was inserted between each biomarker at each study visit, so called 'inter time slices'. For example, an edge was inserted between BMI [1], BMI [2] and BMI [3] to maintain temporal dependence, but also between BMI [1] and Liver PDFF [1] to explicitly model the relationship between obesity status and liver fat deposition. Of course, it could be argued that this edge could be reversed, where the value of BMI is conditionally dependent on the value of liver

PDFF, or even a third, unknown variable (confounder variable). This highlights another limitation of Bayesian-networks, whereby there may naturally be directional relationships or very tight correlations between variables within the given system. Such interactions, however, cannot be modelled here due to the directed and acyclic nature of Bayesian-networks. To test whether this is a significant factor in the application of Bayesian-networks to biological systems, future work could include an analysis to see whether flipping the direction of certain edges causes a significant change in the results obtained from such network. See Fig. 8.3 for the final Bayesian-network structure. Conditional probability queries were estimated using the likelihood weighting technique, a Monte-Carlo approximation technique that utilizes importance sampling.

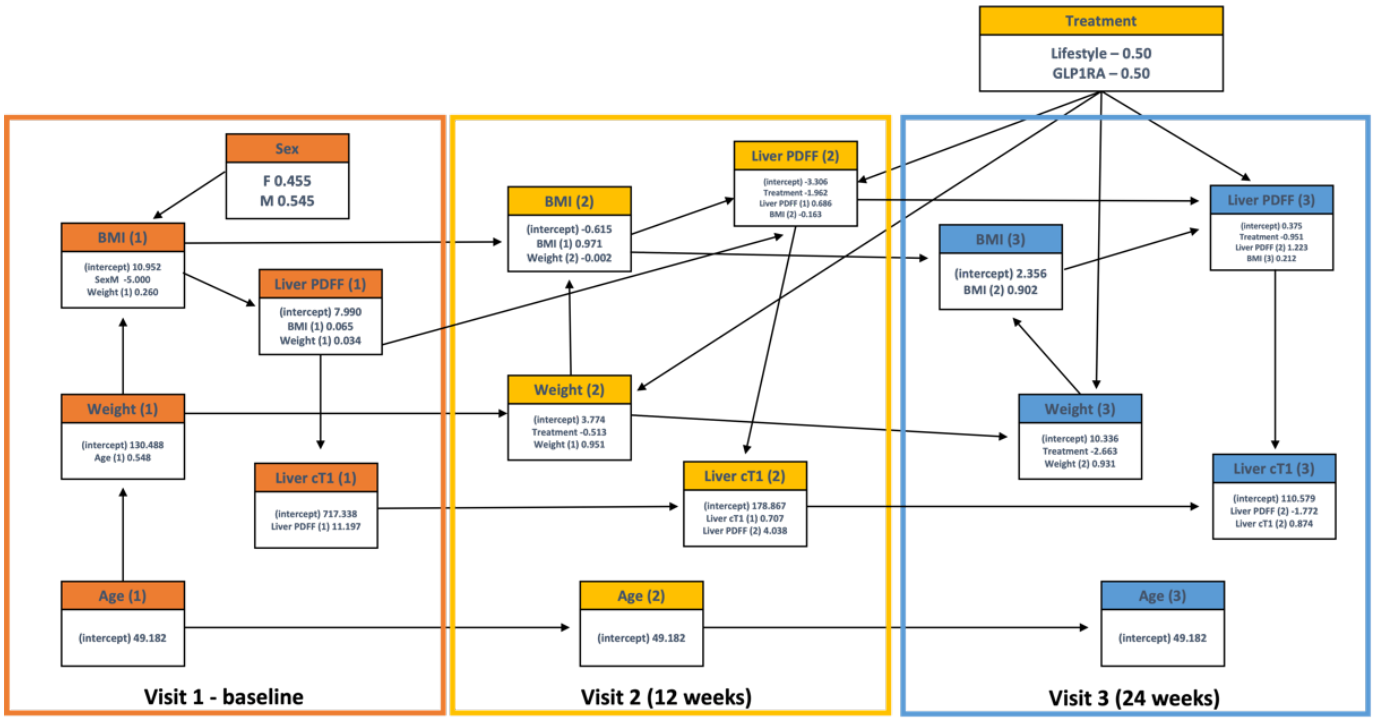


Figure 8.3. Dynamic Bayesian-network structure.

8.2.3. Results

Table 8.1. Descriptive statistics of population. Data is mean ± standard deviation.

Biomarker	**Baseline**			**12 weeks**			**24 weeks**		
	Lifestyle	GLP1-RA	p value	Lifestyle	GLP1-RA	p value	Lifestyle	GLP1-RA	p value
Weight (kg)	105.7 ± 6.6	103.6 ± 4.6	0.99	100.4 ± 6.2	97.9 ± 4.4	0.98	100 ± 5.9	99.3 ± 4.1	1.00
BMI (kg/m^2)	36.2 ± 1.7	34.9 ± 1.7	0.91	34.5 ± 1.6	33.0 ± 1.6	0.88	34.4 ± 1.5	33.4 ± 1.5	0.97
Liver cT1 (ms)	876.3 ± 23.3	855.1 ± 32.6	0.93	819.1 ± 27.1	815.2 ± 28.6	1.00	821.7 ± 29.3	840.6 ± 26.9	0.95
Liver fat (%)	13.2 ± 1.6	13.7 ± 1.9	1.00	7.1 ± 1.5	9.6 ± 1.5	0.66	8.1 ± 1.6	11.6 ± 2.0	0.38

Probabilistic reasoning

We first explored the probability of obesity (BMI >30kg/m^2) at visit 3 (24 weeks) between the two treatment arms. Intervening on the variable 'Treatment' to fix the value to either GLP1-RA or lifestyle, we revealed only a 1% difference in the probability of obesity (72% [GLP1-RA] vs 71% [lifestyle]). This finding was corroborated when estimating the probability of normal weight (6% probability) and overweight (21% probability), where we revealed no differences in the probabilities between the treatments.

However, when estimating the probability of severe steatosis (>10%) at 24 weeks, we revealed an 18% lower probability when fixing the treatment variable to lifestyle treatment vs GLP-1RA (39% vs 57%, respectively). Again, this finding was corroborated when estimating the probability of normal liver fat (<5.6%) at 24 weeks, revealing that lifestyle treatment elicited a 12% higher probability of normal liver fat vs GLP1-RA (34% vs 22%, respectively). When exploring the probability of NASH (defined by: PDFF >5.6% & cT1 >800ms), fixing the treatment variable to lifestyle management again resulted in an 8% and 10% lower probability of NASH at 12 and 24 weeks, respectively, (12 weeks: 47% vs 55%; 24 weeks: 50% vs 60%).

Table 8.2. Probability of obesity status and liver events between lifestyle and GLP-1RA treatment arms.

	GLP-1RA	**Lifestyle**	**% difference**
Obesity status			
Obese	72%	71%	1%
Overweight	6%	6%	-
Normal weight	21%	21%	-
Liver outcomes			
Severe steatosis (>10%) – 24 weeks	57%	39%	18%
Healthy (<5.6%) – 24 weeks	22%	34%	12%
NASH – 12 weeks	55%	47%	8%
NASH – 24 weeks	60%	50%	10%

8.2.4. Discussion

Together, these results indicate that while treatment with GLP1-RA or lifestyle management may result in similar changes in obesity status at 24 weeks, lifestyle management may have a greater efficacy in reducing the probability of elevated liver fat and NASH.

While this pilot study was conducted in a relatively small cohort and contained a small number of clinical variables, it serves as a proof of concept to demonstrate how Bayesian-networks could be used for modelling dynamic changes across study visits and for estimating clinical outcomes to different treatment options. Further work on this topic should expand on this network to include a more comprehensive list of biomarkers, such as the body composition metrics used throughout this Thesis. Prescription of GLP1-RAs has shown to elicit significant reductions in body weight [165], but there exists an open question as to the *type* of weight loss occurring, with concerns over reductions in lean muscle mass in people with existing vulnerability. Therefore, applying DBNs to estimate potential reductions in body composition and particularly lean muscle would allow clinicians to make a more informed clinical decision given the risk of their specific patient. Additional work utilising DBNs may also include estimating the likelihood of meeting trial endpoints to assist with sample size powering, estimating the risk of potential side effects of specific treatments, and estimating the efficacy of combinatorial treatments before study initiation.

8.3. Area of further work 2 – pre-surgical decision support

The second area of further work is the application of Bayesian-networks for pre-surgical screening and decision planning. Surgery is a commonly used treatment option for patients with liver cancer, the long-term survival prospects of which are generally positive. That said, in the week after surgery patients can sometimes develop unforeseen liver problems, requiring hospital re-admittance and prolonged clinical care. The causes of post-surgical complications are multi-faceted, though the health and size of the liver remaining after surgery has emerged as a significant component. Perspectum have developed a pre-surgical screening tool (Hepatica) that provides a 3D quantitative assessment of the patients' liver state, allowing surgeons to make an informed decision on their surgical procedure and post-surgical care. However, despite this support tool, the question posed by the clinician holding such data still remains inherently *counterfactual* – 'what is the probability that my patient develops post-surgical complications, *given* I remove 40% of their liver, or they are 70 years old with secondary comorbidities?'

For this reason, I conjecture that Bayesian-networks are also a promising tool for aiding with surgical decision making and therefore support clinicians with optimising treatment strategies and recovery planning. The example network in Fig. 8.4 contains metrics offered using Hepatica and some routinely collected biomarkers, constructed in collaboration with Dr John Connell – the clinical lead for Hepatica. Using such network, we could pose the following three example questions using the techniques demonstrated throughout this Thesis:

1. What is the probability that my patient develops post-surgical complications, *given* their future liver remnant is 20%, they are aged 65 and have concurrent obesity and type-2 diabetes?
2. What is the probability that my patient requires 2-weeks in hospital post-surgery, *given* they have a Child-Pugh score of Class B, a future liver remnant of 15% and I performed a right hepatectomy procedure?
3. Should I proceed with the planned surgery, *given* my patient has a liver PDFF of 20% (severe steatosis), a liver cT1 of 900ms (severe liver disease activity), and is severely obese?

The above questions are common in pre-surgical planning, and this area of further work illustrates how Bayesian-networks could be merged with surgical planning to optimise patient care. Again, this another example of a multi-disciplinary approach for promoting personalised medicine.

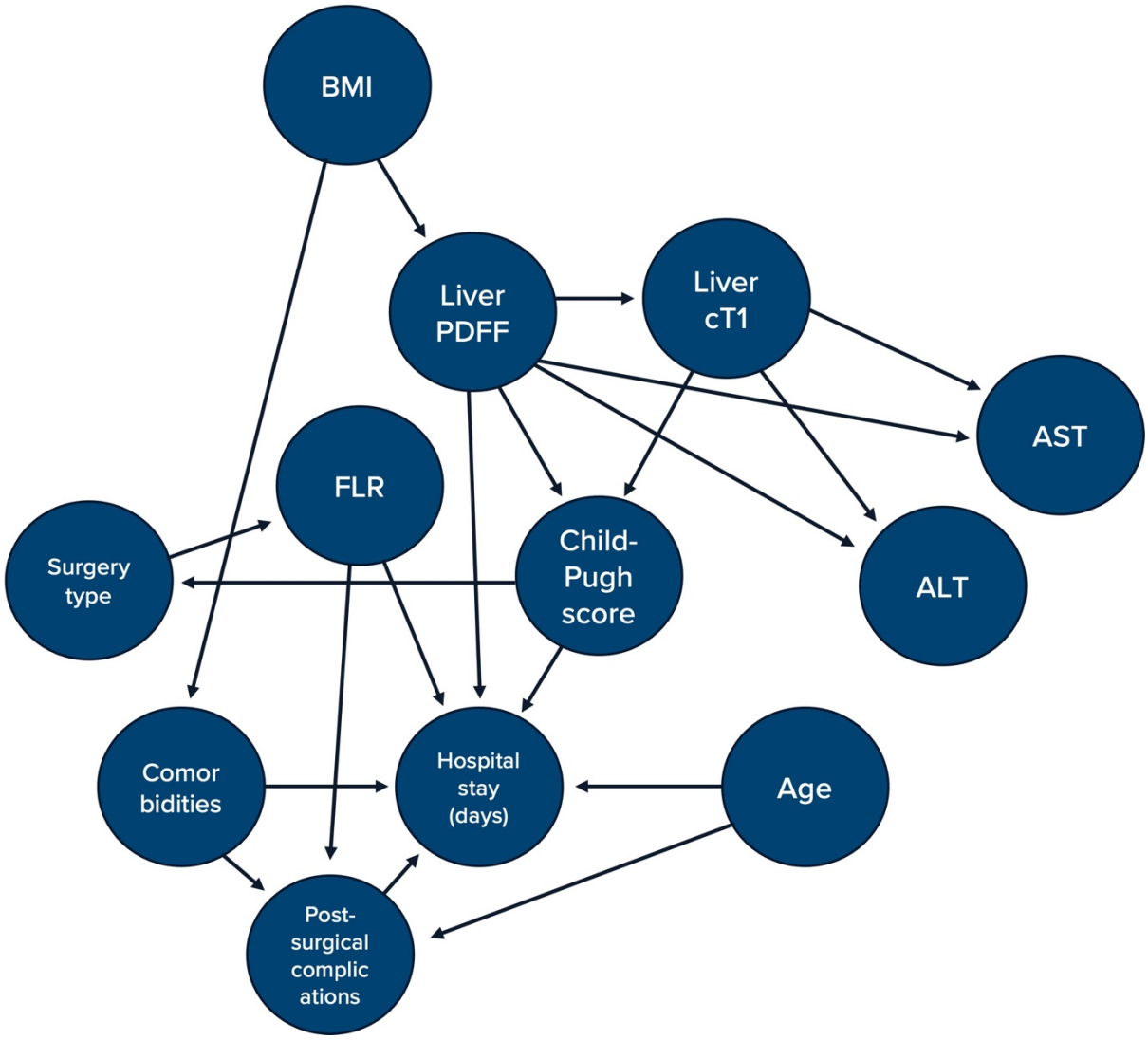


Figure 8.4. Example Bayesian-network using Hepatica biomarkers for liver surgical planning; FLR = future liver remnant.

8.4. Final summary of Thesis

This Thesis sought to examine the application of Bayesian-networks for providing support to clinical decision making. To this end, it applied Bayesian-networks to clinical data derived from multiple sources, namely the UK Biobank, to explore specific phenotypes and clinical

outcomes. However, the Bayesian-networks and work contained within this Thesis were built using small and limited datasets with unbalanced cohorts. This can lead to several potential flaws in the analysis as discussed throughout this Thesis, namely the risk of artefacts and false inference when performing probabilistic inference. Again, the results in this analysis should be viewed with caution. Rather, the objective of this Thesis was to examine the application of Bayesian-networks, not to create novel breakthroughs in physiological or diagnostics, but instead highlight their potential, of which I believe there is plenty when built using sufficient data.

The future of medicine will see the continual merging of AI systems into routine clinical care, not simply replacing the clinician, but as a diagnostic aid when the decision making is too complex or the scale of patient screening too large. However, as we have echoed throughout this Thesis, we must ensure the technology is interpretable for a clinician. The true utility of Bayesian-networks lies in the ability to speak the language of clinicians, model causal relationships and ask counterfactual questions to explore individual phenotypes. Ultimately, Bayesian-networks allow the researcher and clinician to ask the critical question – why, and, what if?

REFERENCES

[1] G. Eknoyan, 'A History of Obesity, or How What Was Good Became Ugly and Then Bad', *Adv. Chronic Kidney Dis.*, vol. 13, no. 4, pp. 421–427, Oct. 2006, doi: 10.1053/j.ackd.2006.07.002.

[2] 'Health matters: obesity and the food environment', *GOV.UK*. https://www.gov.uk/government/publications/health-matters-obesity-and-the-food-environment/health-matters-obesity-and-the-food-environment--2 (accessed Aug. 30, 2023).

[3] A. Hruby and F. B. Hu, 'The Epidemiology of Obesity: A Big Picture', *PharmacoEconomics*, vol. 33, no. 7, pp. 673–689, Jul. 2015, doi: 10.1007/s40273-014-0243-x.

[4] E. Bennett, S. A. E. Peters, and M. Woodward, 'Sex differences in macronutrient intake and adherence to dietary recommendations: findings from the UK Biobank', *BMJ Open*, vol. 8, no. 4, p. e020017, Apr. 2018, doi: 10.1136/bmjopen-2017-020017.

[5] K. McPherson, T. Marsh, and M. Brown, 'Tackling Obesities: Future Choices--Modelling Future Trends in Obesity & Their Impact on Health: (602672011-001)'. 2007. doi: 10.1037/e602672011-001.

[6] C. Baker, 'Obesity statistics', Aug. 2023, Accessed: Aug. 30, 2023. [Online]. Available: https://commonslibrary.parliament.uk/research-briefings/sn03336/

[7] 'Obesity and overweight'. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight (accessed Aug. 30, 2023).

[8] 'Childhood Obesity', *World Obesity Federation*. https://www.worldobesity.org/what-we-do/our-policy-priorities/childhood-obesity (accessed Aug. 30, 2023).

[9] S. R. Daniels, 'Complications of obesity in children and adolescents', *Int. J. Obes.*, vol. 33, no. 1, Art. no. 1, Apr. 2009, doi: 10.1038/ijo.2009.20.

[10] J. J. Reilly, 'Obesity in childhood and adolescence: evidence based clinical and public health perspectives', *Postgrad. Med. J.*, vol. 82, no. 969, pp. 429–437, Jul. 2006, doi: 10.1136/pgmj.2005.043836.

[11] J. Smith, 'New analysis estimates over 21 million UK adults will be obese by 2040', *Cancer Research UK - Cancer News*, May 18, 2022. https://news.cancerresearchuk.org/2022/05/19/new-analysis-estimates-over-21-million-uk-adults-will-be-obese-by-2040/ (accessed Aug. 30, 2023).

[12] 'Tackling obesities: future choices - project report (2nd edition)'.

[13] E.-J. Rhee, 'The Influence of Obesity and Metabolic Health on Vascular Health', *Endocrinol. Metab.*, vol. 37, no. 1, pp. 1–8, Feb. 2022, doi: 10.3803/EnM.2022.101.

[14] W. R. Rowley, C. Bezold, Y. Arikan, E. Byrne, and S. Krohe, 'Diabetes 2030: Insights from Yesterday, Today, and Future Trends', *Popul. Health Manag.*, vol. 20, no. 1, pp. 6–12, Feb. 2017, doi: 10.1089/pop.2015.0181.

[15] R. Kumar, R. N. Priyadarshi, and U. Anand, 'Non-alcoholic Fatty Liver Disease: Growing Burden, Adverse Outcomes and Associations', *J. Clin. Transl. Hepatol.*, vol. 8, no. 1, pp. 76–86, Mar. 2020, doi: 10.14218/JCTH.2019.00051.

[16] A. Abdullah, A. Peeters, M. de Courten, and J. Stoelwinder, 'The magnitude of association between overweight and obesity and the risk of diabetes: a meta-analysis of prospective cohort studies', *Diabetes Res. Clin. Pract.*, vol. 89, no. 3, pp. 309–319, Sep. 2010, doi: 10.1016/j.diabres.2010.04.012.

[17] Prospective Studies Collaboration *et al.*, 'Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies', *Lancet Lond.*

Engl., vol. 373, no. 9669, pp. 1083–1096, Mar. 2009, doi: 10.1016/S0140-6736(09)60318-4.
[18] T. M. Powell-Wiley *et al.*, 'Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association', *Circulation*, May 2021, doi: 10.1161/CIR.0000000000000973.
[19] S. Singh, A. M. Allen, Z. Wang, L. J. Prokop, M. H. Murad, and R. Loomba, 'Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies', *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.*, vol. 13, no. 4, pp. 643-654.e1–9; quiz e39-40, Apr. 2015, doi: 10.1016/j.cgh.2014.04.014.
[20] T. Waddell *et al.*, 'Greater ectopic fat deposition and liver fibroinflammation and lower skeletal muscle mass in people with type 2 diabetes', *Obes. Silver Spring Md*, vol. 30, no. 6, pp. 1231–1238, Jun. 2022, doi: 10.1002/oby.23425.
[21] L. Calzadilla Bertot and L. A. Adams, 'The Natural Course of Non-Alcoholic Fatty Liver Disease', *Int. J. Mol. Sci.*, vol. 17, no. 5, p. 774, May 2016, doi: 10.3390/ijms17050774.
[22] G. Musso *et al.*, 'Association of non-alcoholic fatty liver disease with chronic kidney disease: a systematic review and meta-analysis', *PLoS Med.*, vol. 11, no. 7, p. e1001680, Jul. 2014, doi: 10.1371/journal.pmed.1001680.
[23] A. Mantovani *et al.*, 'Nonalcoholic fatty liver disease increases risk of incident chronic kidney disease: A systematic review and meta-analysis', *Metabolism.*, vol. 79, pp. 64–76, Feb. 2018, doi: 10.1016/j.metabol.2017.11.003.
[24] Y. S. Khalid *et al.*, 'Increased cardiovascular events and mortality in females with NAFLD: a meta-analysis', *Am. J. Cardiovasc. Dis.*, vol. 10, no. 3, pp. 258–271, Aug. 2020.
[25] 'Diabetes and heart disease', *Diabetes UK*. https://www.diabetes.org.uk/guide-to-diabetes/complications/cardiovascular_disease (accessed Aug. 30, 2023).
[26] CDC, 'Heart Disease Facts | cdc.gov', *Centers for Disease Control and Prevention*, May 15, 2023. https://www.cdc.gov/heartdisease/facts.htm (accessed Aug. 30, 2023).
[27] B. Cepeda-Valery, G. S. Pressman, V. M. Figueredo, and A. Romero-Corral, 'Impact of obesity on total and cardiovascular mortality--fat or fiction?', *Nat. Rev. Cardiol.*, vol. 8, no. 4, pp. 233–237, Apr. 2011, doi: 10.1038/nrcardio.2010.209.
[28] A. Morgan *et al.*, 'Disease burden and economic impact of diagnosed non-alcoholic steatohepatitis (NASH) in the United Kingdom (UK) in 2018', *Eur. J. Health Econ.*, vol. 22, no. 4, pp. 505–518, 2021, doi: 10.1007/s10198-020-01256-y.
[29] J.-P. Després *et al.*, 'Abdominal obesity and the metabolic syndrome: contribution to global cardiometabolic risk', *Arterioscler. Thromb. Vasc. Biol.*, vol. 28, no. 6, pp. 1039–1049, Jun. 2008, doi: 10.1161/ATVBAHA.107.159228.
[30] L. Heilbronn, S. R. Smith, and E. Ravussin, 'Failure of fat cell proliferation, mitochondrial function and fat oxidation results in ectopic fat storage, insulin resistance and type II diabetes mellitus', *Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes.*, vol. 28 Suppl 4, pp. S12-21, Dec. 2004, doi: 10.1038/sj.ijo.0802853.
[31] K. Karastergiou, 'The Interplay Between Sex, Ethnicity, and Adipose Tissue Characteristics', *Curr. Obes. Rep.*, vol. 4, no. 2, pp. 269–278, Jun. 2015, doi: 10.1007/s13679-015-0149-8.
[32] L. M. Goff, 'Ethnicity and Type 2 diabetes in the UK', *Diabet. Med.*, vol. 36, no. 8, pp. 927–938, 2019, doi: 10.1111/dme.13895.
[33] P. Gupta *et al.*, 'The Association between Body Composition using Dual energy X-ray Absorptiometry and Type-2 Diabetes: A Systematic Review and Meta-Analysis of

Observational studies', *Sci. Rep.*, vol. 9, p. 12634, Sep. 2019, doi: 10.1038/s41598-019-49162-5.
[34] R. Taylor, 'Type 2 diabetes: etiology and reversibility', *Diabetes Care*, vol. 36, no. 4, pp. 1047–1055, Apr. 2013, doi: 10.2337/dc12-1805.
[35] R. V. Shah *et al.*, 'Visceral adiposity and the risk of metabolic syndrome across body mass index: the MESA Study', *JACC Cardiovasc. Imaging*, vol. 7, no. 12, pp. 1221–1235, Dec. 2014, doi: 10.1016/j.jcmg.2014.07.017.
[36] Z. M. Younossi *et al.*, 'The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: A systematic review and meta-analysis', *J. Hepatol.*, vol. 71, no. 4, pp. 793–801, Oct. 2019, doi: 10.1016/j.jhep.2019.06.021.
[37] J.-P. Després, 'Body Fat Distribution and Risk of Cardiovascular Disease', *Circulation*, vol. 126, no. 10, pp. 1301–1313, Sep. 2012, doi: 10.1161/CIRCULATIONAHA.111.067264.
[38] M. Haberka *et al.*, 'Obesity, Visceral Fat, and Hypertension-Related Complications', *Metab. Syndr. Relat. Disord.*, vol. 16, no. 10, pp. 521–529, Dec. 2018, doi: 10.1089/met.2018.0062.
[39] R. Banerjee *et al.*, 'Multiparametric magnetic resonance for the non-invasive diagnosis of liver disease', *J. Hepatol.*, vol. 60, no. 1, pp. 69–77, Jan. 2014, doi: 10.1016/j.jhep.2013.09.002.
[40] M. Blüher, 'Metabolically Healthy Obesity', *Endocr. Rev.*, vol. 41, no. 3, p. bnaa004, Jun. 2020, doi: 10.1210/endrev/bnaa004.
[41] R. Wagner *et al.*, 'Pathophysiology-based subphenotyping of individuals at elevated risk for type 2 diabetes', *Nat. Med.*, vol. 27, no. 1, Art. no. 1, Jan. 2021, doi: 10.1038/s41591-020-1116-9.
[42] A. Wesolowska-Andersen *et al.*, 'Four groups of type 2 diabetes contribute to the etiological and clinical heterogeneity in newly diagnosed individuals: An IMI DIRECT study', *Cell Rep. Med.*, vol. 3, no. 1, p. 100477, Jan. 2022, doi: 10.1016/j.xcrm.2021.100477.
[43] M. Pavlides *et al.*, 'Multiparametric magnetic resonance imaging for the assessment of non-alcoholic fatty liver disease severity', *Liver Int. Off. J. Int. Assoc. Study Liver*, vol. 37, no. 7, pp. 1065–1073, Jul. 2017, doi: 10.1111/liv.13284.
[44] P. J. Eddowes *et al.*, 'Utility and cost evaluation of multiparametric magnetic resonance imaging for the assessment of non-alcoholic fatty liver disease', *Aliment. Pharmacol. Ther.*, vol. 47, no. 5, pp. 631–644, Mar. 2018, doi: 10.1111/apt.14469.
[45] M. Pavlides *et al.*, 'Multiparametric magnetic resonance imaging predicts clinical outcomes in patients with chronic liver disease', *J. Hepatol.*, vol. 64, no. 2, pp. 308–315, Feb. 2016, doi: 10.1016/j.jhep.2015.10.009.
[46] A. Dennis *et al.*, 'A composite biomarker using multiparametric magnetic resonance imaging and blood analytes accurately identifies patients with non-alcoholic steatohepatitis and significant fibrosis', *Sci. Rep.*, vol. 10, no. 1, p. 15308, Sep. 2020, doi: 10.1038/s41598-020-71995-8.
[47] A. Andersson *et al.*, 'Clinical Utility of Magnetic Resonance Imaging Biomarkers for Identifying Nonalcoholic Steatohepatitis Patients at High Risk of Progression: A Multicenter Pooled Data and Meta-Analysis', *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.*, vol. 20, no. 11, pp. 2451-2461.e3, Nov. 2022, doi: 10.1016/j.cgh.2021.09.041.
[48] K. Imajo *et al.*, 'Quantitative multiparametric magnetic resonance imaging can aid non-alcoholic steatohepatitis diagnosis in a Japanese cohort', *World J. Gastroenterol.*, vol. 27, no. 7, pp. 609–623, Feb. 2021, doi: 10.3748/wjg.v27.i7.609.

[49] W. Shen *et al.*, 'Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image', *J. Appl. Physiol. Bethesda Md 1985*, vol. 97, no. 6, pp. 2333–2338, Dec. 2004, doi: 10.1152/japplphysiol.00744.2004.
[50] E. W. Demerath *et al.*, 'Approximation of total visceral adipose tissue with a single magnetic resonance image', *Am. J. Clin. Nutr.*, vol. 85, no. 2, pp. 362–368, Feb. 2007.
[51] L. Schweitzer *et al.*, 'What is the best reference site for a single MRI slice to assess whole-body skeletal muscle and adipose tissue volumes in healthy adults?', *Am. J. Clin. Nutr.*, vol. 102, no. 1, pp. 58–65, Jul. 2015, doi: 10.3945/ajcn.115.111203.
[52] N. S. Patel *et al.*, 'Effect of Weight Loss on MRI Estimation of Liver Fat and Volume in Patients With Nonalcoholic Steatohepatitis', *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.*, vol. 13, no. 3, pp. 561-568.e1, Mar. 2015, doi: 10.1016/j.cgh.2014.08.039.
[53] F. Magkos *et al.*, 'Effects of moderate and subsequent progressive weight loss on metabolic function and adipose tissue biology in humans with obesity', *Cell Metab.*, vol. 23, no. 4, pp. 591–601, Apr. 2016, doi: 10.1016/j.cmet.2016.02.005.
[54] E. Vilar-Gomez *et al.*, 'Weight Loss Through Lifestyle Modification Significantly Reduces Features of Nonalcoholic Steatohepatitis', *Gastroenterology*, vol. 149, no. 2, pp. 367-378.e5; quiz e14-15, Aug. 2015, doi: 10.1053/j.gastro.2015.04.005.
[55] A. Mardinoglu *et al.*, 'An Integrated Understanding of the Rapid Metabolic Benefits of a Carbohydrate-Restricted Diet on Hepatic Steatosis in Humans', *Cell Metab.*, vol. 27, no. 3, pp. 559-571.e5, Mar. 2018, doi: 10.1016/j.cmet.2018.01.005.
[56] F. Sofi and A. Casini, 'Mediterranean diet and non-alcoholic fatty liver disease: new therapeutic option around the corner?', *World J. Gastroenterol.*, vol. 20, no. 23, pp. 7339–7346, Jun. 2014, doi: 10.3748/wjg.v20.i23.7339.
[57] T. Y. N. Tong, N. J. Wareham, K.-T. Khaw, F. Imamura, and N. G. Forouhi, 'Prospective association of the Mediterranean diet with cardiovascular disease incidence and mortality and its population impact in a non-Mediterranean population: the EPIC-Norfolk study', *BMC Med.*, vol. 14, no. 1, p. 135, Sep. 2016, doi: 10.1186/s12916-016-0677-4.
[58] K. Esposito, M. I. Maiorino, G. Bellastella, P. Chiodini, D. Panagiotakos, and D. Giugliano, 'A journey into a Mediterranean diet and type 2 diabetes: a systematic review with meta-analyses', *BMJ Open*, vol. 5, no. 8, p. e008222, Aug. 2015, doi: 10.1136/bmjopen-2015-008222.
[59] C.-M. Kastorini, H. J. Milionis, K. Esposito, D. Giugliano, J. A. Goudevenos, and D. B. Panagiotakos, 'The effect of Mediterranean diet on metabolic syndrome and its components: a meta-analysis of 50 studies and 534,906 individuals', *J. Am. Coll. Cardiol.*, vol. 57, no. 11, pp. 1299–1313, Mar. 2011, doi: 10.1016/j.jacc.2010.09.073.
[60] E. Bacchi *et al.*, 'Both resistance training and aerobic training reduce hepatic fat content in type 2 diabetic subjects with nonalcoholic fatty liver disease (the RAED2 Randomized Trial)', *Hepatol. Baltim. Md*, vol. 58, no. 4, pp. 1287–1295, Oct. 2013, doi: 10.1002/hep.26393.
[61] C. A. Slentz *et al.*, 'Effects of aerobic vs. resistance training on visceral and liver fat stores, liver enzymes, and insulin resistance by HOMA in overweight adults from STRRIDE AT/RT', *Am. J. Physiol. Endocrinol. Metab.*, vol. 301, no. 5, pp. E1033-1039, Nov. 2011, doi: 10.1152/ajpendo.00291.2011.
[62] A. Wahid *et al.*, 'Quantifying the Association Between Physical Activity and Cardiovascular Disease and Diabetes: A Systematic Review and Meta-Analysis', *J. Am. Heart Assoc.*, vol. 5, no. 9, p. e002495, Sep. 2016, doi: 10.1161/JAHA.115.002495.

[63] H. Hamasaki, 'Daily physical activity and type 2 diabetes: A review', *World J. Diabetes*, vol. 7, no. 12, pp. 243–251, Jun. 2016, doi: 10.4239/wjd.v7.i12.243.
[64] E. A. Richter and N. B. Ruderman, 'AMPK and the biochemistry of exercise: Implications for human health and disease', *Biochem. J.*, vol. 418, no. 2, pp. 261–275, Mar. 2009, doi: 10.1042/BJ20082055.
[65] D. K. McGuire *et al.*, 'Effects of empagliflozin on first and recurrent clinical events in patients with type 2 diabetes and atherosclerotic cardiovascular disease: a secondary analysis of the EMPA-REG OUTCOME trial', *Lancet Diabetes Endocrinol.*, vol. 8, no. 12, pp. 949–959, Dec. 2020, doi: 10.1016/S2213-8587(20)30344-2.
[66] S. P. Marso *et al.*, 'Design of the liraglutide effect and action in diabetes: evaluation of cardiovascular outcome results (LEADER) trial', *Am. Heart J.*, vol. 166, no. 5, pp. 823-830.e5, Nov. 2013, doi: 10.1016/j.ahj.2013.07.012.
[67] H. J. L. Heerspink *et al.*, 'Kidney outcomes associated with use of SGLT2 inhibitors in real-world clinical practice (CVD-REAL 3): a multinational observational cohort study', *Lancet Diabetes Endocrinol.*, vol. 8, no. 1, pp. 27–35, Jan. 2020, doi: 10.1016/S2213-8587(19)30384-5.
[68] J. M. Dennis *et al.*, 'Time trends in prescribing of type 2 diabetes drugs, glycaemic response and risk factors: A retrospective analysis of primary care data, 2010-2017', *Diabetes Obes. Metab.*, vol. 21, no. 7, pp. 1576–1584, Jul. 2019, doi: 10.1111/dom.13687.
[69] K. Pearson, *The Grammar of Science*. Adam and Charles Black, 1900.
[70] P. C. Ilie, S. Stefanescu, and L. Smith, 'The role of vitamin D in the prevention of coronavirus disease 2019 infection and mortality', *Aging Clin. Exp. Res.*, vol. 32, no. 7, pp. 1195–1198, 2020, doi: 10.1007/s40520-020-01570-8.
[71] A. Srivastava and V. Kumar, 'Spurious Correlation? A review of the relationship between Vitamin D and Covid-19 infection and mortality'. medRxiv, p. 2020.05.25.20110338, May 26, 2020. doi: 10.1101/2020.05.25.20110338.
[72] J. Pearl and D. Mackenzie, *The book of why: the new science of cause and effect*. Basic books, 2018.
[73] W. A. Woods, 'What's in a link: Foundations for semantic networks', in *Representation and understanding*, Elsevier, 1975, pp. 35–82.
[74] J. Pearl, 'The solution for the branching factor of the alpha-beta pruning algorithm and its optimality', *Commun. ACM*, vol. 25, no. 8, pp. 559–564, Aug. 1982, doi: 10.1145/358589.358616.
[75] S. Andreassen, M. Woldbye, B. Falck, and S. K. Andersen, 'MUNIN: a causal probabilistic network for interpretation of electromyographic findings', in *Proceedings of the 10th international joint conference on Artificial intelligence - Volume 1*, in IJCAI'87. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., Aug. 1987, pp. 366–372.
[76] D. E. Heckerman, E. J. Horvitz, and B. N. Nathwani, 'Toward Normative Expert Systems: Part I The Pathfinder Project', *Methods Inf. Med.*, vol. 31, no. 02, pp. 90–105, 1992, doi: 10.1055/s-0038-1634867.
[77] J. Pearl, M. Glymour, and N. P. Jewell, *Causal Inference in Statistics: A Primer*. John Wiley & Sons, 2016.
[78] M. Scutari, 'Dirichlet Bayesian Network Scores and the Maximum Relative Entropy Principle'. arXiv, Mar. 05, 2018. Accessed: Sep. 05, 2023. [Online]. Available: http://arxiv.org/abs/1708.00689
[79] G. F. Cooper and E. Herskovits, 'A Bayesian method for the induction of probabilistic networks from data', *Mach. Learn.*, vol. 9, no. 4, pp. 309–347, Oct. 1992, doi: 10.1007/BF00994110.

[80] R. W. Robinson, 'Counting unlabeled acyclic digraphs', C. H. C. Little, Ed., in Lecture Notes in Mathematics, vol. 622. Berlin, Heidelberg: Springer Berlin Heidelberg, 1977, pp. 28–43. doi: 10.1007/BFb0069178.
[81] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*, vol. 81. 1993. doi: 10.1007/978-1-4612-2748-9.
[82] T. S. Verma and J. Pearl, 'On the Equivalence of Causal Models'. arXiv, Mar. 27, 2013. doi: 10.48550/arXiv.1304.1108.
[83] G. Schwarz, 'Estimating the Dimension of a Model', *Ann. Stat.*, vol. 6, no. 2, pp. 461–464, 1978.
[84] U. E. N. C. for E. Assessment, 'Information theory and an extension of the maximum likelihood principle', Mar. 15, 2009. https://hero.epa.gov/hero/index.cfm/reference/details/reference_id/591 (accessed Aug. 29, 2023).
[85] B. Selman and C. P. Gomes, 'Hill-climbing Search', in *Encyclopedia of Cognitive Science*, L. Nadel, Ed., 1st ed.Wiley, 2006. doi: 10.1002/0470018860.s00015.
[86] T. Verma and J. Pearl, 'Equivalence and synthesis of causal models', in *Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence*, in UAI '90. USA: Elsevier Science Inc., Jul. 1990, pp. 255–270.
[87] I. Tsamardinos, C. Aliferis, and A. Statnikov, *Algorithms for Large Scale Markov Blanket Discovery*. 2003, p. 381.
[88] C. Meek, 'Causal Inference and Causal Explanation with Background Knowledge'. arXiv, Feb. 20, 2013. doi: 10.48550/arXiv.1302.4972.
[89] I. Tsamardinos, L. E. Brown, and C. F. Aliferis, 'The max-min hill-climbing Bayesian network structure learning algorithm', *Mach. Learn.*, vol. 65, no. 1, pp. 31–78, Oct. 2006, doi: 10.1007/s10994-006-6889-7.
[90] R. Chelouah and P. Siarry, 'Tabu Search applied to global optimization', *Eur. J. Oper. Res.*, vol. 123, no. 2, pp. 256–270, Jun. 2000, doi: 10.1016/S0377-2217(99)00255-6.
[91] P. Spirtes, C. Glymour, S. N, and Richard, 'Causation, Prediction, and Search', 2000. https://philarchive.org/rec/SPICPA-2 (accessed Sep. 04, 2023).
[92] D. Colombo and M. H. Maathuis, 'Order-independent constraint-based causal structure learning'. arXiv, Sep. 27, 2013. doi: 10.48550/arXiv.1211.3295.
[93] S. Yaramakala and D. Margaritis, 'Speculative Markov Blanket Discovery for Optimal Feature Selection', *Fifth IEEE Int. Conf. Data Min. ICDM05*, pp. 809–812, 2005, doi: 10.1109/ICDM.2005.134.
[94] M. Gasse, A. Aussem, and H. Elghazel, 'A hybrid algorithm for Bayesian network structure learning with application to multi-label learning', *Expert Syst. Appl.*, vol. 41, no. 15, pp. 6755–6772, Nov. 2014, doi: 10.1016/j.eswa.2014.04.032.
[95] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. 2009.
[96] D. P. Kroese and R. Y. Rubinstein, 'Monte Carlo methods', *WIREs Comput. Stat.*, vol. 4, no. 1, pp. 48–58, 2012, doi: 10.1002/wics.194.
[97] S. Russell and P. Norvig, 'Artificial Intelligence: a modern approach, 4th US ed', *Univ. Calif. Berkeley*, 2021.
[98] H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, 'Bayesian Inference Identifies Combination Therapeutic Targets in Breast Cancer', *IEEE Trans. Biomed. Eng.*, vol. 66, no. 9, pp. 2684–2692, Sep. 2019, doi: 10.1109/TBME.2019.2894980.
[99] A. Bradley, R. Van der Meer, and C. J. McKay, 'A prognostic Bayesian network that makes personalized predictions of poor prognostic outcome post resection of pancreatic ductal adenocarcinoma', *PloS One*, vol. 14, no. 9, p. e0222270, 2019, doi: 10.1371/journal.pone.0222270.

[100] S. Walczak and V. Velanovich, 'An Evaluation of Artificial Neural Networks in Predicting Pancreatic Cancer Survival', *J. Gastrointest. Surg. Off. J. Soc. Surg. Aliment. Tract*, vol. 21, no. 10, pp. 1606–1612, Oct. 2017, doi: 10.1007/s11605-017-3518-7.
[101] Y. Cao, M. Raoof, E. Szabo, J. Ottosson, and I. Näslund, 'Using Bayesian Networks to Predict Long-Term Health-Related Quality of Life and Comorbidity after Bariatric Surgery: A Study Based on the Scandinavian Obesity Surgery Registry', *J. Clin. Med.*, vol. 9, no. 6, p. 1895, Jun. 2020, doi: 10.3390/jcm9061895.
[102] J. Xie, Y. Liu, X. Zeng, W. Zhang, and Z. Mei, 'A Bayesian network model for predicting type 2 diabetes risk based on electronic health records', *Mod. Phys. Lett. B*, vol. 31, no. 19–21, p. 1740055, Jul. 2017, doi: 10.1142/S0217984917400553.
[103] Y. Sase *et al.*, 'Characteristics of Type-2 Diabetics Who are Prone to High-Cost Medical Care Expenses by Bayesian Network', *Int. J. Environ. Res. Public. Health*, vol. 17, no. 15, p. 5271, Jul. 2020, doi: 10.3390/ijerph17155271.
[104] K. Orphanou, A. Stassopoulou, and E. Keravnou, 'Risk Assessment for Primary Coronary Heart Disease Event Using Dynamic Bayesian Networks', J. H. Holmes, R. Bellazzi, L. Sacchi, and N. Peek, Eds., in Lecture Notes in Computer Science, vol. 9105. Cham: Springer International Publishing, 2015, pp. 161–165. doi: 10.1007/978-3-319-19551-3_20.
[105] Y. Li *et al.*, 'Prediction models for acute kidney injury in patients with gastrointestinal cancers: a real-world study based on Bayesian networks', *Ren. Fail.*, vol. 42, no. 1, pp. 869–876, Nov. 2020, doi: 10.1080/0886022X.2020.1810068.
[106] S. Tian, M. Bi, Y. Bi, X. Che, and Y. Liu, 'A Bayesian Network Analysis of the Probabilistic Relationships Between Various Obesity Phenotypes and Cardiovascular Disease Risk in Chinese Adults: Chinese Population-Based Observational Study', *JMIR Med. Inform.*, vol. 10, no. 3, p. e33026, Mar. 2022, doi: 10.2196/33026.
[107] S. Xu *et al.*, 'Modeling interrelationships between health behaviors in overweight breast cancer survivors: Applying Bayesian networks', *PloS One*, vol. 13, no. 9, p. e0202923, 2018, doi: 10.1371/journal.pone.0202923.
[108] M. R. Neves *et al.*, 'Causal Dynamic Bayesian Networks for the Management of Glucose Control in Gestational Diabetes', in *2021 IEEE 9th International Conference on Healthcare Informatics (ICHI)*, Aug. 2021, pp. 31–40. doi: 10.1109/ICHI52183.2021.00018.
[109] A. Hill *et al.*, 'A Bayesian Network Decision Support Tool for Low Back Pain Using a RAND Appropriateness Procedure: Proposal and Internal Pilot Study', *JMIR Res. Protoc.*, vol. 10, no. 1, p. e21804, Jan. 2021, doi: 10.2196/21804.
[110] A. Fahmi *et al.*, 'Causal Bayesian Networks for Medical Diagnosis: A Case Study in Rheumatoid Arthritis', in *2020 IEEE International Conference on Healthcare Informatics (ICHI)*, Nov. 2020, pp. 1–7. doi: 10.1109/ICHI48887.2020.9374327.
[111] A. I. F. Poon and J. J. Y. Sung, 'Opening the black box of AI-Medicine', *J. Gastroenterol. Hepatol.*, vol. 36, no. 3, pp. 581–584, Mar. 2021, doi: 10.1111/jgh.15384.
[112] A. Aussem, A. Tchernof, S. R. de Morais, and S. Rome, 'Analysis of lifestyle and metabolic predictors of visceral obesity with Bayesian Networks', *BMC Bioinformatics*, vol. 11, no. 1, p. 487, Sep. 2010, doi: 10.1186/1471-2105-11-487.
[113] C. A. Matteoni, Z. M. Younossi, T. Gramlich, N. Boparai, Y. C. Liu, and A. J. McCullough, 'Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity', *Gastroenterology*, vol. 116, no. 6, pp. 1413–1419, Jun. 1999, doi: 10.1016/s0016-5085(99)70506-8.

[114] E. Chang, M. Varghese, and K. Singer, 'Gender and Sex Differences in Adipose Tissue', *Curr. Diab. Rep.*, vol. 18, no. 9, p. 69, Jul. 2018, doi: 10.1007/s11892-018-1031-3.
[115] L. J. M. Heyens, D. Busschots, G. H. Koek, G. Robaeys, and S. Francque, 'Liver Fibrosis in Non-alcoholic Fatty Liver Disease: From Liver Biopsy to Non-invasive Biomarkers in Diagnosis and Treatment', *Front. Med.*, vol. 8, 2021, Accessed: Aug. 29, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fmed.2021.615978
[116] M.-J. Honka *et al.*, 'Insulin-stimulated glucose uptake in skeletal muscle, adipose tissue and liver: a positron emission tomography study', *Eur. J. Endocrinol.*, vol. 178, no. 5, pp. 523–531, May 2018, doi: 10.1530/EJE-17-0882.
[117] R. A. DeFronzo, 'From the Triumvirate to the Ominous Octet: A New Paradigm for the Treatment of Type 2 Diabetes Mellitus', *Diabetes*, vol. 58, no. 4, pp. 773–795, Apr. 2009, doi: 10.2337/db09-9028.
[118] F. Colace, M. De Santo, M. Vento, and P. Foggia, *Bayesian Network Structural Learning from Data: An Algorithms Comparison.* 2004, p. 530.
[119] A. D. Booth *et al.*, 'Subcutaneous adipose tissue accumulation protects systemic glucose tolerance and muscle metabolism', *Adipocyte*, vol. 7, no. 4, pp. 261–272, 2018, doi: 10.1080/21623945.2018.1525252.
[120] M. V. Chakravarthy, T. Waddell, R. Banerjee, and N. Guess, 'Nutrition and Nonalcoholic Fatty Liver Disease: Current Perspectives', *Gastroenterol. Clin. North Am.*, vol. 49, no. 1, pp. 63–94, Mar. 2020, doi: 10.1016/j.gtc.2019.09.003.
[121] A. Al-Mrabeh, 'Pathogenesis and remission of type 2 diabetes: what has the twin cycle hypothesis taught us?', *Cardiovasc. Endocrinol. Metab.*, vol. 9, no. 4, pp. 132–142, Dec. 2020, doi: 10.1097/XCE.0000000000000201.
[122] K. W. M. Abeysekera *et al.*, 'Prevalence of steatosis and fibrosis in young adults in the UK: a population-based study', *Lancet Gastroenterol. Hepatol.*, vol. 5, no. 3, pp. 295–305, Mar. 2020, doi: 10.1016/S2468-1253(19)30419-4.
[123] 'Liver disease in numbers – key facts and statistics', *British Liver Trust*. https://britishlivertrust.org.uk/information-and-support/statistics/ (accessed Aug. 30, 2023).
[124] S. Mayor, 'Socioeconomic disadvantage is linked to obesity across generations, UK study finds', *BMJ*, vol. 356, p. j163, Jan. 2017, doi: 10.1136/bmj.j163.
[125] J. R. Petrie, T. J. Guzik, and R. M. Touyz, 'Diabetes, Hypertension, and Cardiovascular Disease: Clinical Insights and Vascular Mechanisms', *Can. J. Cardiol.*, vol. 34, no. 5, pp. 575–584, May 2018, doi: 10.1016/j.cjca.2017.12.005.
[126] P. Saeedi *et al.*, 'Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition', *Diabetes Res. Clin. Pract.*, vol. 157, p. 107843, Nov. 2019, doi: 10.1016/j.diabres.2019.107843.
[127] P. Golabi, M. Otgonsuren, L. de Avila, M. Sayiner, N. Rafiq, and Z. M. Younossi, 'Components of metabolic syndrome increase the risk of mortality in nonalcoholic fatty liver disease (NAFLD)', *Medicine (Baltimore)*, vol. 97, no. 13, p. e0214, Mar. 2018, doi: 10.1097/MD.0000000000010214.
[128] Z. M. Younossi, T. Gramlich, C. A. Matteoni, N. Boparai, and A. J. McCullough, 'Nonalcoholic fatty liver disease in patients with type 2 diabetes', *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.*, vol. 2, no. 3, pp. 262–265, Mar. 2004, doi: 10.1016/s1542-3565(04)00014-x.

[129] T. Tada *et al.*, 'Type 2 diabetes mellitus: A risk factor for progression of liver fibrosis in middle-aged patients with non-alcoholic fatty liver disease', *J. Gastroenterol. Hepatol.*, vol. 34, no. 11, pp. 2011–2018, Nov. 2019, doi: 10.1111/jgh.14734.
[130] H. B. El-Serag, T. Tran, and J. E. Everhart, 'Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma', *Gastroenterology*, vol. 126, no. 2, pp. 460–468, Feb. 2004, doi: 10.1053/j.gastro.2003.10.065.
[131] Z. M. Younossi *et al.*, 'Economic and Clinical Burden of Nonalcoholic Steatohepatitis in Patients With Type 2 Diabetes in the U.S', *Diabetes Care*, vol. 43, no. 2, pp. 283–289, Feb. 2020, doi: 10.2337/dc19-1113.
[132] E. Levelt *et al.*, 'Ectopic and Visceral Fat Deposition in Lean and Obese Patients With Type 2 Diabetes', *J. Am. Coll. Cardiol.*, vol. 68, no. 1, pp. 53–63, Jul. 2016, doi: 10.1016/j.jacc.2016.03.597.
[133] N. EICHERT *et al.*, '60-OR: High Prevalence of Multiorgan Steatosis and Fibroinflammation, Identified by Multiparametric Magnetic Resonance Imaging, in People with Type 2 Diabetes', *Diabetes*, vol. 71, no. Supplement_1, pp. 60-OR, Jun. 2022, doi: 10.2337/db22-60-OR.
[134] S. J. Yu *et al.*, 'Visceral Obesity Predicts Significant Fibrosis in Patients With Nonalcoholic Fatty Liver Disease', *Medicine (Baltimore)*, vol. 94, no. 48, p. e2159, Dec. 2015, doi: 10.1097/MD.0000000000002159.
[135] A. Schäffler, J. Schölmerich, and C. Büchler, 'Mechanisms of disease: adipocytokines and visceral adipose tissue--emerging role in nonalcoholic fatty liver disease', *Nat. Clin. Pract. Gastroenterol. Hepatol.*, vol. 2, no. 6, pp. 273–280, Jun. 2005, doi: 10.1038/ncpgasthep0186.
[136] M. Małodobra-Mazur, A. Cierzniak, D. Pawełka, K. Kaliszewski, J. Rudnicki, and T. Dobosz, 'Metabolic Differences between Subcutaneous and Visceral Adipocytes Differentiated with an Excess of Saturated and Monounsaturated Fatty Acids', *Genes*, vol. 11, no. 9, p. 1092, Sep. 2020, doi: 10.3390/genes11091092.
[137] A. Gastaldelli and K. Cusi, 'From NASH to diabetes and from diabetes to NASH: Mechanisms and treatment options', *JHEP Rep. Innov. Hepatol.*, vol. 1, no. 4, pp. 312–328, Oct. 2019, doi: 10.1016/j.jhepr.2019.07.002.
[138] E. Zhang, Y. Zhao, and H. Hu, 'Impact of Sodium Glucose Cotransporter 2 Inhibitors on Nonalcoholic Fatty Liver Disease Complicated by Diabetes Mellitus', *Hepatol. Commun.*, vol. 5, no. 5, pp. 736–748, Apr. 2021, doi: 10.1002/hep4.1611.
[139] P. Newsome *et al.*, 'Effect of semaglutide on liver enzymes and markers of inflammation in subjects with type 2 diabetes and/or obesity', *Aliment. Pharmacol. Ther.*, vol. 50, no. 2, pp. 193–203, Jul. 2019, doi: 10.1111/apt.15316.
[140] 'Cases in the UK | Coronavirus in the UK', Aug. 24, 2023. https://coronavirus.data.gov.uk/details/cases (accessed Aug. 30, 2023).
[141] E. Ladds *et al.*, 'Persistent symptoms after Covid-19: qualitative study of 114 "long Covid" patients and draft quality principles for services', *BMC Health Serv. Res.*, vol. 20, no. 1, p. 1144, Dec. 2020, doi: 10.1186/s12913-020-06001-y.
[142] A. Roca-Fernández *et al.*, 'Hepatic Steatosis, Rather Than Underlying Obesity, Increases the Risk of Infection and Hospitalization for COVID-19', *Front. Med.*, vol. 8, p. 636637, 2021, doi: 10.3389/fmed.2021.636637.
[143] 'Tackling obesity: empowering adults and children to live healthier lives', *GOV.UK*. https://www.gov.uk/government/publications/tackling-obesity-government-strategy/tackling-obesity-empowering-adults-and-children-to-live-healthier-lives (accessed Aug. 30, 2023).

[144] H. Ogata *et al.*, 'Impact of visceral fat on the prognosis of coronavirus disease 2019: an observational cohort study', *BMC Infect. Dis.*, vol. 21, no. 1, p. 1240, Dec. 2021, doi: 10.1186/s12879-021-06958-z.
[145] K. M. Bunnell *et al.*, 'Body composition predictors of outcome in patients with COVID-19', *Int. J. Obes.*, vol. 45, no. 10, Art. no. 10, Oct. 2021, doi: 10.1038/s41366-021-00907-1.
[146] A. Rawshani *et al.*, 'Severe COVID-19 in people with type 1 and type 2 diabetes in Sweden: A nationwide retrospective cohort study', *Lancet Reg. Health Eur.*, vol. 4, p. 100105, May 2021, doi: 10.1016/j.lanepe.2021.100105.
[147] A. Rajpal, L. Rahimi, and F. Ismail-Beigi, 'Factors leading to high morbidity and mortality of COVID-19 in patients with type 2 diabetes', *J. Diabetes*, vol. 12, no. 12, pp. 895–908, Dec. 2020, doi: 10.1111/1753-0407.13085.
[148] R. Butcher and N. Fenton, 'Extending the range of symptoms in a Bayesian Network for the Predictive Diagnosis of COVID-19'. medRxiv, p. 2020.10.22.20217554, Oct. 26, 2020. doi: 10.1101/2020.10.22.20217554.
[149] M. Neil, N. Fenton, M. Osman, and S. McLachlan, 'Bayesian network analysis of Covid-19 data reveals higher infection prevalence rates and lower fatality rates than widely reported', *J. Risk Res.*, vol. 23, no. 7–8, pp. 866–879, Aug. 2020, doi: 10.1080/13669877.2020.1778771.
[150] N. E. Fenton *et al.*, 'A Bayesian network model for personalised COVID19 risk assessment and contact tracing'. medRxiv, p. 2020.07.15.20154286, Jan. 05, 2021. doi: 10.1101/2020.07.15.20154286.
[151] R. Cartin-Ceba *et al.*, 'Evidence showing lipotoxicity worsens outcomes in covid-19 patients and insights about the underlying mechanisms', *iScience*, vol. 25, no. 5, p. 104322, May 2022, doi: 10.1016/j.isci.2022.104322.
[152] M. Naeem *et al.*, 'Pathogenetic Mechanisms of Liver-Associated Injuries, Management, and Current Challenges in COVID-19 Patients', *Biomolecules*, vol. 13, no. 1, p. 99, Jan. 2023, doi: 10.3390/biom13010099.
[153] B. M. Volk *et al.*, 'Reduced COVID-19 severity elicited by weight loss from a medically supervised ketogenic diet in a geographically diverse ambulatory population with type 2 diabetes and obesity', *BMJ Nutr. Prev. Health*, p. e000444, Jul. 2022, doi: 10.1136/bmjnph-2022-000444.
[154] R. B. Luo *et al.*, 'How bariatric surgery affects liver volume and fat density in NAFLD patients', *Surg. Endosc.*, vol. 32, no. 4, pp. 1675–1682, Apr. 2018, doi: 10.1007/s00464-017-5846-9.
[155] A. Flint *et al.*, 'Randomised clinical trial: semaglutide versus placebo reduced liver steatosis but not liver stiffness in subjects with non-alcoholic fatty liver disease assessed by magnetic resonance imaging', *Aliment. Pharmacol. Ther.*, vol. 54, no. 9, pp. 1150–1161, Nov. 2021, doi: 10.1111/apt.16608.
[156] K. M. Gadde and S. B. Heymsfield, 'Targeting visceral adiposity with pharmacotherapy', *Lancet Diabetes Endocrinol.*, vol. 9, no. 9, pp. 551–552, Sep. 2021, doi: 10.1016/S2213-8587(21)00204-7.
[157] A. J. London, 'Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability', *Hastings Cent. Rep.*, vol. 49, no. 1, pp. 15–21, Jan. 2019, doi: 10.1002/hast.973.
[158] J. Varghese, 'Artificial Intelligence in Medicine: Chances and Challenges for Wide Clinical Adoption', *Visc. Med.*, vol. 36, no. 6, pp. 443–449, Dec. 2020, doi: 10.1159/000511930.

[159] K. Tunyasuvunakool *et al.*, 'Highly accurate protein structure prediction for the human proteome', *Nature*, vol. 596, no. 7873, Art. no. 7873, Aug. 2021, doi: 10.1038/s41586-021-03828-1.
[160] P. P. Massion *et al.*, 'Assessing the Accuracy of a Deep Learning Method to Risk Stratify Indeterminate Pulmonary Nodules', *Am. J. Respir. Crit. Care Med.*, vol. 202, no. 2, pp. 241–249, Jul. 2020, doi: 10.1164/rccm.201903-0505OC.
[161] R. Nevola *et al.*, 'GLP-1 Receptor Agonists in Non-Alcoholic Fatty Liver Disease: Current Evidence and Future Perspectives', *Int. J. Mol. Sci.*, vol. 24, no. 2, p. 1703, Jan. 2023, doi: 10.3390/ijms24021703.
[162] J. Schattenberg *et al.*, 'Non-invasive tests of liver injury, inflammation and fibrosis are improved by efruxifermin and correlate with histological improvements in F2-F3 NASH patients: secondary analysis of Ph2b HARMONY study', *J. Hepatol.*, vol. 78, pp. S808–S809, Jun. 2023, doi: 10.1016/S0168-8278(23)02265-1.
[163] S. A. Harrison *et al.*, 'A phase 2, proof of concept, randomised controlled trial of berberine ursodeoxycholate in patients with presumed non-alcoholic steatohepatitis and type 2 diabetes', *Nat. Commun.*, vol. 12, p. 5503, Sep. 2021, doi: 10.1038/s41467-021-25701-5.
[164] P. Ladyzynski, M. Molik, and P. Foltynski, 'Dynamic Bayesian networks for prediction of health status and treatment effect in patients with chronic lymphocytic leukemia', *Sci. Rep.*, vol. 12, no. 1, Art. no. 1, Feb. 2022, doi: 10.1038/s41598-022-05813-8.
[165] M. Jensterle, M. Rizzo, M. Haluzík, and A. Janež, 'Efficacy of GLP-1 RA Approved for Weight Management in Patients With or Without Diabetes: A Narrative Review', *Adv. Ther.*, vol. 39, no. 6, pp. 2452–2467, Jun. 2022, doi: 10.1007/s12325-022-02153-x.

APPENDIX

1. Overview of example automated network construction algorithms.

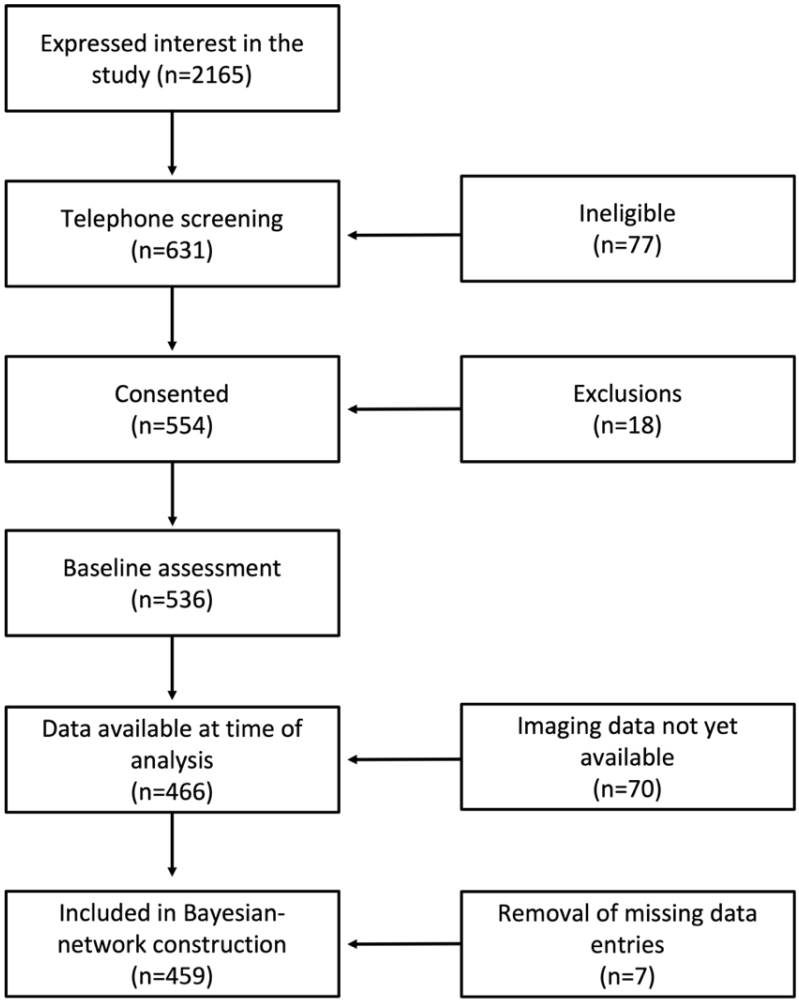
Score and search algorithms
Greedy search Hill-Climbing (HC): A greedy search hill-climb selects the local edge function (edge insertion, removal, or reversal) that elicits the greatest improvement in network score as determined by criterion such as the BIC and AIC. The search terminates and the DAG is returned once no local edge function can be found that improves the network score [85].
Tabu search (TS): Similar to HC, Tabu search selects the local edge function that elicits the greatest improvement in network score. However, TS employs a guided local search procedure that rejects any move within the search space that has already visited. This is achieved through maintaining a 'tabu list' of visited spaces to avoid a locally optimal, yet globally sub-optimal, network score [90].
Constraint-based algorithms
PC (stable) (PC): The PC stable algorithm is a modification of the PC algorithm detailed by [91]. This modification removes all or part of the order dependence, a complication of the original PC algorithm where the network output has shown to be dependent on the order in which the variables are introduced. The PC stable modification has demonstrated superior performance when applied to higher dimensionality datasets [92].
Fast Incremental Association (Fast.IAMB): The Fast Incremental Association Markov Blanket algorithm is a modification of the original Incremental Association Markov Blanket algorithm detailed by [87]. Fast.IAMB employs a heuristic to recover the Markov-blanket of a target variable faster and more reliably, without adversely affecting the discovered Markov blanket [93].
Hybrid algorithms
Max-Min Hill-Climbing (MMHC): The Max-Min Hill-Climbing algorithm first returns the network skeleton (edges, without orientation) through the local discovery Max-Min Parents and Children (MMPC) algorithm. Next, network edges are orientated using a greedy Bayesian-scoring hill-climbing search, where the search is constrained to only consider adding an edge if returned by the MMPC algorithm in the first instance [89].
Hybrid Parents and Children (H2PC): Like the MMHC algorithm, the Hybrid Parents and Children algorithm firstly learns the network skeleton through tests of conditional independence and performs a Bayesian-scoring greedy hill-climbing search to orientate network edges. The difference between the H2PC and MMHC algorithms is in the CI approach for returning the network skeleton, where H2PC instead applies an ensemble method combining the *Data-Efficient*

Parents and Children Superset, Data-Efficient Spouses Superset and Interleaved Incremental Association Parents and Children to return the parents and children of a target variable [94].

2. Discretization thresholds used in 'Applying Bayesian-networks to study concomitant non-alcoholic steatohepatitis in people with Type-2 diabetes (Chapter 6).

Liver PDFF (%)	[0] - <5.6% [1] – 5.6-10% [2] - >10%
Liver cT1 (ms)	[0] - <800ms [1] – 800-875ms [3] - >875ms
Pancreatic PDFF (%)	[0] - <6.6% [1] - >6.6%
VAT (cm^2)	**Male** [0] - <210(cm^2) [1] - >210(cm^2) **Female** [0] - <140(cm^2) [1] - >140(cm^2)
SAT (cm^2)	**Male** [0] - <220(cm^2) [1] - >220(cm^2) **Female** [0] - <400(cm^2) [1] - >400(cm^2)
SMI (cm^2/m)	**Male** [0] - >44(cm^2/m) [1] - <44(cm^2/m) **Female** [0] - >31(cm^2/m) [1] - <31(cm^2/m)
Body mass index (kg/m^2)	[0] - <25(kg/m^2) [1] – 25-30 (kg/m^2) [2] – 30-40(kg/m^2) [3] - >40(kg/m^2)
Age (yrs)	[0] - <40yrs [1] – 40-50yrs [2] – 50-60yrs [3] – 60-70yrs [4] - >70yrs
HbA1c (mmol/mol)	**T2D** [0] - <62mmol/mol [1] - >62mmol/mol **Non-T2D** [0] - <38mmol/mol [1] - >38mmol/mol

3. Cohort flow diagram from COVERSCAN analysis in 'Applying Bayesian-networks to study hospitalisation risk following acute COVID-19' (Chapter 7); Waddell et al. (2023). *Frontiers in Bioinformatics.*



4. Discretization thresholds used in 'Applying Bayesian-networks to study hospitalisation risk following acute COVID-19' (Chapter 7).

Variable	Discretisation
Liver PDFF (%) Relevant publication: Roca-Fernández, A., Dennis, A., Nicholls, R., McGonigle, J., Kelly, M., Banerjee, R., Banerjee, A. and Sanyal, A.J., 2021. Hepatic steatosis, rather than underlying obesity, increases the risk of infection and hospitalization for COVID-19. *Frontiers in Medicine*, *8*, p.636637.	<5% - normal (n = 370) 5-10% - mild steatosis (n = 40) >10% - severe steatosis (n = 49)
VAT (cm^2) Elevated measures defined as values greater than the 75th percentile derived from descriptive statistics.	**Male** Normal - <210(cm^2) Elevated - >210(cm^2) **Female** Normal - <140(cm^2) Elevated - >140(cm^2) **Total** Normal (n = 364) Elevated (n = 95)
SAT (cm^2) Elevated measures defined as values greater than the 75th percentile derived from descriptive statistics.	**Male** Normal - <220(cm^2) Elevated - >220(cm^2) **Female** Normal - <400(cm^2) Elevated - >400(cm^2) **Total** Normal (n = 312) Elevated (n = 147)
SMI (cm^2/m) Reduced measures defined as values less than the 25th percentile derived from descriptive statistics.	**Male** Normal- >44(cm^2/m) Reduced - <44(cm^2/m) **Female** Normal - >31(cm^2/m) Reduced - <31(cm^2/m) **Total** Normal (n = 422) Reduced (n = 37)
Body mass index (kg/m^2)	Normal weight - <25(kg/m^2) (n = 210) Overweight – 25-30 (kg/m^2) (n = 148) Obese – >30(kg/m^2) (n = 101)
Age (yrs)	<40yrs (n = 146) 40-50yrs (n = 163) 50-60yrs (n = 113) >60yrs (n = 37)