

## **Inferring causal pathways between metabolic processes and liver fat accumulation: an IMI DIRECT study**

Naeimeh Atabaki-Pasdar<sup>1†</sup>, Hugo Pomares-Millan<sup>1†</sup>, Robert W Koivula<sup>1,2</sup>, Andrea Tura<sup>3</sup>, Andrew Brown<sup>4</sup>, Ana Viñuela<sup>5,6</sup>, Leandro Agudelo<sup>7,8</sup>, Daniel Coral<sup>1</sup>, Sabine van Oort<sup>9</sup>, Kristine Allin<sup>10</sup>, Elizaveta Chabanova<sup>11</sup>, Henna Cederberg<sup>12,13</sup>, Federico De Masi<sup>14</sup>, Petra Elders<sup>15</sup>, Juan Fernandez Tajés<sup>1</sup>, Ian M Forgie<sup>16</sup>, Tue H Hansen<sup>10,17</sup>, Alison Heggie<sup>18</sup>, Angus Jones<sup>19,20</sup>, Tarja Kokkola<sup>21</sup>, Anubha Mahajan<sup>22</sup>, Timothy J McDonald<sup>19,23</sup>, Donna McEvoy<sup>24</sup>, Konstantinos Tsirigos<sup>25,26</sup>, Harriet Teare<sup>27</sup>, Jagadish Vangipurapu<sup>13</sup>, Henrik Vestergaard<sup>26,28</sup>, Jerzy Adamski<sup>29,30,31</sup>, Joline WJ Beulens<sup>9</sup>, Søren Brunak<sup>10,25</sup>, Emmanouil Dermitzakis<sup>5</sup>, Torben Hansen<sup>26</sup>, Andrew T Hattersley<sup>19,20</sup>, Markku Laakso<sup>13,21</sup>, Oluf Pedersen<sup>10</sup>, Martin Ridderstråle<sup>32</sup>, Hartmut Ruetten<sup>33</sup>, Femke Rutters<sup>9</sup>, Jochen M Schwenk<sup>34</sup>, Mark Walker<sup>18</sup>, Giuseppe N Giordano<sup>1</sup>, Mattias Ohlsson<sup>35,36</sup>, Ramneek Gupta<sup>14</sup>, Andrea Mari<sup>3</sup>, Mark I McCarthy<sup>2,22,37</sup>, E Louise Thomas<sup>38</sup>, Jimmy D Bell<sup>38</sup>, Imre Pavo<sup>39</sup>, Ewan R Pearson<sup>16</sup>, Paul W Franks<sup>1,40\*</sup>

### **Author information/affiliations**

1. Genetic and Molecular Epidemiology Unit, Lund University Diabetes Centre, Department of Clinical Sciences, Clinical Research Centre, Lund University, Skåne University Hospital, Malmö, Malmö, Sweden
2. Radcliffe Department of Medicine, Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, U.K.
3. CNR Institute of Neuroscience, Padova, Italy
4. Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee, Scotland, U.K.
5. Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland
6. Faculty of Medical Sciences, Biosciences Institute, Newcastle University, Newcastle, U.K.
7. Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge MA, USA
8. Broad Institute of MIT and Harvard, Cambridge MA, USA
9. Department of Epidemiology and Data Science, Amsterdam UMC–Location VUmc, Amsterdam Public Health Research Institute, Amsterdam, the Netherlands
10. Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

**NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

11. Department of Diagnostic Radiology, Copenhagen University Hospital Herlev  
Gentofte, Herlev, Denmark
12. Department of Endocrinology, Abdominal Centre, Helsinki University Hospital and  
University of Helsinki, Helsinki, Finland
13. Internal Medicine, Institute of Clinical Medicine, University of Eastern Finland,  
Kuopio, Finland
14. Department of Health Technology, Technical University of Denmark, Kongens  
Lyngby, Denmark
15. Department of General Practice, Amsterdam UMC–Location VUmc, Amsterdam  
Public Health Research Institute, Amsterdam, the Netherlands
16. Population Health and Genomics, School of Medicine, University of Dundee, Dundee,  
U.K.
17. Department of Medical Endocrinology and Metabolism, Rigshospitalet, Copenhagen  
University Hospital, Denmark
18. Faculty of Medical Sciences, Translational and Clinical Research Institute, Newcastle  
University, Newcastle, U.K.
19. Institute of Biomedical and Clinical Science, University of Exeter Medical School,  
Exeter, U.K.
20. Diabetes and Endocrinology, Royal Devon and Exeter NHS Foundation Trust, Exeter,  
U.K.
21. Department of Medicine, University of Eastern Finland and Kuopio University  
Hospital, Kuopio, Finland
22. Wellcome Centre for Human Genetics, University of Oxford, Oxford, U.K.
23. Blood Sciences, Royal Devon and Exeter NHS Foundation Trust, Exeter, U.K.
24. Diabetes Research Network, Royal Victoria Infirmary, Newcastle upon Tyne, U.K.
25. Department of Systems Biology, Center for Biological Sequence Analysis, Technical  
University of Denmark, Kongens Lyngby, Denmark
26. Novo Nordisk Foundation Center for Protein Research, University of Copenhagen,  
Copenhagen, Denmark
27. Centre for Health, Law and Emerging Technologies (HeLEX), Faculty of Law,  
University of Oxford, Oxford, U.K.
28. Bornholms Hospital, Rønne, Denmark
29. Institute of Experimental Genetics, Helmholtz Zentrum München, German Research  
Center for Environmental Health, Ingolstädter Landstraße 1, 85764 Neuherberg,  
Germany
30. Department of Biochemistry, Yong Loo Lin School of Medicine, National University  
of Singapore, 8 Medical Drive, Singapore 117597, Singapore

31. Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, 1000 Ljubljana, Slovenia
32. Clinical Obesity, Department of Clinical Sciences, Lund University, Malmö, Sweden
33. R&D Global Development, Translational Medicine & Clinical Pharmacology, Sanofi Deutschland GmbH, Frankfurt, Germany
34. Affinity Proteomics, Science for Life Laboratory, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Solna, Sweden
35. Computational Biology and Biological Physics Unit, Department of Astronomy and Theoretical Physics, Lund University, Lund, Sweden
36. Center for Applied Intelligent Systems Research, Halmstad University, Halmstad, Sweden
37. Oxford NIHR Biomedical Research Centre, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Oxford, U.K.
38. School of Life Sciences, Research Centre for Optimal Health, University of Westminster, London, U.K.
39. Eli Lilly Regional Operations GmbH, Vienna, Austria
40. Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA

† Joint first authors.

\*Corresponding author: Paul W. Franks,

Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Clinical Research Centre, Lund University, Jan Waldenströmsgata 35, Skåne University Hospital, Malmö SE-20502, Sweden

E-Mail: [paul.franks@med.lu.se](mailto:paul.franks@med.lu.se)

Orcid: [orcid.org/0000-0002-0520-7604](https://orcid.org/0000-0002-0520-7604)

## ABSTRACT

Type 2 diabetes (T2D) and non-alcoholic fatty liver disease (NAFLD) often co-occur. Defining causal pathways underlying this relationship may help optimize the prevention and treatment of both diseases. Thus, we assessed the strength and magnitude of the putative causal pathways linking dysglycemia and fatty liver, using a combination of causal inference methods.

Measures of glycemia, insulin dynamics, magnetic resonance imaging (MRI)-derived abdominal and liver fat content, serological biomarkers, lifestyle, and anthropometry were obtained in participants from the IMI DIRECT cohorts (n=795 with new onset T2D and 2234 individuals free from diabetes). UK Biobank (n=3641) was used for modelling and replication purposes. Bayesian networks were employed to infer causal pathways, with causal validation using two-sample Mendelian randomization.

Bayesian networks fitted to IMI DIRECT data identified higher basal insulin secretion rate (BasalISR) and MRI-derived excess visceral fat (VAT) accumulation as the features of dysmetabolism most likely to cause liver fat accumulation; the unconditional probability of fatty liver (>5%) increased significantly when conditioning on high levels of BasalISR and VAT (by 23%, 32% respectively; 40% for both). Analyses in UK Biobank yielded comparable results. MR confirmed most causal pathways predicted by the Bayesian networks.

Here, BasalISR had the highest causal effect on fatty liver predisposition, providing mechanistic evidence underpinning the established association of NAFLD and T2D. BasalISR may represent a pragmatic biomarker for NAFLD prediction in clinical practice.

## INTRODUCTION

Around 20-30% of adults in high-income countries have non-alcoholic fatty liver disease (NAFLD), with rates increasing worldwide<sup>1</sup>. For several decades, excessive intrahepatic fat accumulation has been recognized as a risk factor for type 2 diabetes (T2D)<sup>2</sup>, yet the evidence for this stem largely from cross-sectional studies<sup>3</sup>, which are prone to bias, confounding and reverse causation. Determining the causal nature of these relationships and their underlying metabolic mechanisms is necessary to optimize the prevention and treatment of NAFLD and T2D.

NAFLD is highly heritable ( $h^2 \sim 40\%$ )<sup>4</sup>, suggesting that, as with T2D, inherited DNA variants affect predisposition. Nevertheless, both diseases are also caused by numerous obesogenic environmental factors<sup>5</sup>. Excessive weight gain can cause fatty liver by initially raising circulating free fatty acid concentrations, which are then taken up by hepatocytes and stored as triacylglycerol. This in turn can trigger a cycle of liver-specific insulin resistance and excessive *de novo* hepatic glucose and lipid production<sup>6-8</sup>. This cycle of dyslipidemia and dysglycemia is postulated to link obesity, fatty liver and diabetes<sup>9-11</sup>.

Fatty liver pathology ranges from simple steatosis, through non-alcoholic steatohepatitis (NASH), liver fibrosis and ultimately cirrhosis. T2D and cardiovascular disease often co-occur with NAFLD, but the causal relationships are poorly defined<sup>12</sup>. Here, we examined a range of putative causal pathways linking the development of NAFLD with T2D using Bayesian network (BN) and bidirectional Mendelian randomization (MR) analyses.

## RESULTS

The primary analyses were conducted within IMI DIRECT, a multicenter prospective cohort study of 3029 European-ancestry adults diagnosed with T2D (n=795) or without T2D (n=2234); the latter group included participants with normal and prediabetic glycemia (ascertained by HbA1c, fasting glucose or 2-hour glucose). Of these participants, 1070 had the required variables for a complete case analysis with the mean (SD) age of 62 (5.0) years and 77.5% being male. In addition, the UK Biobank was utilized with the analysis sample restricted to white European ethnicity (n= 442580) and complete cases (n= 3641); in UK Biobank, the mean (SD) age of the cohort was 56.8 (8.0) years and 47.2% were male.

The following section describes the BNs in the IMI DIRECT and UK Biobank cohorts (subgroup analyses were performed in participants with and without T2D and in both sexes). The bidirectional two-sample MR results that were statistically significant after Benjamini-Hochberg False Discovery Rate (FDR) correction, at 5% threshold<sup>13</sup> per trait, were investigated for the directed association (arcs) that were present in the BNs. Detailed results for all bidirectional MR analyses performed here, including information about the published genome-wide association studies (GWAS) from which the genetic instruments are derived, are reported separately and per trait (S1 Table). The subset of statistically significant MR results is summarized in Table 1.

**Table 1. Two-sample MR exposure-outcome significant associations that achieved corrected threshold and sensitivity analysis**

Exposure	Outcome	# SNPs	IVW				MR-Egger			Sensitivity analysis		
			$\beta$	$\beta$ SE	P-value	P-value FDR corrected	$\beta$	$\beta$ SE	P-value	MR-Egger Intercept	MR-Egger Intercept P-value	MR-PRESSO global test P-value
<b>Glycemic traits</b>												
HbA1C	HDL	11	0.107	0.037	0.004	0.010	0.205	0.091	0.050	-0.004	0.236	0.543
HbA1C	ISI	11	-0.317	0.118	0.007	0.019	-0.291	0.321	0.389	-0.001	0.931	0.830
<b>Adiposity/obesity</b>												
BMI	Serum albumin level	62	-0.573	0.221	0.010	0.024	-0.873	0.638	0.177	0.017	0.232	0.046
BMI	Fasting blood insulin	67	0.172	0.019	6.25E-20	3.92E-19	0.126	0.055	0.025	0.001	0.464	0.073
BMI	Fasting blood glucose	66	0.058	0.024	0.016	0.039	0.109	0.070	0.124	-0.001	0.672	0.157
Waist circumference	Serum albumin levels	35	-0.678	0.272	0.013	0.032	-0.701	0.934	0.458	0.014	0.437	0.875
Waist circumference	Fasting blood insulin	41	0.232	0.025	1.11E-20	7.02E-20	0.141	0.087	0.114	0.002	0.19	0.064
Whole body fat mass	Modified Stumvoll Insulin Sensitivity Index (model adjusted for BMI)	184	0.111	0.045	0.013	0.032	0.133	0.137	0.333	0.0002	0.877	0.078
<b>Blood pressure</b>												
SBP	Two-hour glucose challenge	108	0.271	0.108	0.012	0.030	0.462	0.416	0.269	-0.004	0.615	0.497

We considered as significant if the directions of the estimates by IVW and MR-Egger were consistent, IVW method passed the FDR corrected threshold ( $p < 0.05$ ), and no significant pleiotropy tested by the MR-Egger intercept ( $p > 0.05$ ). Outlier SNPs were removed by MR-PRESSO at  $p < 0.05$  global test. IVW: inverse variance weighted; MR-PRESSO: MR pleiotropy residual sum and outlier; ISI corresponds to Modified Stumvoll Insulin Sensitivity Index; BMI, body mass index; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; SBP, systolic blood pressure; SNPs, single nucleotide polymorphisms.

### *Bayesian Network and Mendelian randomization analyses in the IMI DIRECT cohorts*

Prior to building the BNs, we applied cluster analysis to the input variables (the clustered variables are more likely to be present in the BN) (S1 Fig). A cluster comprising abdominal fat, body mass index (BMI), waist circumference and insulin measurements was observed in the Pearson correlation plot and the heatmap cluster analyses. These variables were initially selected based on their association with liver fat within the IMI DIRECT cohorts<sup>14</sup> or from existing literature<sup>15</sup>. We then proceeded with structural and parameter learnings using the variables as nodes to build the BN. To have a more stable network with information on the strengths of the arcs amongst the nodes, we performed network averaging of the bootstrapped BN samples. Table 2 reports the conditional density and parameters for each node of the constructed BN. Fig 1A shows the averaged BN and Fig 1B shows the cumulative distribution function of the arc strengths. Detailed information of those arcs with strength and directional probabilities equal to or greater than the significant threshold (0.505, learned from the data) is reported in S2 Table along with the available MR results. Basal insulin secretion rate (BasalISR), visceral adipose tissue (VAT) and 2-hour insulin (TwoInsulin) were identified as the causal upstream (parent) nodes, whereas liver iron, alanine transaminase (ALT) and gamma-glutamyl transpeptidase (GGTP) were defined as the downstream (child) effects of liver fat, given the averaged BN estimates. *Markov blanket* of liver fat (the set of nodes that includes adequate knowledge for inference) was identified as liver iron, TwoInsulin, BasalISR, VAT, subcutaneous adipose tissue (SAT), BMI, glycated hemoglobin A1C (HbA1c), fasting glucose, oral glucose insulin sensitivity (OGIS)<sup>16</sup>, total fasting plasma GLP-1 (TotGLP1min0), mean insulin clearance (Clins), fasting triglycerides (TG), ALT, aspartate transaminase (AST) and GGTP (green nodes in Fig 1A). A more robust parsimonious network was generated by restricting the arcs to those with both strength and direction probabilities above or equal to 0.8 (Fig 1C); BasalISR, VAT and TwoInsulin remained as the causal nodes and liver iron as the



downstream effect of liver fat. We then checked the averaged BN arcs with the two-sample MR results when available.

**Table 2. Conditional density and parameters for each node of the constructed average Bayesian network using IMI DIRECT inverse normal transformed data (n=1070).**

<b>Outcome   Exposures</b>
<b>LiverIron   SAT + LiverFat + HbA1c</b> Coefficients: (Intercept)= -4.159e-06, SAT=-1.282e-01, LiverFat=4.755e-01, HbA1c=-1.044e-01, SE:0.897
<b>PancIron   PancFat</b> Coefficients: (Intercept)=1.657e-07, PancFat=2.082e-01, SE=0.978
<b>PancFat   VAT</b> Coefficients: (Intercept)= 4.421e-07, VAT=3.296e-01, SE=0.944
<b>VAT   BasalISR</b> Coefficients: (Intercept)=-2.429e-07, BasalISR=5.272e-01, SE=0.849
<b>SAT   BMI</b> Coefficients: (Intercept)=-2.576e-08, BMI=7.429e-01, SE= 0.669
<b>LiverFat   VAT + TwoInsulin + BasalISR</b> Coefficients: (Intercept)=6.842e-07, VAT=2.264e-01, TwoInsulin=1.242e-01, BasalISR=3.402e-01, SE= 0.815
<b>Waist   VAT + TwoInsulin</b> Coefficients: (Intercept)=-0.00005, VAT=0.639, TwoInsulin=0.0601, SE= 0.747
<b>BMI   Waist + Insulin</b> Coefficients: (Intercept)=4.562e-05, Waist=7.735e-01, Insulin=1.310e-01, SE=0.532
<b>SBP   TwoGlucose</b> Coefficients: (Intercept)=-5.999e-07, TwoGlucose=1.934e-01, SE=0.981
<b>Glucose</b> Coefficients: (Intercept)=-2.147e-06, SE= 0.999
<b>TwoGlucose   HbA1c + TwoInsulin + BasalISR + OGIS + Clins + GlucoseSens</b> Coefficients: (Intercept)=9.620e-06, HbA1c=1.134e-01, TwoInsulin=4.586e-01, BasalISR=-1.800e-01, OGIS=-6.220e-01, Clins=3.865e-01, GlucoseSens=-1.889e-01, SE=0.602
<b>Insulin</b> Coefficients: (Intercept)=2.904e-06, SE= 0.999
<b>TwoInsulin   Insulin</b> Coefficients: (Intercept)=-1.383e-06, Insulin=6.385e-01, SE=0.769
<b>BasalISR   Insulin + OGIS</b> Coefficients: (Intercept)=-1.994e-06, Insulin=6.868e-01, OGIS=-2.444e-01, SE=0.521
<b>OGIS   Glucose + Insulin + TwoInsulin + Clins</b> Coefficients: (Intercept)=-1.119e-06, Glucose=-5.301e-01, Insulin=7.494e-02, TwoInsulin=5.031e-01, Clins=2.562e-01, SE= 0.397
<b>Clins   Insulin + TwoInsulin</b> Coefficients: (Intercept)=1.735e-06, Insulin=-5.600e-01, TwoInsulin=-2.325e-01, SE=0.683

---

**TotGLP1min0 | VAT + SAT + BasalISR**

Coefficients: (Intercept)=4.303e-08, VAT=2.206e-01, SAT=-1.593e-01, BasalISR=3.439e-01, SE=0.894

---

**Glucagonmin0 | TotGLP1min0 + HDL**

Coefficients: (Intercept)=2.097e-06, TotGLP1min0=2.691e-01, HDL=-1.938e-01, SE=0.935

---

**HDL | SAT + Insulin + TG + AST**

Coefficients: (Intercept)=-6.144e-06, SAT=3.617e-02, Insulin=-2.614e-01, TG=-3.025e-01, AST=1.675e-01, SE=0.898

---

**TG | VAT + Insulin + BasalISR**

Coefficients: (Intercept)=6.214e-06, VAT=2.002e-01, Insulin=-1.722e-01, BasalISR=4.882e-01, SE=0.872

---

**ALT | LiverFat + OGIS**

Coefficients: (Intercept)=0.0001, LiverFat=0.256, OGIS=-0.303, SE=0.880

---

**AST | ALT**

Coefficients: (Intercept)=2.102e-05, ALT=5.571e-01, SE=0.830

---

**GGTP | TG + AST**

Coefficients: (Intercept)=5.174e-05, TG=1.856e-01, AST=2.623e-01, SE=0.937

---

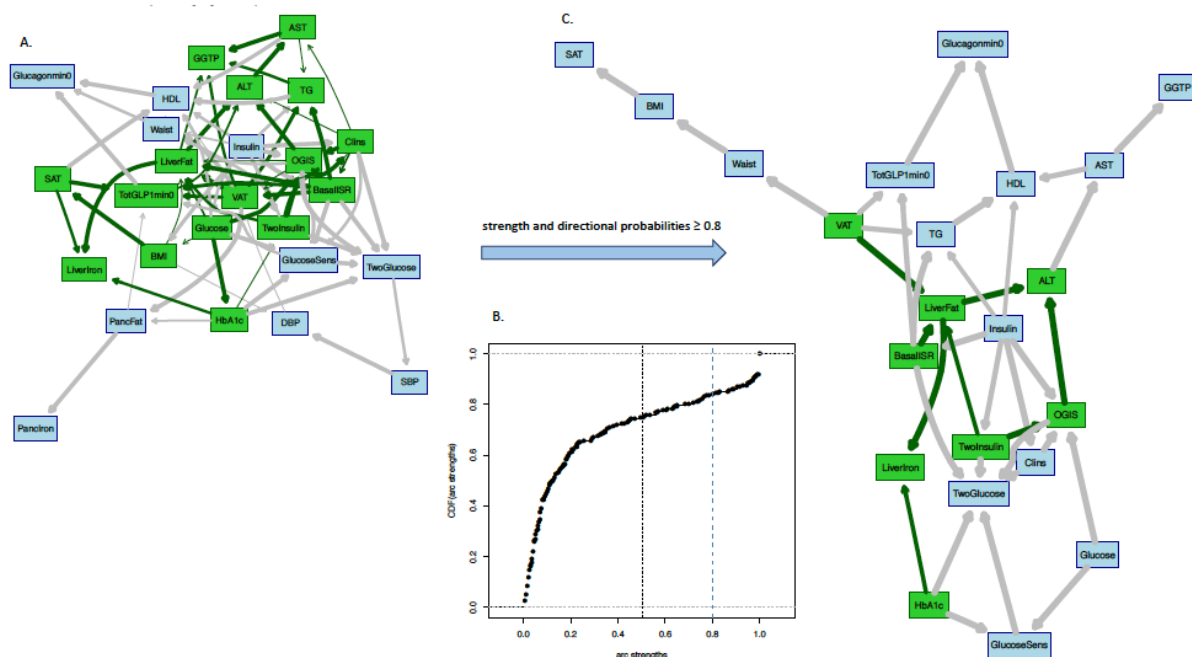
**GlucoseSens | HbA1c + Glucose + BasalISR + OGIS**

Coefficients: (Intercept)=-8.620e-06, HbA1c=-2.056e-01, Glucose=-4.695e-01, BasalISR=2.469e-01, OGIS=-2.982e-01, SE=0.864

---

ALT, alanine transaminase; AST, aspartate transaminase; BasalISR, insulin secretion at the beginning of the oral glucose tolerance test/ mixed-meal tolerance test; BMI, body mass index; Clins, mean insulin clearance during the oral glucose tolerance test/mixed meal tolerance test, calculated as (mean insulin secretion)/(mean insulin concentration); DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucagonmin0, fasting glucagon concentration; Glucose, fasting glucose from venous plasma samples; GlucoseSens, glucose sensitivity, slope of the dose-response relating insulin secretion to glucose concentration; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples; OGIS, oral glucose insulin sensitivity index according to the method of Mari et al. <sup>16</sup>; PancFat, pancreas fat; PancIron, pancreas iron; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; SE, standard error; TG, fasting triglycerides; TotGLP1min0, concentration of fasting total GLP-1 in plasma; TwoGlucose, 2-hour glucose after oral glucose tolerance test/mixed-meal tolerance test; TwoInsulin, 2-hour insulin; VAT, visceral adipose tissue.

---



**Fig1. Panel A:** Averaged Bayesian network of the bootstrapped samples among the variables of IMI DIRECT combined cohorts (data are inverse normal transformed,  $n=1070$ ). **Panel B:** Cumulative distribution function of the arc strengths presented in the Bayesian network of Panel A. Black vertical line is the threshold (0.505) learned from data, equal or above which the arcs are identified as statistically significant. **Panel C:** Subset of Bayesian networks in Panel A including only arcs with strength and directional probabilities  $\geq 0.8$ . Nodes in green highlight the Markov blanket of liver fat that includes the nodes with adequate information to stand as a separate Bayesian network.

ALT, alanine transaminase; AST, aspartate transaminase; BasalISR, insulin secretion at the beginning of the oral glucose tolerance test/ mixed-meal tolerance test; BMI, body mass index; Clins, mean insulin clearance during the oral glucose tolerance test/mixed meal tolerance test, calculated as (mean insulin secretion)/(mean insulin concentration); DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucagonmin0, fasting glucagon concentration; Glucose, fasting glucose from venous plasma samples; GlucoseSens, glucose sensitivity, slope of the dose–response relating insulin secretion to glucose concentration; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples; OGIS, oral glucose insulin sensitivity index according to the method of Mari et al.<sup>16</sup>; PancFat, pancreas fat; PancIron, pancreas iron; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; TotGLP1min0, concentration of fasting total GLP-1 in plasma; TwoGlucose, 2-hour glucose after oral glucose tolerance test/mixed-meal tolerance test; TwoInsulin, 2-hour insulin; VAT, visceral adipose tissue.

In the MR analyses, several arcs revealed in the BN were consistent in directionality and magnitude (results shown in SD units, otherwise stated). Among these arcs, waist circumference had a likely causal link to BMI ( $\beta = 1.01$ ;  $P_{\text{FDRcorrected}} = 4.10\text{E-}198$ ,  $P_{\text{Egger Intercept}} = 2.97\text{E-}01$ ). Additionally, as in the BN, causal effects were inferred for SBP to DBP ( $\beta = 0.64$ ;  $P_{\text{FDRcorrected}} = 5.95\text{E-}62$ ,  $P_{\text{Egger Intercept}} = 5.2\text{E-}01$ ). Moreover, AST (per  $\mu\text{g/L}$  increase) was causally associated with GGTP ( $\mu\text{g/L}$ ) ( $\beta = 0.2$ ;  $P_{\text{FDRcorrected}} = 1.14\text{E-}06$ ,  $P_{\text{Egger Intercept}} = 1.52\text{E-}$

01) and waist circumference was inversely causally related with HDL ( $\beta = -0.31$ ;  $P_{\text{FDRcorrected}} = 1.79\text{E-}10$ ,  $P_{\text{Egger Intercept}} = 8.84\text{E-}01$ ). A likely causal effect of TG on HDL ( $\beta = -0.12$ ;  $P_{\text{FDRcorrected}} = 7.04\text{E-}04$ ,  $P_{\text{Egger Intercept}} = 1.55\text{E-}01$ ) was also observed.

#### *Subgroup analysis: diabetes vs. non-diabetes & female vs. male (in IMI DIRECT)*

In order to elucidate causal networks that might be specific to diabetes or gender, analyses were reproduced within each subgroup. In the constructed BNs, BasalISR and VAT appeared to be the causal parental nodes for liver fat in both cohorts (S2 Fig, panels A and C). Subsets of these BNs with arcs having strengths and directional probabilities above or equal to 0.8 left only BasalISR as the causal parental node in both cohorts (S2 Fig, panels B and D). We then continued with the sensitivity analysis by gender, as illustrated in Figure S3. In the female's BN, BasalISR and BMI were the causal parental nodes for liver fat, whereas in the male's BN, BasalISR, VAT, TwoInsulin and HbA1c were the causal variables (S3 Fig, panel A and C). These BNs with extra restrictions on the strength and directional probabilities above or equal to 0.8 resulted in only BasalISR as the causal parental node for liver fat (S3 Fig, panel B and D). The Pearson correlation and the heatmap cluster analyses, performed in the sensitivity analyses resulted in similar clusters to the combined cohorts' one (see S1, S4 and S5 Figs). Detailed information of arcs with strength and directional probabilities above or equal to the BNs' thresholds (0.51, 0.5, 0.5 and 0.48 in the non-diabetes, diabetes, female and male BNs, respectively) are reported in S2 Table along with the corresponding MR results.

#### *Bayesian networks and posterior probabilistic inference (IMI DIRECT)*

In order to calculate the posterior probabilities (updated probabilities given prior knowledge) of the conditional and unconditional queries, all continuous variables were first discretized into three levels (low, moderate and high) using Hartemink's method<sup>17</sup> where variables are transformed to discrete variables with the same number of levels (see S3 Table). Liver fat was

considered as the variable of interest (*event*) where it was conditioned on some defined parental *evidence*; high BasalISR, high VAT and having high levels of both. Fig 2 summarizes unconditional and posterior conditional probabilities of liver fat on the combined, diabetes, non-diabetes, female and male groups of the IMI DIRECT BNs. A significant increase was observed after conditioning on high levels of BasalISR and VAT in the unconditional probability of high liver fat level on the random observations generated by the combined cohorts BNs: 23% increase after conditioning on VAT, 32% increase after conditioning on BasalISR, and 40% after conditioning on both. Similarly, diabetes, non-diabetes, female and male groups had substantial increments on the high liver fat level probability after conditioning on high levels of VAT, BasalISR and the combined effect. The probability of having high liver fat after conditioning on high levels of VAT and BasalISR was 0.84, 0.97, 0.86, 0.79 and 0.86 in the combined, diabetes, non-diabetes, female and male groups, respectively in the BNs constructed in IMI DIRECT.

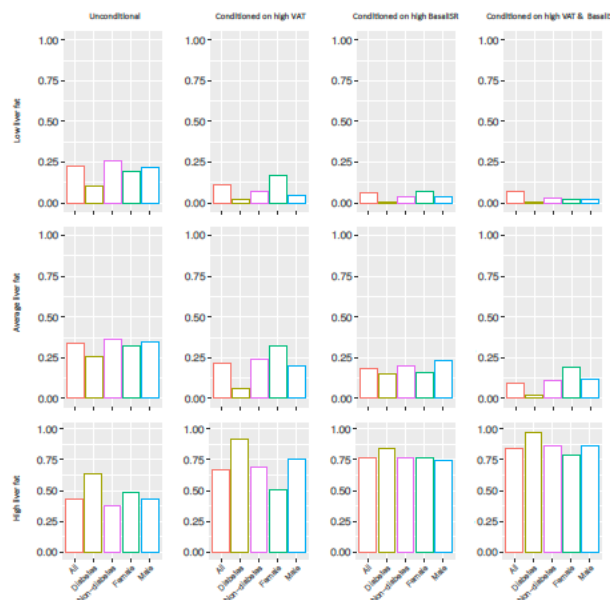


Fig2. Unconditional and posterior conditional probabilities of liver fat on the combined (n=1070), diabetes (n=283), non-diabetes (n=787), female (n=239) and male (n=831) groups of the IMI DIRECT Bayesian networks (data are inverse normal transformed). Y axis represents the unconditional and posterior conditional probabilities of low (0.1,1.95], average (1.95,4.5] or high (4.5,37.6] levels of liver fat in rows. The posterior probabilities were obtained after conditioning on high levels of basal insulin secretion rate (BasalISR) (141,425] (figures in 2<sup>nd</sup> column), high levels of visceral adipose tissue (VAT) (8.76,14.5] (figures in 3<sup>rd</sup> column) and high levels of both (figures in 4<sup>th</sup> column). Continuous variables were discretized with Hartemink's method<sup>17</sup>.

*Bayesian network and Mendelian randomization assessment of the Twin-cycle model (IMI DIRECT)*

Taylor's Twin-cycle model, a hypothesized mechanistic model linking NAFLD and T2D, postulates that chronic energy excess leads to accumulation of fat in the liver and subsequently the pancreas, which drives dysglycemia and diabetes. Here, we made a comparison of two models both using the same set of variables: i) an unsupervised BN and ii) the model hypothesized by Taylor. The bidirectional two-sample MR analyses was also used to test causal relationships between all possible combinations of selected variables. The unsupervised BN constructed using the variables of the Twin-cycle model is presented in Fig 3A, where panel B reports the strength and directional probabilities of the network's arcs. The arcs, mimicking the original Twin-cycle model, include fasting glucose and OGIS towards BasalISR, from glucose sensitivity and OGIS to 2-hour glucose, and from BasalISR to liver fat. Among those arcs not following the Twin-cycle model, TG to liver fat, glucose sensitivity to OGIS and fasting glucose to glucose sensitivity had weak directional probabilities (0.57, 0.66 and 0.60, respectively), which may help explain why the metabolic network predicted in our model is not fully compatible with the network hypothesized in the Twin-cycle model. Through the conducted MR analyses, only the association of fasting glucose to BasalISR could be tested, yet it did not reach the multiple testing corrected significance threshold ( $P_{\text{FDRcorrected}}=0.59$ ).

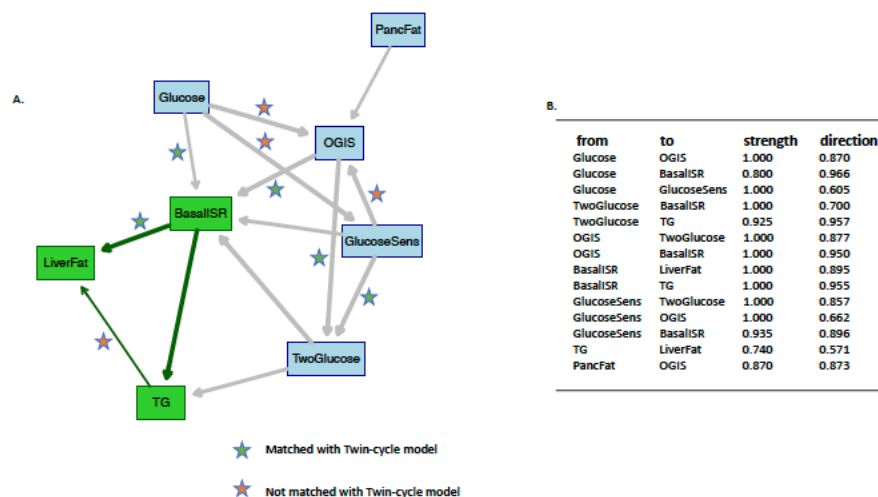


Fig3. **Panel A:** averaged Bayesian network of the bootstrapped samples among the variables of the Twin-cycle model using data from IMI DIRECT combined cohorts (data are inverse normal transformed,  $n=1264$ ). **Panel B:** strength and directional probabilities of the arcs presented in the Bayesian network of the panel A. Nodes in green highlight the Markov blanket of liver fat that includes the nodes with adequate information to stand as a separate Bayesian network.

BasalISR, insulin secretion at the beginning of the oral glucose tolerance test/ mixed-meal tolerance test; Glucose, fasting glucose from venous plasma samples; GlucoseSens, glucose sensitivity, slope of the dose–response relating insulin secretion to glucose concentration; OGIS, oral glucose insulin sensitivity index according to the method of Mari et al. <sup>16</sup>; PancFat, pancreas fat; TG, fasting triglycerides; TwoGlucose, 2-hour glucose after oral glucose tolerance test/mixed-meal tolerance test; TwoInsulin, 2-hour insulin.

### *Bayesian network and Mendelian randomization in the UK Biobank cohort*

As with the IMI DIRECT, we replicated the BN analyses in the UK Biobank dataset where data permitted. We first checked the heatmap cluster among the inverse normal transformed variables (S6 Fig A). The graphical connection of the variables with Pearson correlation  $\geq 0.4$  is also presented in S6 Fig B, identifying a cluster of abdominal fat, BMI, waist circumference and liver enzymes. As with the IMI DIRECT analyses, we performed network averaging of the bootstrapped BN samples in the UK Biobank dataset. Table 3 reports the conditional density and parameters for each node of the UK Biobank’s constructed BN. Panel A in Fig 4 shows the averaged BN and panel B illustrates the cumulative distribution function of the arc strengths. Detailed information of those arcs with strength and directional probabilities equal to or greater than the significant threshold (0.5) is reported in S2 Table, along with the available MR results.

VAT, SAT, waist circumference, ALT and Albumin were identified as the causal parental nodes, and the effect child nodes were HbA1c and TG. The Markov blanket of liver fat included VAT, SAT, waist circumference, whole-body fat mass, systolic blood pressure (SBP), ALT, GGTP, HDL, TG, Albumin, Bilirubin and HbA1c. Fig 4, Panel C shows the BN with only the strong arcs (both strength and directional probabilities  $\geq 0.8$ ), where only VAT and SAT remained as the causal nodes of liver fat. We then checked the averaged BN arcs with the two-sample MR results where possible.

Among these arcs, the consistent associations between the MR and BN were waist circumference to BMI (as described above for IMI DIRECT, results shown in SD units, otherwise stated). The association of HbA1c, preceded by fasting glucose (per mmol/L) was also noted as described above. BMI, and waist circumference were causally related to whole-body fat mass ( $\beta = 0.76$ ,  $P_{\text{FDRcorrected}} = 5.88\text{E-}171$ ,  $P_{\text{Egger Intercept}} = 3.4\text{E-}01$ ;  $\beta = 0.87$ ,  $P_{\text{FDRcorrected}} = 4.65\text{E-}139$ ,  $P_{\text{Egger Intercept}} = 6.44\text{E-}01$ , respectively). The associations linked with HDL, were those related to adiposity/anthropometry, i.e., waist circumference (as described above), and BMI ( $\beta = -0.31$ ,  $P_{\text{FDRcorrected}} = 1.79\text{E-}10$ ,  $P_{\text{Egger Intercept}} = 2.97\text{E-}01$ ;  $\beta = -0.25$ ,  $P_{\text{FDRcorrected}} = 3.96\text{E-}07$ ,  $P_{\text{Egger Intercept}} = 8.84\text{E-}01$ , respectively); moreover, ALT was associated with HDL ( $\beta = -0.28$ ,  $P_{\text{FDRcorrected}} = 3.61\text{E-}03$ ,  $P_{\text{Egger Intercept}} = 4.84\text{E-}01$ ). Waist circumference, HDL, GGTP, and whole-body fat mass were all associated with TG ( $\beta = 0.22$ ,  $P_{\text{FDRcorrected}} = 1.59\text{E-}08$ ,  $P_{\text{Egger Intercept}} = 2.97\text{E-}01$ ;  $\beta = -0.34$ ,  $P_{\text{FDRcorrected}} = 1.81\text{E-}08$ ,  $P_{\text{Egger Intercept}} = 1.04\text{E-}01$ ;  $\beta = 0.21$ ,  $P_{\text{FDRcorrected}} = 2.83\text{E-}05$ ,  $P_{\text{Egger Intercept}} = 1.02\text{E-}01$ ;  $\beta = 0.12$ ,  $P_{\text{FDRcorrected}} = 1.73\text{E-}03$ ,  $P_{\text{Egger Intercept}} = 1.09\text{E-}01$ , respectively).

---

**Table 3. Conditional density and parameters for each node of the constructed average Bayesian network using UK biobank inverse normal transformed data (n=3641). SE: standard error.**

---

**SAT | BMI**

Coefficients: (Intercept)=-2.892e-09, BMI=7.603e-01, SE=0.649

---



---

**VAT | BMI + Bilirubin**

Coefficients: (Intercept)=-1.265e-07, BMI=7.221e-01, Bilirubin=1.101e-01, SE=0.688

---

**LiverFat | SAT + VAT + Waist + ALT + Albumin**

Coefficients: (Intercept)=4.094e-05, SAT=2.049e-01, VAT=6.234e-01, Waist=-1.585e-01, ALT=1.122e-01, Albumin=6.906e-02, SE=0.734

---

**Bodyfat Mass | SAT + VAT + DBP + BMI + Waist + HbA1c + Bilirubin**

Coefficients: (Intercept)=-0.00002, SAT=0.671, VAT=-0.246, DBP=0.040, BMI=0.119, Waist=0.376, HbA1c= 0.043, Bilirubin=-0.032, SE=0.472

---

**DBP | VAT + Waist + Albumin + GGTP**

Coefficients: (Intercept)=3.306e-05, VAT=1.240e-01, Waist=1.854e-01, Albumin=1.001e-01, GGTP=8.225e-02, SE=0.934

---

**SBP | SAT + DBP + AST + GGTP + HbA1c**

Coefficients: (Intercept)=2.925e-06, SAT=-7.294e-02, DBP=6.818e-01, AST=4.260e-02, GGTP=6.249e-02, HbA1c=8.602e-02, SE=0.702

---

**BMI**

Coefficients: (Intercept)=1.084e-09, SE=0.999

---

**Waist | SAT + BMI**

Coefficients: (Intercept)=0.0001, SAT=-0.260, BMI=0.941, SE=0.646

---

**ALT | SAT + VAT + Waist + Albumin**

Coefficients: (Intercept)=-3.533e-05, SAT=-1.708e-01, VAT=3.211e-01, Waist=2.731e-01, Albumin=1.401e-01, SE=0.852

---

**Albumin**

Coefficients: (Intercept)=6.566e-08, SE=0.999

---

**AST | SAT + ALT + Bilirubin**

Coefficients: (Intercept)=1.945e-05, SAT=-1.266e-01, ALT=7.208e-01, Bilirubin=5.342e-02, SE=0.683

---

**GGTP | SAT + VAT + Waist + ALT + Albumin**

Coefficients: (Intercept)=-3.640e-06, SAT=-1.026e-01, VAT=1.648e-01, Waist=1.099e-01, ALT=4.966e-01, Albumin=8.686e-02, SE=0.751

---

**Glucose | SBP + GGTP**

Coefficients: (Intercept)=-7.250e-06, SBP=1.468e-01, GGTP=1.017e-01, SE=0.980

---

**HbA1c | LiverFat + Albumin + GGTP + Bilirubin**

Coefficients: (Intercept)=2.546e-06, LiverFat=1.243e-01, Albumin=-6.173e-02, GGTP=1.419e-01, Bilirubin=-1.643e-01, SE=0.962

---

**HDL | SAT + VAT + Bodyfat Mass + SBP + BMI + Waist + ALT + Albumin + GGTP**

Coefficients: (Intercept)=3.837e-05, SAT=1.724e-01, VAT=-2.712e-01, Bodyfat Mass=9.114e-02, SBP=7.267e-02, BMI=1.069e-01, Waist=-3.377e-01, ALT=-1.094e-01, Albumin=6.870e-02, GGTP=9.059e-02, SE=0.830

---

**Bilirubin | SAT + Waist + Albumin**

Coefficients: (Intercept)=-2.523e-05, SAT=-2.555e-01, Waist=2.027e-01, Albumin=1.800e-01, SE=0.949

---

**TG | VAT + LiverFat + Bodyfat Mass + SBP + Waist + Albumin + GGTP + HDL + Bilirubin**

Coefficients: (Intercept)=-1.456e-05, VAT=4.584e-02, LiverFat=1.662e-01, Bodyfat Mass=5.114e-02, SBP=4.485e-02, Waist=5.891e-02, Albumin=5.935e-02, GGTP=1.755e-01, HDL=-3.164e-01, Bilirubin=-1.116e-01, SE=0.803

---

ALT, alanine transaminase; AST, aspartate transaminase; BMI, body mass index; DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucose, fasting glucose from venous plasma samples; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples;

SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; VAT, visceral adipose tissue.

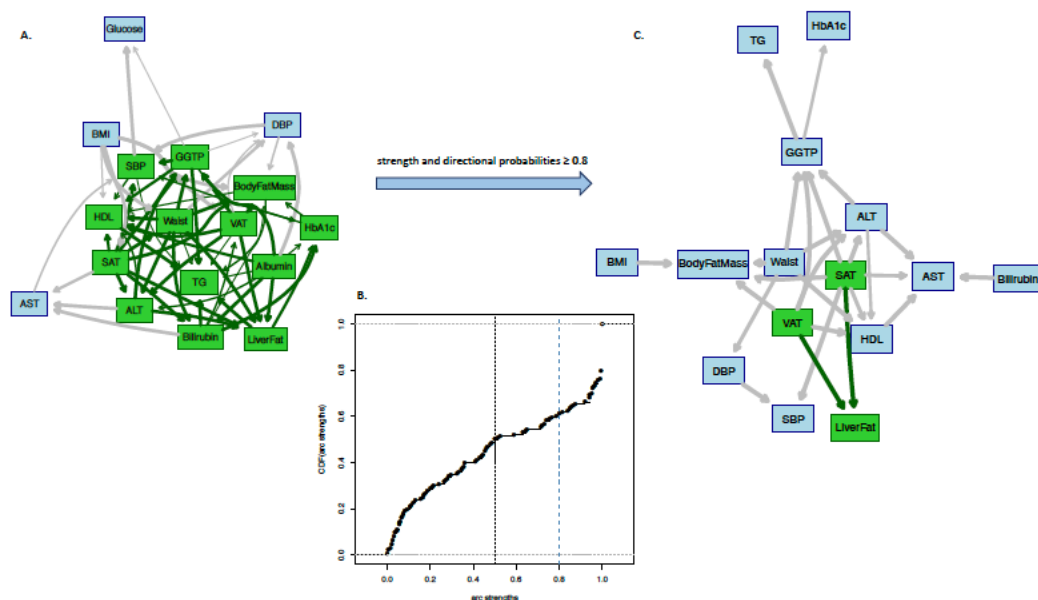


Fig4. **Panel A:** Averaged Bayesian network of the bootstrapped samples among the variables of UK Biobank (data are inverse normal transformed,  $n=3641$ ). **Panel B:** Cumulative distribution function of the arc strengths presented in the Bayesian network of Panel A. Black vertical line is the threshold (0.505) learned from data, equal or above which the arcs are identified as statistically significant. **Panel C:** Subset of Bayesian networks in Panel A including only arcs with strength and directional probabilities  $\geq 0.8$ . Nodes in green highlight the Markov blanket of liver fat that includes the nodes with adequate information to stand as a separate Bayesian network.

ALT, alanine transaminase; AST, aspartate transaminase; BMI, body mass index; DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucose, fasting glucose from venous plasma samples; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; VAT, visceral adipose tissue.

#### *Subgroup analysis: diabetes vs. non-diabetes - female vs. male (in UK Biobank)*

As with the IMI DIRECT analyses reported above, we built BNs separately in participants diagnosed with T2D (S7 Fig, lower panel) and without diabetes (S7 Fig, upper panel). VAT, SAT and albumin were identified as the parental nodes of liver fat in both groups; ALT and waist circumference were only present as the parents in the group without T2D (S7 Fig panels A and C). The subset of these BNs with arcs having strengths and directional probabilities  $\geq 0.8$  left only VAT as the causal node for liver fat in the group with T2D and VAT and SAT in the group without T2D (S7 Fig panels B and D). We then continued with the sensitivity analysis

by gender, as illustrated in S8 Fig. In the female's BN, VAT, TG, ALT, whole-body fat mass, HbA1c and BMI were the causal parental nodes for liver fat, whereas in the male's BN, VAT, whole-body fat mass and Albumin were the causal variables (S8 Fig, panel A and C). These BNs with extra restrictions on the strength and directional probabilities above or equal to 0.8 resulted in only VAT as the causal parental node for liver fat (S8 Fig, panel B and D). The Pearson correlation and the heatmap cluster analyses, performed in the sensitivity analyses of the UK Biobank resulted in similar clusters to the combined cohorts' one (see S6, S9 and S10 Figs). Detailed information of arcs with strength and directional probabilities above or equal to the UK Biobank BNs' thresholds (0.49, 0.5, 0.49 and 0.48 in the non-diabetes, diabetes, female and male BNs, respectively) are reported in S2 Table with their corresponding MR results.

#### *Bayesian networks and the posterior probabilistic inference (UK Biobank)*

Owing to their role as causal parental nodes for liver fat in the UK Biobank BNs, high levels of VAT and SAT (see S3 Table) were investigated for their conditional effects on posterior probabilities of liver fat levels. Fig 5 summarizes unconditional and posterior conditional probabilities of liver fat derived using complete case data, diabetes, non-diabetes, female and male subgroups of the UK Biobank BNs. A statistically significant increase was observed after conditioning on high levels of VAT and SAT in the unconditional probability of high liver fat level on the random observations generated by the BNs from the complete case data, 37% increase after conditioning on VAT, 18% increase after conditioning on SAT, and 41% after conditioning on both. Similarly, diabetes, non-diabetes, female and male groups had substantial increments on the high liver fat level probability after conditioning on high levels of VAT, SAT and the combined effect. The probability of having high liver fat after conditioning on high levels of VAT and SAT was 0.70, 0.97, 0.79, 0.74 and 0.72 in the complete case data, diabetes, non-diabetes, female and male groups of the UK Biobank constructed BNs.

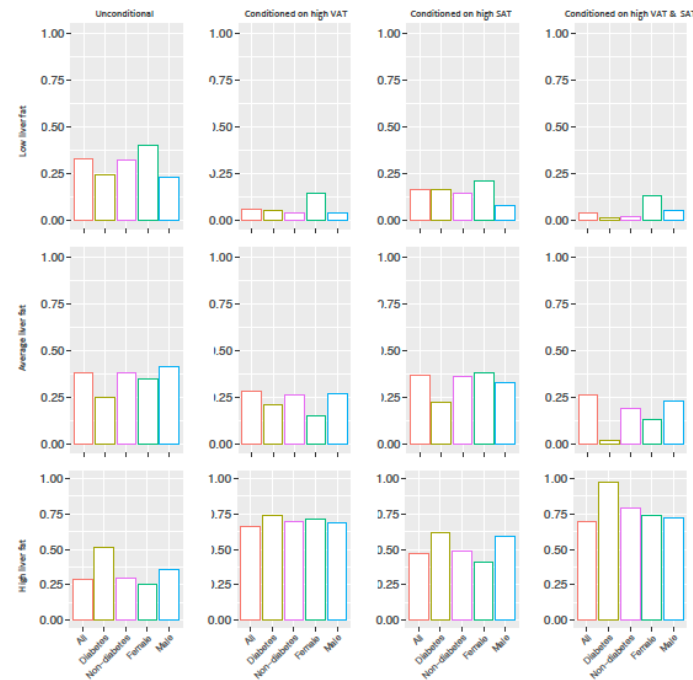


Fig5. Unconditional and posterior conditional probabilities of liver fat on the combined (n=3641), diabetes (n=104), non-diabetes (n=3531), female (n=1921) and male (n=1720) groups of the UK biobank Bayesian networks (data are inverse normal transformed). Y axis represents the unconditional and posterior conditional probabilities of low (0,1.438], average (1.438,3.619] or high (3.619,46.049] levels of liver fat in rows. The posterior probabilities were obtained after conditioning on high levels of visceral adipose tissue (VAT) (5.958,14.408] (figures in 2<sup>nd</sup> column), high levels of subcutaneous adipose tissue (SAT) (7.808,23.476] (figures in 3<sup>rd</sup> column) and high levels of both (figures in 4<sup>th</sup> column). Continuous variables were discretized with Hartemink's method<sup>17</sup>.

## DISCUSSION

We interrogated the putative causal pathways between metabolic processes and liver fat accumulation through Bayesian models and a series of bidirectional MR analyses. BNs fitted to each dataset separately suggested that VAT and BasalISR (only available in IMI DIRECT) causally affect liver fat accumulation. Reassuringly, VAT and BasalISR were also identified as the strongest causal determinants through different subgroup analyses as the closest determinants of liver fat accumulation. In addition, the probabilistic inference analyses showed a substantial increase in the posterior probability of having high levels of liver fat after conditioning on VAT and BasalISR (~30%-40%). Furthermore, the MR analysis suggests that several nominal directional associations between hepatic biomarkers, glycemic and adiposity measures are likely to be causal (summarized in S1 Table).

The epidemiological association of NAFLD and T2D has been widely reported, yet it is unclear if this relationship is causal, and, if so, whether it is uni- or bi-directional, and what the underlying mechanisms might be. MR has been widely applied to epidemiological datasets to help minimize confounding and reverse-causation that are common weaknesses of such studies<sup>18,19</sup>. In a recent study from Liu et.al, the causal relationships between NAFLD, T2D and obesity were explored using a bidirectional 2-sample MR analysis in UK Biobank<sup>9</sup>. Their results showed that genetically-determined NAFLD causes T2D and central obesity, whereas genetically-determined T2D, overall and central obesity cause metabolic NAFLD. Another smaller-scale MR analysis<sup>20</sup> suggested the causal role of hepatic fat accumulation in the development of chronic liver disease and insulin resistance. The hepatic enzymes ALT and AST were found to be causally associated with T2D through bidirectional MR analyses that assessed the causal relationship of NAFLD and T2D<sup>21</sup>, something that was not observed in our analyses. BN is an established causal inference method<sup>22,23</sup>, yet it has rarely been utilized to

study metabolic traits<sup>24</sup>. Within these rare exceptions, a reciprocal association between NAFLD and metabolic syndrome was suggested through a simplified BN applied to a Chinese cohort<sup>25</sup>. Aiming to predict complications of T2D, BN models were built from physiological risk factors of the disease<sup>26</sup>. BN have also been used to study relationships between risk factors for T2D in European-ancestry cohorts<sup>27</sup>.

Through the main BN analyses reported here, BasalISR was the strongest causal driver of liver fat accumulation. This may be explained by hepatic insulin sensitivity (not measured in IMI DIRECT nor UK Biobank) or as a direct effect of BasalISR as shown in a recent study by van Vliet et al.<sup>28</sup> where a clinical association between obesity and basal insulin secretion rate was observed even in the absence of insulin resistance<sup>28</sup>. Our analyses also showed that in people with diabetes, although BasalISR was an important factor, VAT had a bigger causal impact on liver fat accumulation (Fig 2- compare 2<sup>nd</sup> and 3<sup>rd</sup> figures in the last row). In the UK Biobank analyses, where BasalISR was unavailable, VAT was the strongest determinant of fatty liver. This, combined with findings in the diabetes sub-cohort of IMI DIRECT, suggests that VAT is likely to be a strong causal determinant of fatty liver, especially when diabetes has progressed, and insulin secretion is diminished. Our findings were also consistent with other studies that show that excess visceral adiposity is an important predictor of NAFLD, much more so than subcutaneous adiposity<sup>29</sup> (Fig 5- compare 2<sup>nd</sup> and 3<sup>rd</sup> figures in the last row).

In addition to using BNs for causal inference, we used MR to validate the factors involved in the metabolic processes and liver fat pathways. We found that glycemic traits were at least nominally associated with liver enzymes (i.e., AST and GGTP), but the reciprocal effect was not detected; however, AST was causally related with insulin sensitivity, though this was not detected in pleiotropy/outlier sensitivity analyses. Although there are no robust genetic

instruments for VAT and SAT<sup>30</sup>, which would be necessary for an MR analysis assessing these traits, this could be done for whole-body fat mass. Here, the MR analysis supports a causal association of whole-body fat mass with insulin sensitivity, consistent with the findings from the BNs. Moreover, larger waist circumference and higher BMI were causally associated with fasting insulin levels, yet there was no evidence of an effect of these traits on insulin secretion.

In a recent analysis from IMI DIRECT consortium<sup>31</sup>, we used structural equation modelling to test the hypothesis that physical activity affects glucose regulation in a manner consistent with the Twin-cycle hypothesis<sup>6,7</sup>. Here, we constructed a BN that was agnostic to the hypothesized Twin-cycle model and showed good, albeit not complete, agreement between the two models, providing further validation of the Twin-cycle hypothesis. Note that BN is acyclical, thus preventing complete recapitulation of the Twin cycle hypothesis using this approach. Moreover, instead of hepatic insulin sensitivity (as used in the Twin-cycle model), OGIS, the insulin sensitivity index, was used as a proxy, which might explain some of the observed differences. Furthermore, there are currently no genetic instruments available that adequately characterize pancreatic fat<sup>32</sup>, thus preventing its inclusion in our MR analyses.

Although BN and MR have key limitations, they are largely orthogonal methods and their combination provides a fairly robust causal inference framework. To avoid any overlap in the two samples utilized in MR analysis (exposure and outcome), we restricted the analyses primarily to the non-UK Biobank repositories. However, when we had both exposure and outcome derived from the same repositories, we considered UK Biobank for one of them. Low statistical power can also be a problem in MR analyses, owing to the conditional nature of the calculations performed. Specifically, MR depends on the proportion of variance in a phenotypic trait that is explained by the instrumental variables (SNPs) used in the GWAS; thus, a limited

number of instrumental variables is unlikely to capture a true effect for complex exposures. Moreover, established methodological caveats of MR, such as pleiotropy and heterogeneity, were assessed to effectively disentangle the effect of each trait; however, we cannot completely rule out that our instrumental variables have other phenotypic effects. We adopted significance criteria with a conservative p-value threshold; the concordance with two methods, and no statistical evidence of pleiotropy. Within the same context, utilizing BNs for causal inference needs careful consideration, the assumption being there are no latent variables in the model. Here, we focused on probabilistic inference and computed the posterior probability of liver fat from the observed BNs after conditioning on different pieces of evidence (high levels of BasalISR, VAT and/or SAT). It is noteworthy that the constructed BNs should be considered as probable causal models, since they are derived from available data and knowledge.

Another limitation of our approach was that to learn the structure and parameters of the BNs, a complete case analysis was required, which diminished the sample sizes both in UK Biobank and IMI DIRECT considerably. Amongst those from UK Biobank, liver iron and liver inflammation factors were two interesting variables that we elected not to include in the BN analyses, due to high missing rates (~70%). However, as sensitivity analyses, we tried building BN in the subset including liver iron and liver inflammation factor, and it had a downstream association from liver fat, similar to the IMI DIRECT network. Subsetting the UK Biobank datasets to only those diagnosed with T2D left very few individuals for the complete case analyses. As such, the BN model for the diabetic subset is sparse and possibly underpowered. In a recent study, Scutari reviewed how BNs can model data with incomplete observations<sup>33</sup>, which can be a focus for future research. Multicollinearity, where input variables are highly correlated, can be an issue in some pathway analyses. However, the type of structural learning in a BN makes it robust against such limitation. This property makes BN an intuitive visual tool



to express associations among several variables in a complex network. We did not apply any force node in building the BNs and they were built unsupervised. However, semi-supervised networks feeding with known directional associations (e.g., from previous MR analyses) can be considered as future work. Expansion of the Bayesian workspace to omics<sup>34</sup> would be a logical extension of the current work, especially considering the multi-omics data (genetic, transcriptomic, proteomic and metabolomic) available in IMI DIRECT consortium.

In summary, the combination of BN and MR analyses deployed here provided a powerful causal inference framework through which a causal model of metabolic homeostasis in fatty liver disease could be developed and validated. The key findings are that insulin secretion rate and visceral adiposity are major causal drivers of fatty liver, until, at least, diabetes develops, at which point visceral adiposity becomes the dominant causal agent. These findings may aid the targeted prevention of fatty liver disease.

## **MATERIALS AND METHODS**

### *IMI DIRECT cohorts and measures*

We utilized data from the IMI DIRECT consortium, prospective cohort studies of 795 adults with T2D and 2234 without the disease. Written informed consent was provided by all participants at enrollment and the regional research ethics committees of each clinical study center have approved the study protocol separately<sup>35,36</sup>. The participants were extensively assessed with measures focused on glycemia, insulin dynamics, organ-specific adiposity, serological biomarkers, lifestyle, anthropometry and other clinical features. Frequently sampled mixed-meal tolerance tests (MMTTs) and 75 g frequently sampled oral glucose tolerance tests (fsOGTTs) were carried out in the diabetes and non-diabetes cohorts, respectively. Basal insulin secretion rate (BasalISR) was estimated as the product of insulin clearance and fasting plasma

insulin as defined elsewhere<sup>37</sup>. Adiposity, including liver fat, pancreas fat, visceral fat and subcutaneous fat, was assessed using a T2\*-based multiecho technique from MRI scans<sup>38</sup>. An overview of participant characteristics is shown in S3 Table in the non-diabetes, diabetes and combined cohorts of the IMI DIRECT. More details of the study design and the core characteristics are provided elsewhere<sup>35,36</sup>.

#### *UK Biobank cohort and measures*

UK Biobank is a prospective cohort study with more than 500,000 adults aged 37–73 years and recruited between 2006 and 2010<sup>39</sup>. The current analysis was conducted using data obtained via the UK Biobank Access Application process (project number 18274). We used this dataset for both modelling and replication of the primary findings reported herein from the IMI DIRECT cohorts. UK Biobank and IMI DIRECT have the same protocol and procedure for quantification of the MRI-derived abdominal fat<sup>40</sup>. The field numbers for the UK Biobank variables and an overview of UK Biobank participants' characteristics are shown in S3 Table.

#### *Bayesian Network analysis*

We utilized BNs to build the graphical models for inferring causal pathways. BNs are probabilistic graphical models that are built from joint probability distributions of random variables using Bayesian inference. The network *structure* is depicted graphically by directed acyclic graphs (DAGs), where nodes represent the random variables with directed arcs affirming the conditional dependencies and missing arcs defining the conditional independencies in the model<sup>41-43</sup>.

Building a BN can be summarized into a two-step learning process (equation 1); structural learning where the network structure is learned from data and parameter learning where local distributions and parameters are inferred from the learned structure in the first step:

$$p(M|D) = p(G, \theta|D) = p(G|D).p(\theta|G, D) \quad (1)$$

Where  $p(M|D)$  is the likelihood of the BN model according to the structure  $G$  and parameter  $\theta$  given data  $D$ . The structural learning in our analyses has been undertaken by the *score-based* approach, as it deals better with the small sample size and the possible noise in the data<sup>44</sup>. The heuristic search using the hill-climbing technique<sup>45</sup> was applied, searching towards the structure with the maximum goodness of fit score. Scoring of the BN structures was done using the *Bayesian Information Criterion* (BIC) score method as follows:

$$BIC = \log \left( p(D|\hat{\theta}, G) \right) - \frac{n_p}{2} \log(N)$$

where,  $p(D|\theta, G)$  is the probability of data given parameters and the graph,  $n_p$  is the number of parameters with  $N$  denoting the sample size. We elected the BIC method as it penalizes networks with many arcs, leading to a simpler graph with fewer false positives<sup>45</sup>.

Numeric variables were all inverse normalized to fit a Gaussian BN with a multivariate normal distribution<sup>46,47</sup>. Such networks require that the nodes are normally distributed to function properly. Furthermore, to attain a stable structure from data, which is resistance to the network perturbation, we performed model averaging. As such, data were resampled using bootstrapping, following which a separate structure was built for each bootstrap sample, and the averaged network was then constructed based on the frequency of the possible arcs amongst all networks.

The *Markov condition* is an assumption that must be satisfied when making the BNs, which states a node can be conditionally independent of the whole network given its Markov blanket (the set of nodes that includes all the information required for separating it from the rest of graph)<sup>48</sup>. In our networks, the Markov blanket of the liver fat node, which includes the upstream (parent) nodes, the downstream (children) nodes and other parents of those children, were derived. After learning the structure of the network, the parameters were fitted using their maximum likelihood estimate.

To further investigate a constructed BN after its structure ( $G$ ) and parameter ( $\theta$ ) learning, *queries* on the variable of interest ( $X$ ), known as *event*, can be defined. Here, we focused on liver fat as our event and computed its posterior probability after conditioning on an *evidence* ( $E$ ) to further understand its causal effect on liver fat,  $p(X|E, G, \theta)$ . To compute the conditional posterior probabilities, variables were required to be discrete. To discretize the continuous variables, Hartemink's method was deployed which helps to preserve the initial dependencies among the variables while defining the intervals<sup>17</sup>.

### *Mendelian randomization analysis*

The MR approach has gained popularity because it is less susceptible to confounding and reverse causation than most other observational studies<sup>49,50</sup>, and is often likened to a randomized controlled trial (RCT) owing to the random allocation of most confounding variables across the levels of the exposure instrument (genotype). Specifically, the random assortment of alleles during meiosis allows exploration of causal relationships between various traits and diseases under the assumption that genetic variants [known as instrumental variables (IV)] are suitable proxies for given risk factors, moreover, the instruments should influence the outcome only through the exposure and not through any other confounding<sup>18</sup>. Bidirectional MR involves assessment of the reciprocal exposure-outcome relationship to assess the directionality

in causation. Thus, a systematic bidirectional MR study leveraging the latest public GWAS data would help triangulate causality of directed arcs in BN analysis.

Two-sample MR analysis was deployed, where each trait was investigated for potential causal associations. Between the selected metabolic traits, we assessed bidirectionally, i.e. used as exposure and as outcome<sup>50</sup>, we leveraged from the latest summary statistics from publicly available sources i.e. GWAS catalog<sup>51</sup> and MR-CIEU<sup>52</sup>. Operationally, not all traits defined in IMI DIRECT or UK Biobank have a corresponding GWAS published. We prioritized GWAS restricted to European population (see S1 Table) based on i) latest release, ii) sample size, and iii) if no published GWAS was obtained or not enough number of independent variants were identified we used the UK Biobank data<sup>53</sup> from Neale Lab (<http://www.nealelab.is/blog/2017/9/11/details-and-considerations-of-the-uk-biobank-gwas>).

Additionally, in case of sample overlap between two datasets, we preemptively selected one trait from UK Biobank and overlap was calculated with the maximum sample-overlapping rate<sup>55</sup>, as reported by Liu and colleagues<sup>9</sup>. Instrumental variables were prioritized at GWAS-significance threshold ( $p\text{-value} < 5 \times 10^{-8}$ ) and proxies were used if the genetic variants were in linkage disequilibrium (LD) at  $r^2 \geq 0.8$  in any of the two-samples, yet, to avoid LD within the instruments, we performed LD-clumping restricted to  $r^2 < 0.2$ , a 1000 kb window for the final sets.

We used the inverse variance weighted (IVW) method as the main analysis to estimate the effects of the instrumental variables, when there were sufficient number of genetic variants. Furthermore, we used MR-Egger method to minimize false positive associations. To quantify heterogeneity, horizontal pleiotropy, and detect outliers, we used the MR-Egger intercept, the Q statistic and the MR Pleiotropy Residual Sum and Outlier (MR-PRESSO) global test at p

level of  $>0.05$  when appropriate<sup>55</sup>. MR-Egger, as part of the regression, provides a consistent estimate of the causal effect. Moreover, when the InSIDE (Instrument Strength Independent of Direct Effect) assumption holds, the intercept provides an unbiased average pleiotropic effect that should not significantly differ from the null, yet, when this is unmet, its value represents an estimate of the directional horizontal pleiotropy<sup>56</sup>. We considered MR findings to be statistically significant if: i) the causal association amongst IVW, and Egger were directionally concordant, ii) the IVW (main method) passed the FDR corrected<sup>13</sup> threshold ( $p < 0.05$ ) for multiple-testing, and iii) no statistical evidence of heterogeneity and/or pleiotropy (PEgger Intercept, and PMR-PRESSO Global  $p > 0.05$ ).

All the statistical analyses were undertaken using R software version 3.6.2<sup>57</sup>, the BNs were built using the *bnlearn* package<sup>58,59</sup> and the MR analysis using the *TwoSampleMR*<sup>52</sup> and MR-PRESSO packages<sup>55</sup>. Variables from IMI DIRECT were adjusted for center effect in a linear model including each variable per model. The residuals were then extracted from these models and were ranked normalized to fulfill the Gaussian BN assumption. The distribution of the IMI DIRECT and UK Biobank variables prior to transformation is depicted in S11 and S12 Figs, respectively.

## REFERENCES

- 1 Younossi, Z. *et al.* Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. *Nature reviews Gastroenterology & hepatology* **15**, 11-20 (2018).
- 2 Ludwig, J., Viggiano, T. R., McGill, D. B. & Oh, B. Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clinic Proceedings* **55**, 434-438 (1980).
- 3 Mantovani, A., Byrne, C. D., Bonora, E. & Targher, G. Nonalcoholic fatty liver disease and risk of incident type 2 diabetes: a meta-analysis. *Diabetes care* **41**, 372-382 (2018).
- 4 Schwimmer, J. B. *et al.* Heritability of nonalcoholic fatty liver disease. *Gastroenterology* **136**, 1585-1592 (2009).

- 5 Lonardo, A., Leoni, S., Alswat, K. A. & Fouad, Y. History of nonalcoholic fatty liver disease. *International Journal of Molecular Sciences* **21**, 5888 (2020).
- 6 Taylor, R. Type 2 diabetes: etiology and reversibility. *Diabetes care* **36**, 1047-1055 (2013).
- 7 Taylor, R. Pathogenesis of type 2 diabetes: tracing the reverse route from cure to cause. *Diabetologia* **51**, 1781-1789 (2008).
- 8 Utzschneider, K. M. & Kahn, S. E. The role of insulin resistance in nonalcoholic fatty liver disease. *The Journal of Clinical Endocrinology & Metabolism* **91**, 4753-4761 (2006).
- 9 Liu, Z. *et al.* Causal relationships between NAFLD, T2D and obesity have implications for disease subphenotyping. *Journal of hepatology* **73**, 263-276 (2020).
- 10 Dongiovanni, P. & Valenti, L. A nutrigenomic approach to non-alcoholic fatty liver disease. *International journal of molecular sciences* **18**, 1534 (2017).
- 11 Parisinos, C. A. *et al.* Genome-wide and Mendelian randomisation studies of liver MRI yield insights into the pathogenesis of steatohepatitis. *Journal of Hepatology* (2020).
- 12 Byrne, C. D. & Targher, G. NAFLD: a multisystem disease. *Journal of hepatology* **62**, S47-S64 (2015).
- 13 Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* **57**, 289-300 (1995).
- 14 Atabaki-Pasdar, N. *et al.* Predicting and elucidating the etiology of fatty liver disease: A machine learning modeling and validation study in the IMI DIRECT cohorts. *PLoS medicine* **17**, e1003149 (2020).
- 15 Leite, N. C., Villela-Nogueira, C. A., Cardoso, C. R. & Salles, G. F. Non-alcoholic fatty liver disease and diabetes: from physiopathological interplay to diagnosis and treatment. *World journal of gastroenterology: WJG* **20**, 8377 (2014).
- 16 Mari, A., Pacini, G., Murphy, E., Ludvik, B. & Nolan, J. J. A model-based method for assessing insulin sensitivity from the oral glucose tolerance test. *Diabetes care* **24**, 539-548 (2001).
- 17 Hartemink, A. J. *Principled computational methods for the validation discovery of genetic regulatory networks*, Massachusetts Institute of Technology, (2001).
- 18 Lawlor, D. A., Harbord, R. M., Sterne, J. A., Timpson, N. & Davey Smith, G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* **27**, 1133-1163, doi:10.1002/sim.3034 (2008).
- 19 Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* **23**, R89-98, doi:10.1093/hmg/ddu328 (2014).
- 20 Dongiovanni, P. *et al.* Causal relationship of hepatic fat with liver damage and insulin resistance in nonalcoholic fatty liver. *Journal of internal medicine* **283**, 356-370 (2018).
- 21 De Silva, N. M. G. *et al.* Liver function and risk of type 2 diabetes: bidirectional mendelian randomization study. *Diabetes* **68**, 1681-1691 (2019).
- 22 Nadkarni, S. & Shenoy, P. P. A Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research* **128**, 479-498 (2001).
- 23 Heckerman, D. Bayesian networks for data mining. *Data mining and knowledge discovery* **1**, 79-119 (1997).
- 24 Franks, P. W. & Atabaki-Pasdar, N. Causal inference in obesity research. *Journal of internal medicine* **281**, 222-232 (2017).
- 25 Zhang, Y. *et al.* Identification of reciprocal causality between non-alcoholic fatty liver disease and metabolic syndrome by a simplified Bayesian network in a Chinese population. *BMJ open* **5** (2015).



- 26 Liu, S., Zhang, R., Shang, X. & Li, W. Analysis for warning factors of type 2 diabetes mellitus complications with Markov blanket based on a Bayesian network model. *Computer Methods and Programs in Biomedicine* **188**, 105302 (2020).
- 27 Sambo, F. *et al.* in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2119-2122 (IEEE).
- 28 van Vliet, S. *et al.* Obesity is associated with increased basal and postprandial  $\beta$ -cell insulin secretion even in the absence of insulin resistance. *Diabetes* **69**, 2112-2119 (2020).
- 29 Radmard, A. R. *et al.* Assessment of abdominal fat distribution in non-alcoholic fatty liver disease by magnetic resonance imaging: a population-based study. *Archives of Iranian medicine* **19**, 0-0 (2016).
- 30 Cha, E. D. *et al.* Using Adipose Measures from Health Care Provider-Based Imaging Data for Discovery. *Journal of obesity* **2018** (2018).
- 31 Koivula, R. W. *et al.* The role of physical activity in metabolic homeostasis before and after the onset of type 2 diabetes: an IMI DIRECT study. *Diabetologia*, 1-13 (2020).
- 32 Pinnick, K. E. *et al.* Pancreatic ectopic fat is characterized by adipocyte infiltration and altered lipid composition. *Obesity* **16**, 522-530 (2008).
- 33 Scutari, M. Bayesian network models for incomplete and dynamic data. *Statistica Neerlandica* (2020).
- 34 Pourret, O., Naïm, P. & Marcot, B. *Bayesian networks: a practical guide to applications*. (John Wiley & Sons, 2008).
- 35 Koivula, R. W. *et al.* Discovery of biomarkers for glycaemic deterioration before and after the onset of type 2 diabetes: descriptive characteristics of the epidemiological studies within the IMI DIRECT Consortium. *Diabetologia* **62**, 1601-1615 (2019).
- 36 Koivula, R. W. *et al.* Discovery of biomarkers for glycaemic deterioration before and after the onset of type 2 diabetes: rationale and design of the epidemiological studies within the IMI DIRECT Consortium. *Diabetologia* **57**, 1132-1142, doi:10.1007/s00125-014-3216-x (2014).
- 37 Mari, A., Tura, A., Gastaldelli, A. & Ferrannini, E. Assessing insulin secretion by modeling in multiple-meal tests: role of potentiation. *Diabetes* **51**, S221-S226 (2002).
- 38 Thomas, E. L., Fitzpatrick, J., Malik, S., Taylor-Robinson, S. D. & Bell, J. D. Whole body fat: content and distribution. *Progress in nuclear magnetic resonance spectroscopy* **73**, 56-80 (2013).
- 39 Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *Plos med* **12**, e1001779 (2015).
- 40 Wilman, H. R. *et al.* Characterisation of liver fat in the UK Biobank cohort. *PloS one* **12**, e0172921 (2017).
- 41 Pearl, J. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. (Elsevier, 2014).
- 42 Korb, K. B. & Nicholson, A. E. *Bayesian artificial intelligence*. (CRC press, 2010).
- 43 Koller, D. & Friedman, N. *Probabilistic graphical models: principles and techniques*. (MIT press, 2009).
- 44 Scutari, M., Graafland, C. E. & Gutiérrez, J. M. in *International Conference on Probabilistic Graphical Models*. 416-427.
- 45 Tsamardinos, I., Brown, L. E. & Aliferis, C. F. The max-min hill-climbing Bayesian network structure learning algorithm. *Machine learning* **65**, 31-78 (2006).
- 46 Su, C., Andrew, A., Karagas, M. R. & Borsuk, M. E. Using Bayesian networks to discover relations between genes, environment, and disease. *BioData mining* **6**, 6 (2013).



- 47 Hofmann, R. & Tresp, V. in *Advances in neural information processing systems*. 500-506.
- 48 Statnikov, A., Lemeir, J. & Aliferis, C. F. Algorithms for discovery of multiple Markov boundaries. *The Journal of Machine Learning Research* **14**, 499-566 (2013).
- 49 VanderWeele, T. J., Tchetgen Tchetgen, E. J., Cornelis, M. & Kraft, P. Methodological challenges in mendelian randomization. *Epidemiology (Cambridge, Mass.)* **25**, 427-435, doi:10.1097/EDE.0000000000000081 (2014).
- 50 Pierce, B. L. & Burgess, S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *American journal of epidemiology* **178**, 1177-1184 (2013).
- 51 MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic acids research* **45**, D896-D901 (2017).
- 52 Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**, doi:10.7554/eLife.34408 (2018).
- 53 Millard, L. A., Davies, N. M., Gaunt, T. R., Davey Smith, G. & Tilling, K. Software Application Profile: PHESANT: a tool for performing automated phenome scans in UK Biobank. *International Journal of Epidemiology* (2017).
- 54 Burgess, S., Davies, N. M. & Thompson, S. G. Bias due to participant overlap in two-sample Mendelian randomization. *Genetic epidemiology* **40**, 597-608 (2016).
- 55 Verbanck, M., Chen, C.-y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nature genetics* **50**, 693-698 (2018).
- 56 Burgess, S., Bowden, J., Fall, T., Ingelsson, E. & Thompson, S. G. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology (Cambridge, Mass.)* **28**, 30-42, doi:10.1097/EDE.0000000000000559 (2017).
- 57 Team, R. C. R: A Language and Environment for Statistical Computing. 887 R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> **888** (2017).
- 58 Scutari, M. & Denis, J.-B. *Bayesian networks: with examples in R*. (CRC press, 2014).
- 59 Scutari, M. Learning Bayesian networks with the bnlearn R package. *arXiv preprint arXiv:0908.3817* (2009).

## SUPPLEMENTARY INFORMATION

**S1 Fig.** A. Heatmap cluster among the variables of IMI DIRECT combined cohorts (data are inverse normal transformed, n=1070). B. The graphical connection of the variables with Pearson correlation equal or above 0.4.

OGIS, oral glucose insulin sensitivity index according to the method of Mari *et al.* <sup>16</sup>; Clins, mean insulin clearance during the oral glucose tolerance test/mixed meal tolerance test, calculated as (mean insulin secretion)/(mean insulin concentration); HDL, fasting high-density lipoprotein cholesterol; GlucoseSens, glucose sensitivity, slope of the dose-response relating insulin secretion to glucose concentration; PancIron, pancreas iron; PancFat, pancreas fat; SBP, mean systolic blood pressure; DBP, mean diastolic blood pressure; Glucagonmin0, fasting glucagon concentration; TotGLP1min0, concentration of fasting total GLP-1 in plasma; HbA1c, glycated hemoglobin A1C; Glucose, fasting glucose from venous plasma samples; TwoGlucose, 2-hour glucose after oral glucose tolerance test/mixed-meal tolerance test; GGTP, gamma-glutamyl transpeptidase; AST, aspartate transaminase; ALT, alanine transaminase; SAT, subcutaneous adipose tissue; BMI, body mass index; TG, fasting triglycerides; VAT,

visceral adipose tissue; TwoInsulin, 2-hour insulin; Insulin, fasting insulin from venous plasma samples; BasalISR, insulin secretion at the beginning of the oral glucose tolerance test/ mixed-meal tolerance test.

**S2 Fig.** Averaged Bayesian network of the bootstrapped samples among the variables of IMI DIRECT cohorts, panel A-B for non-diabetic (n=787) and panel C-D for diabetic (n=283) cohorts. Data are inverse normal transformed and the Bayesian networks in panels B and D show only the arcs with strength and directional probabilities  $\geq 0.8$ .

OGIS, oral glucose insulin sensitivity index according to the method of Mari et al.<sup>16</sup>; Clins, mean insulin clearance during the oral glucose tolerance test/mixed meal tolerance test, calculated as (mean insulin secretion)/(mean insulin concentration); HDL, fasting high-density lipoprotein cholesterol; GlucoseSens, glucose sensitivity, slope of the dose–response relating insulin secretion to glucose concentration; PancIron, pancreas iron; PancFat, pancreas fat; SBP, mean systolic blood pressure; DBP, mean diastolic blood pressure; Glucagonmin0, fasting glucagon concentration; TotGLP1min0, concentration of fasting total GLP-1 in plasma; HbA1c, glycated hemoglobin A1C; Glucose, fasting glucose from venous plasma samples; TwoGlucose, 2-hour glucose after oral glucose tolerance test/mixed-meal tolerance test; GGTP, gamma-glutamyl transpeptidase; AST, aspartate transaminase; ALT, alanine transaminase; SAT, subcutaneous adipose tissue; BMI, body mass index; TG, fasting triglycerides; VAT, visceral adipose tissue; TwoInsulin, 2-hour insulin; Insulin, fasting insulin from venous plasma samples; BasalISR, insulin secretion at the beginning of the oral glucose tolerance test/ mixed-meal tolerance test.

**S3 Fig.** Averaged Bayesian network of the bootstrapped samples among the variables of IMI DIRECT cohorts, panel A-B for female (n=293) and panel C-D for male (n=831) groups. Data are inverse normal transformed and the Bayesian networks in panels B and D show only the arcs with strength and directional probabilities  $\geq 0.8$ .

OGIS, oral glucose insulin sensitivity index according to the method of Mari et al.<sup>16</sup>; Clins, mean insulin clearance during the oral glucose tolerance test/mixed meal tolerance test, calculated as (mean insulin secretion)/(mean insulin concentration); HDL, fasting high-density lipoprotein cholesterol; GlucoseSens, glucose sensitivity, slope of the dose–response relating insulin secretion to glucose concentration; PancIron, pancreas iron; PancFat, pancreas fat; SBP, mean systolic blood pressure; DBP, mean diastolic blood pressure; Glucagonmin0, fasting glucagon concentration; TotGLP1min0, concentration of fasting total GLP-1 in plasma; HbA1c, glycated hemoglobin A1C; Glucose, fasting glucose from venous plasma samples; TwoGlucose, 2-hour glucose after oral glucose tolerance test/mixed-meal tolerance test; GGTP, gamma-glutamyl transpeptidase; AST, aspartate transaminase; ALT, alanine transaminase; SAT, subcutaneous adipose tissue; BMI, body mass index; TG, fasting triglycerides; VAT, visceral adipose tissue; TwoInsulin, 2-hour insulin; Insulin, fasting insulin from venous plasma samples; BasalISR, insulin secretion at the beginning of the oral glucose tolerance test/ mixed-meal tolerance test.

**S4 Fig.** Heatmap cluster and the graphical connection of the IMI DIRECT variables with Pearson correlation equal or above 0.4, panel A-B for non-diabetic (n=787) and panel C-D for diabetic (n=283) cohorts. Data are inverse normal transformed.

OGIS, oral glucose insulin sensitivity index according to the method of Mari et al.<sup>16</sup>; Clins, mean insulin clearance during the oral glucose tolerance test/mixed meal tolerance test, calculated as (mean insulin secretion)/(mean insulin concentration); HDL, fasting high-density lipoprotein cholesterol; GlucoseSens, glucose sensitivity, slope of the dose–response relating

insulin secretion to glucose concentration; PancIron, pancreas iron; PancFat, pancreas fat; SBP, mean systolic blood pressure; DBP, mean diastolic blood pressure; Glucagonmin0, fasting glucagon concentration; TotGLP1min0, concentration of fasting total GLP-1 in plasma; HbA1c, glycated hemoglobin A1C; Glucose, fasting glucose from venous plasma samples; TwoGlucose, 2-hour glucose after oral glucose tolerance test/mixed-meal tolerance test; GGTP, gamma-glutamyl transpeptidase; AST, aspartate transaminase; ALT, alanine transaminase; SAT, subcutaneous adipose tissue; BMI, body mass index; TG, fasting triglycerides; VAT, visceral adipose tissue; TwoInsulin, 2-hour insulin; Insulin, fasting insulin from venous plasma samples; BasalISR, insulin secretion at the beginning of the oral glucose tolerance test/ mixed-meal tolerance test.

**S5 Fig.** Heatmap cluster and the graphical connection of the IMI DIRECT variables with Pearson correlation equal or above 0.4, panel A-B for female (n=293) and panel C-D for male (n=831) groups. Data are inverse normal transformed.

OGIS, oral glucose insulin sensitivity index according to the method of Mari et al.<sup>16</sup>; Clins, mean insulin clearance during the oral glucose tolerance test/mixed meal tolerance test, calculated as (mean insulin secretion)/(mean insulin concentration); HDL, fasting high-density lipoprotein cholesterol; GlucoseSens, glucose sensitivity, slope of the dose–response relating insulin secretion to glucose concentration; PancIron, pancreas iron; PancFat, pancreas fat; SBP, mean systolic blood pressure; DBP, mean diastolic blood pressure; Glucagonmin0, fasting glucagon concentration; TotGLP1min0, concentration of fasting total GLP-1 in plasma; HbA1c, glycated hemoglobin A1C; Glucose, fasting glucose from venous plasma samples; TwoGlucose, 2-hour glucose after oral glucose tolerance test/mixed-meal tolerance test; GGTP, gamma-glutamyl transpeptidase; AST, aspartate transaminase; ALT, alanine transaminase; SAT, subcutaneous adipose tissue; BMI, body mass index; TG, fasting triglycerides; VAT, visceral adipose tissue; TwoInsulin, 2-hour insulin; Insulin, fasting insulin from venous plasma samples; BasalISR, insulin secretion at the beginning of the oral glucose tolerance test/ mixed-meal tolerance test.

**S6 Fig.** A. Heatmap cluster among the variables of the UK Biobank (data are inverse normal transformed, n=3641). B. The graphical connection of the variables with Pearson correlation equal or above 0.4.

ALT, alanine transaminase; AST, aspartate transaminase; BMI, body mass index; DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucose, fasting glucose from venous plasma samples; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; VAT, visceral adipose tissue.

**S7 Fig.** Averaged Bayesian network of the bootstrapped samples among the variables of the UK Biobank, panel A-B for non-diabetic (n=3531) and panel C-D for diabetic (n=104) cohorts. Data are inverse normal transformed and the Bayesian networks in panels B and D show only the arcs with strength and directional probabilities  $\geq 0.8$ .

ALT, alanine transaminase; AST, aspartate transaminase; BMI, body mass index; DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucose, fasting glucose from venous plasma samples; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; VAT, visceral adipose tissue.

**S8 Fig.** Averaged Bayesian network of the bootstrapped samples among the variables of the UK Biobank, panel A-B for female (n=1921) and panel C-D for male (n=1720) groups. Data are inverse normal transformed and the Bayesian networks in panels B and D show only the arcs with strength and directional probabilities  $\geq 0.8$ .

ALT, alanine transaminase; AST, aspartate transaminase; BMI, body mass index; DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucose, fasting glucose from venous plasma samples; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; VAT, visceral adipose tissue.

**S9 Fig.** Heatmap cluster and the graphical connection of the UK Biobank variables with Pearson correlation equal or above 0.4, panel A-B for non-diabetic (n=3531) and panel C-D for diabetic (n=104) groups. Data are inverse normal transformed.

ALT, alanine transaminase; AST, aspartate transaminase; BMI, body mass index; DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucose, fasting glucose from venous plasma samples; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; VAT, visceral adipose tissue.

**S10 Fig.** Heatmap cluster and the graphical connection of the UK Biobank variables with Pearson correlation equal or above 0.4, panel A-B for female (n=1921) and panel C-D for male (n=1720) groups. Data are inverse normal transformed.

ALT, alanine transaminase; AST, aspartate transaminase; BMI, body mass index; DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucose, fasting glucose from venous plasma samples; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; VAT, visceral adipose tissue.

**S11 Fig.** Distribution of the variables prior to transformation in the IMI DIRECT combined cohorts (n=1070). The blue curve shows the normally transformed distribution of the variables. OGIS, oral glucose insulin sensitivity index according to the method of Mari et al.<sup>16</sup>; Clins, mean insulin clearance during the oral glucose tolerance test/mixed meal tolerance test, calculated as (mean insulin secretion)/(mean insulin concentration); HDL, fasting high-density lipoprotein cholesterol; GlucoseSens, glucose sensitivity, slope of the dose-response relating insulin secretion to glucose concentration; PancIron, pancreas iron; PancFat, pancreas fat; SBP, mean systolic blood pressure; DBP, mean diastolic blood pressure; Glucagonmin0, fasting glucagon concentration; TotGLP1min0, concentration of fasting total GLP-1 in plasma; HbA1c, glycated hemoglobin A1C; Glucose, fasting glucose from venous plasma samples; TwoGlucose, 2-hour glucose after oral glucose tolerance test/mixed-meal tolerance test; GGTP, gamma-glutamyl transpeptidase; AST, aspartate transaminase; ALT, alanine transaminase; SAT, subcutaneous adipose tissue; BMI, body mass index; TG, fasting triglycerides; VAT, visceral adipose tissue; TwoInsulin, 2-hour insulin; Insulin, fasting insulin from venous plasma samples; BasalISR, insulin secretion at the beginning of the oral glucose tolerance test/ mixed-meal tolerance test.

**S12 Fig.** Distribution of the variables prior to transformation in the UK Biobank cohort (n=3641). The blue curve shows the normally transformed distribution of the variables.

ALT, alanine transaminase; AST, aspartate transaminase; BMI, body mass index; DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; Glucose, fasting glucose from venous plasma samples; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; Insulin, fasting insulin from venous plasma samples; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; VAT, visceral adipose tissue.

**S1 Table.** The 2-sample Mendelian randomization exposure-outcome associations reported per outcome trait.

ALT, alanine transaminase; AST, aspartate transaminase; BMI, body mass index; DBP, mean diastolic blood pressure; GGTP, gamma-glutamyl transpeptidase; HbA1c, glycated hemoglobin A1C; HDL, fasting high-density lipoprotein cholesterol; IDI, insulin disposition index; ISI, insulin sensitivity index; SAT, subcutaneous adipose tissue; SBP, mean systolic blood pressure; TG, fasting triglycerides; VAT, visceral adipose tissue.

**S2 Table.** Information on the strength and directional probabilities of the arcs for the combined, diabetes, non-diabetes, female and male groups of the IMI DIRECT and UK Biobank Bayesian networks. The arcs are limited for those with strength and directional probabilities equal to or greater than the model's significant threshold. The corresponding Mendelian randomization is reported when available.

**S3 Table.** Characteristics of the IMI DIRECT and UK Biobank in the non-diabetes, diabetes and combined cohorts and the list of the variables used in the analyses with their meanings. All the continuous variables were discretized into three levels (Low, Average and High) with Hartemink's method for the probabilistic inference analyses and are presented in this table.

**Competing interests:** HR is an employee and shareholder of Sanofi. MIM: The views expressed in this article are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health. MIM has served on advisory panels for Pfizer, NovoNordisk and Zoe Global, has received honoraria from Merck, Pfizer, Novo Nordisk and Eli Lilly, and research funding from Abbvie, Astra Zeneca, Boehringer Ingelheim, Eli Lilly, Janssen, Merck, NovoNordisk, Pfizer, Roche, Sanofi Aventis, Servier, and Takeda. As of June 2019, MIM is an employee of Genentech, and a holder of Roche stock. AM is a consultant for Lilly and has received research grants from several diabetes drug companies. PWF has received research grants from numerous diabetes drug companies and fess as consultant from Novo Nordisk, Lilly, and Zoe Global Ltd. He is currently the Scientific Director in Patient Care at the Novo Nordisk Foundation. Other authors declare non competing interests.

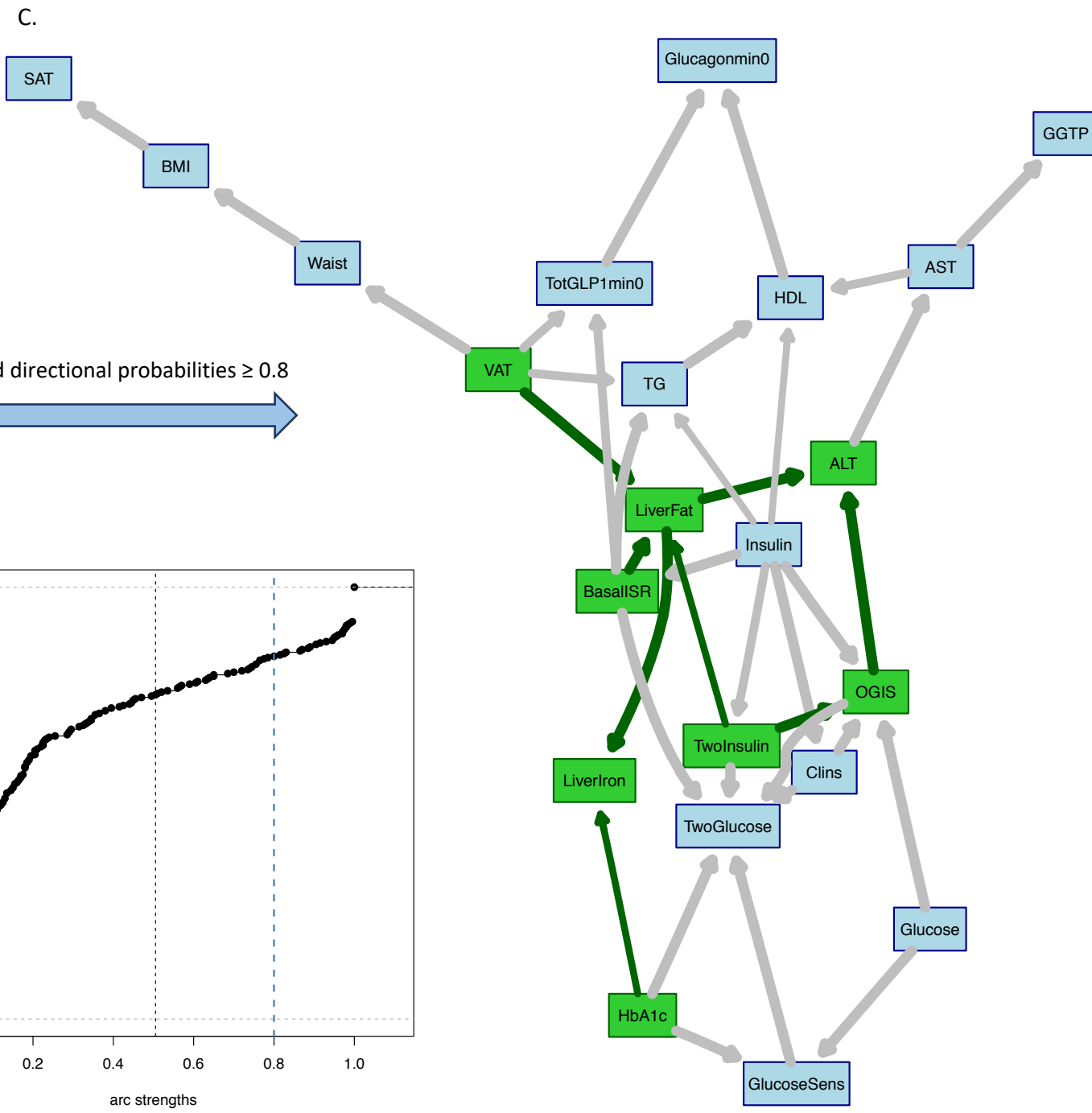
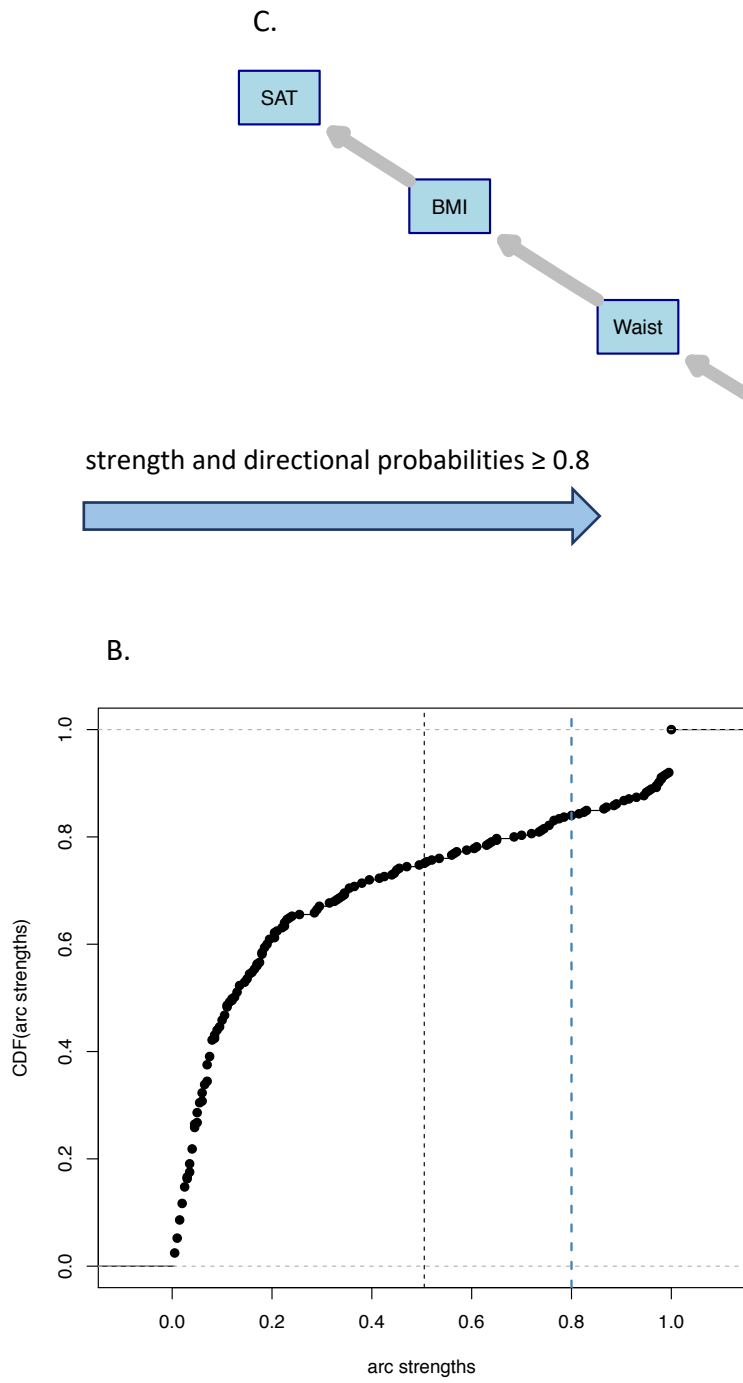
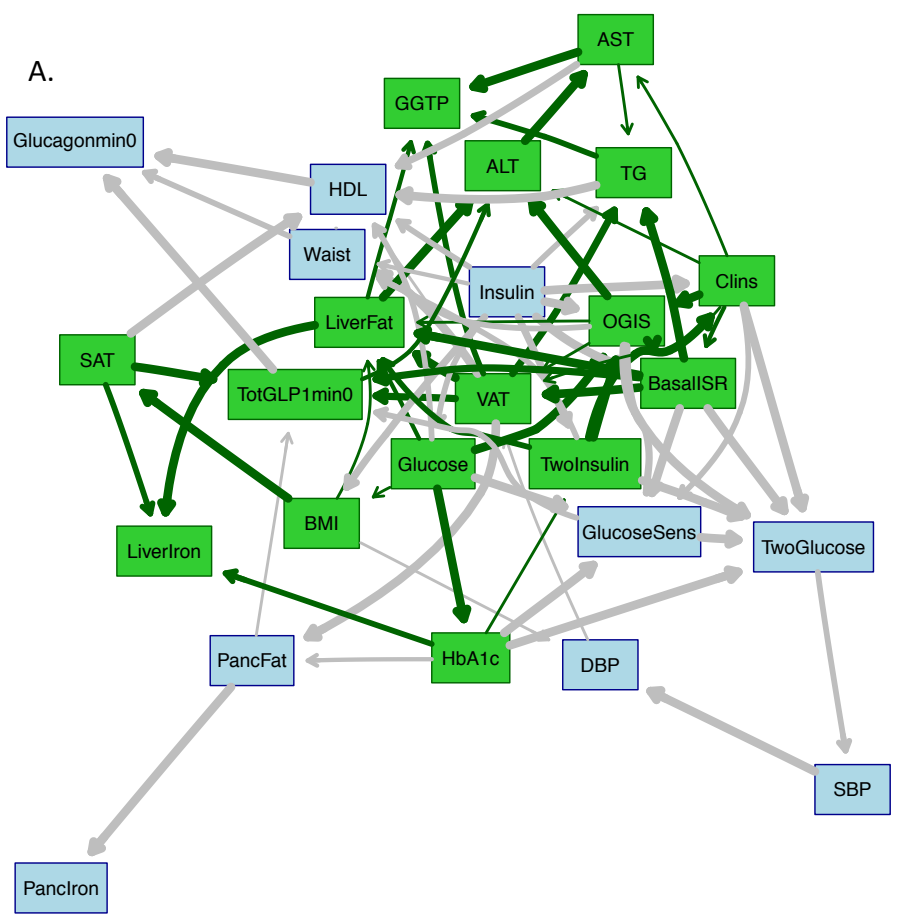
### **Ethics statements**

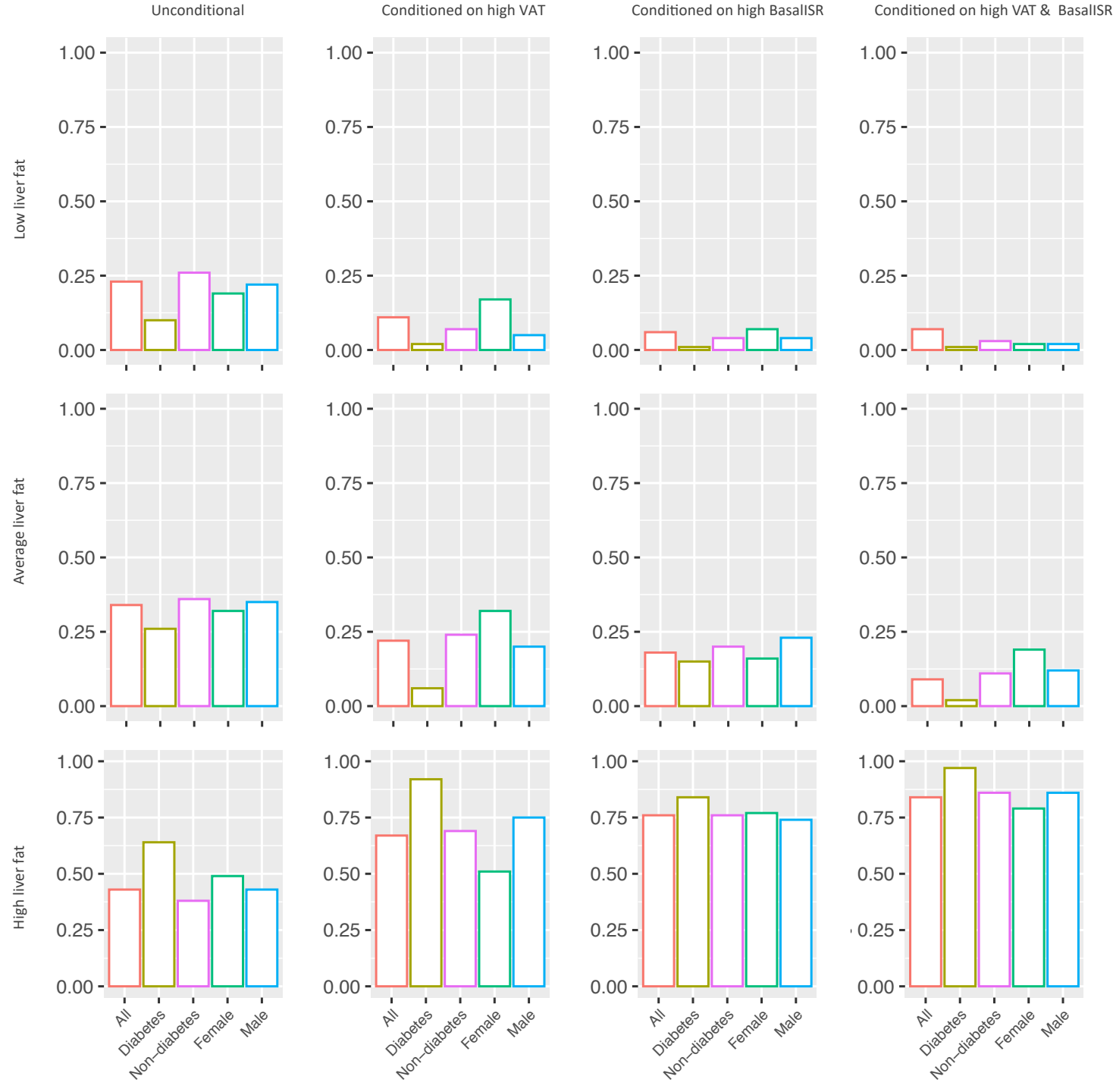
Approval for the study protocol was obtained from each of the regional research ethics review boards separately (Lund, Sweden: 20130312105459927, Copenhagen, Denmark: H-1-2012-166 and H-1-2012-100, Amsterdam, Netherlands: NL40099.029.12, Newcastle, Dundee and Exeter, UK: 12/NE/0132), and all participants provided written informed consent at enrolment. The research conformed to the ethical principles for medical research involving human participants outlined in the Declaration of Helsinki.

### **Acknowledgments**



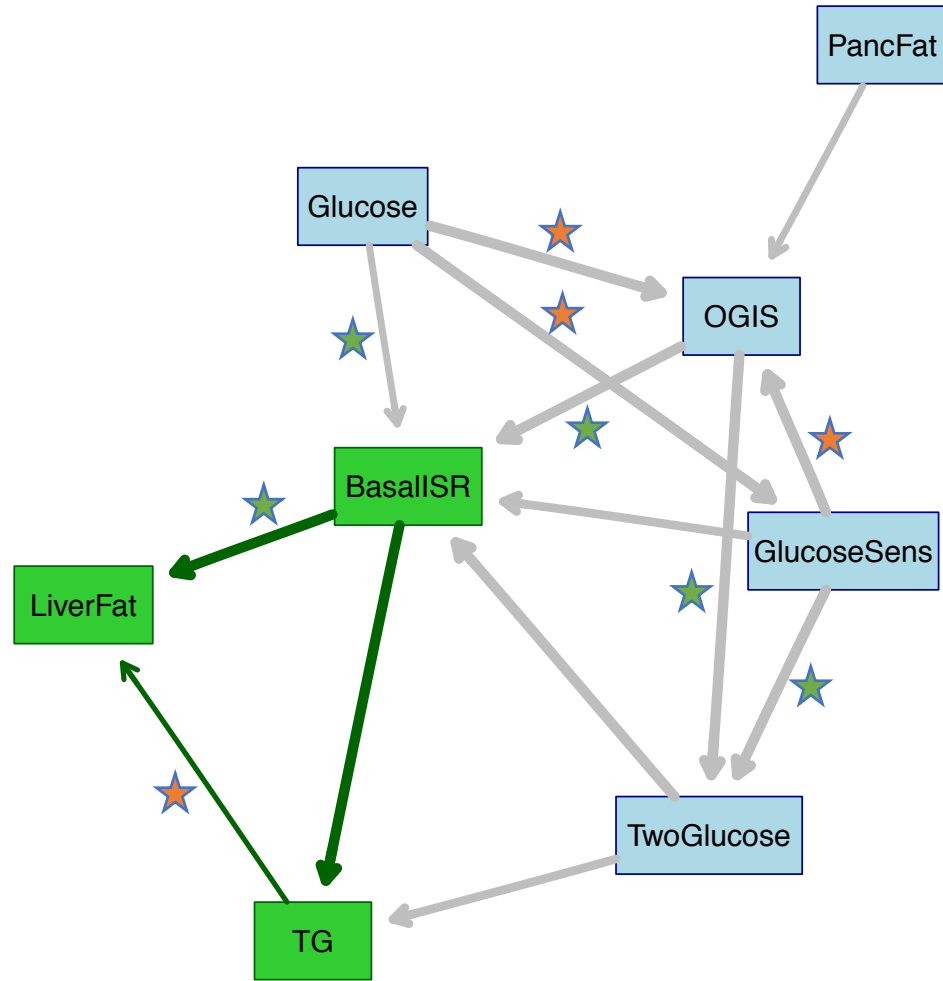
The work leading to this publication has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n°115317 (DIRECT, <https://directdiabetes.org/>), resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. We thank all the participants and study center staff in IMI DIRECT for their contribution to the study. We thank all the participants in the UK Biobank. This research was conducted using the UK Biobank resource (application ID: 18274).







A.

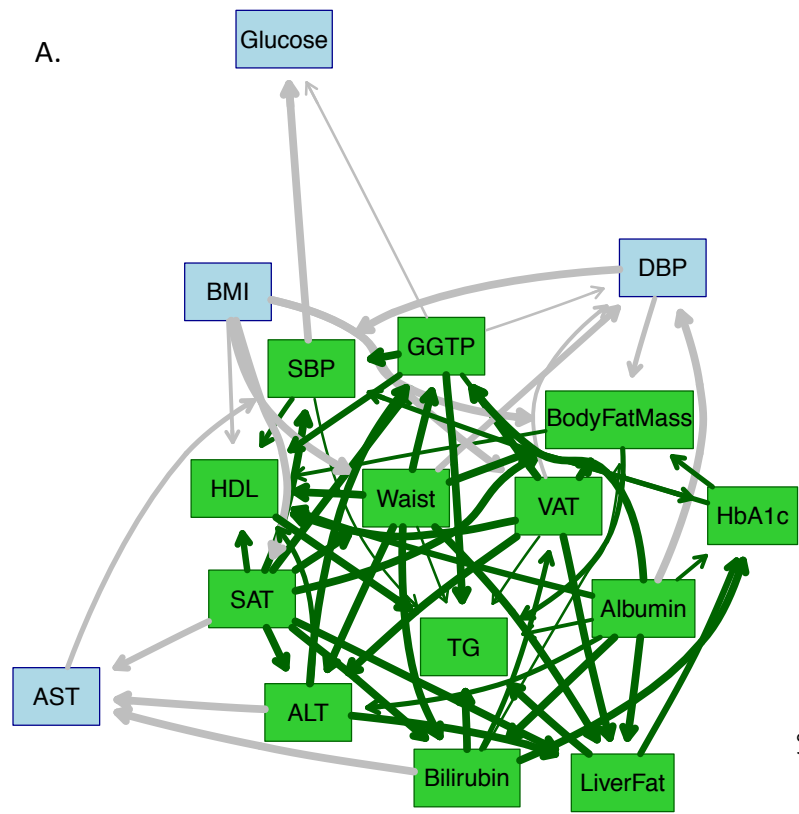


★ Matched with Twin-cycle model

★ Not matched with Twin-cycle model

B.

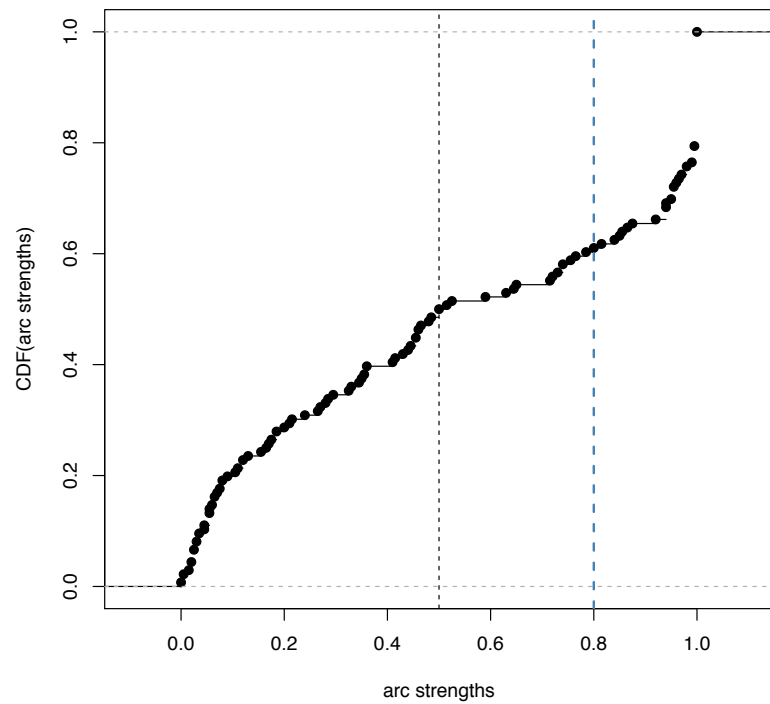
from	to	strength	direction
Glucose	OGIS	1.000	0.870
Glucose	BasalISR	0.800	0.966
Glucose	GlucoseSens	1.000	0.605
TwoGlucose	BasalISR	1.000	0.700
TwoGlucose	TG	0.925	0.957
OGIS	TwoGlucose	1.000	0.877
OGIS	BasalISR	1.000	0.950
BasalISR	LiverFat	1.000	0.895
BasalISR	TG	1.000	0.955
GlucoseSens	TwoGlucose	1.000	0.857
GlucoseSens	OGIS	1.000	0.662
GlucoseSens	BasalISR	0.935	0.896
TG	LiverFat	0.740	0.571
PancFat	OGIS	0.870	0.873



strength and directional probabilities  $\geq 0.8$



B.



C.

