Unravelling the triangular relationship between polycystic ovary syndrome, cardiometabolic disease and de novo lipogenesis

UNRAVELLING THE TRIANGULAR RELATIONSHIP BETWEEN POLYCYSTIC OVARY SYNDROME, CARDIOMETABOLIC DISEASE AND DE NOVO LIPOGENESIS

It's all in the genes

Layout: Tiny Wouters
Cover design: Ingrid Simons & Pomme Simons
Printing: Gildeprint

ISBN: 978-94-6419-472-2

The research described in this thesis was supported by the European Foundation for the Study of Diabetes / Sanofi. Financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged.

UNRAVELLING THE TRIANGULAR RELATIONSHIP BETWEEN POLYCYSTIC OVARY SYNDROME, CARDIOMETABOLIC DISEASE AND DE NOVO LIPOGENESIS

It's all in the genes

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. dr. Pamela Habibović,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen op
woensdag 15 juni 2022 om 16.00 uur

door

Pomme Ingrid Huberdina Gertrudis Simons

Promotoren

Prof. dr. Martijn C.G.J. Brouwers
Prof. dr. Coen D.A. Stehouwer

Copromotor

Dr. Olivier Valkenburg

Beoordelingscommissie

Prof. dr. Marc E.A. Spaanderman (voorzitter)
Prof. dr. Joop S.E. Laven (Erasmus Medisch Centrum)
Dr. Ruth Meex
Prof. dr. Harold Snieder (Rijksuniversiteit Groningen)
Prof. dr. Maurice Zeegers

Table of contents

1

CHAPTER ONE

General introduction

Polycystic ovary syndrome: definition, prevalence and health burden

Polycystic ovary syndrome (PCOS) is a complex, endocrine disorder[1,2]. Women with PCOS can be characterized by a polycystic ovarian morphology, an irregular menstrual cycle, acne, hirsutism, and metabolic disturbances including obesity, insulin resistance and dyslipidaemia[1,2], although the clinical presentation varies substantially between patients with PCOS[2].

The first definition of PCOS was developed in 1990 by the National Institute of Health and defined as the presence of hyperandrogenism and oligo-ovulation[3]. This definition was subsequently adapted during an expert consensus meeting to form the 2003 Rotterdam criteria stating that PCOS can be diagnosed in the presence of two of three criteria; oligo-ovulation, biochemical or clinical hyperandrogenism, and a polycystic ovarian morphology[4]. More recently, the Androgen Excess Society (2006) has proposed to make the definition more focused on the role of androgens in PCOS by suggesting that the diagnosis cannot be made in the absence of biochemical or clinical hyperandrogenism[5]. Although the Rotterdam criteria remain the most commonly accepted and widely implemented diagnostic criteria[6], the debate about the exact definition is ongoing, highlighting the heterogeneous nature of the disorder[7].

PCOS is the most common endocrine disorder amongst premenopausal women, with an estimated prevalence of approximately 10%[8]. PCOS has a great impact on society, through lost work productivity, decreasing quality of life, and burdening the healthcare system[2,9]. Yet the high prevalence of PCOS is worrisome not only for its societal implications, but in particular given its effects on an individual's health. Women with PCOS are at increased risk of developing several complications including subfertility, psychological disorders, endometrial and ovarian cancers, and cardiometabolic disease[1].

Polycystic ovary syndrome and the risk of cardiometabolic disorders

Women with PCOS are at 2 to 3-fold increased risk of developing coronary artery disease and type 2 diabetes[10,11]. Thus far, however, it remains uncertain whether PCOS per se is causal in increasing the risk of cardiometabolic disease. Alternatively, metabolic disturbances that are highly prevalent in PCOS may be common aetiological factors that predispose to both the risk of PCOS and other cardiometabolic disease.

Approximately half of all women with PCOS are obese[12]. Mendelian randomization analyses have reported that genetically predicted risk of obesity was associated with an increased risk of PCOS, while the reverse was not true, that is genetically predicted risk of PCOS did not associate with an increased risk of obesity[13,14]. In addition, it has been well-established that obesity is a risk factor for cardiometabolic disease including coronary artery disease and type 2 diabetes[15]. Therefore, it is of interest to study whether obesity, and downstream complications such as non-alcoholic fatty liver disease (NAFLD; see below), may play a role in increasing the risk of PCOS and other cardiometabolic disorders (Figure 1.1).

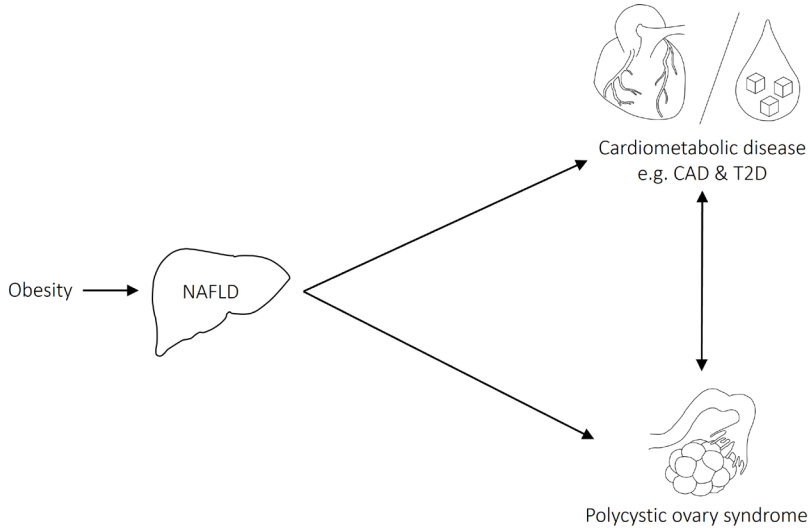

Figure 1.1. The association between polycystic ovary syndrome and cardiometabolic disease (e.g. coronary artery disease [CAD] and type 2 diabetes [T2D]), and the potential role of obesity and non-alcoholic fatty liver disease (NAFLD) herein.

Obesity, non-alcoholic fatty liver disease and de novo lipogenesis

Obesity, defined as the excess accumulation of body fat, is associated with many comorbidities, including hypertension, dyslipidaemia, sleep apnoea, osteoarthritis, insulin resistance and malignancies[16]. Another particularly common comorbidity of obesity is NAFLD[17]. NAFLD has long been viewed as a benign, incidental finding, and marker of obesity and other cardiometabolic disease[18]. However, increasingly, NAFLD is understood to play a central role in metabolism, and actively contribute to the relationship between obesity and cardiometabolic disease[19-21].

Non-alcoholic fatty liver disease

NAFLD comprises a spectrum of histological abnormalities associated with an excess of intrahepatic lipid storage in the absence of excessive alcohol consumption or other underlying hepatic disease[22]. The first stage of NAFLD, i.e. hepatic steatosis, is the accumulation of intrahepatic lipids in more than 5% of hepatocytes[22]. Hepatic steatosis can progress to more advanced stages of NAFLD, characterized by lobular or portal inflammation and ballooning with or without fibrosis (i.e. non-alcoholic steatohepatitis) and can be a precursor for liver failure and hepatocellular carcinoma[22].

The liver plays a central role in lipid metabolism, by contributing to lipid uptake, synthesis, oxidation, export and storage[23]. The accumulation of excessive amounts of intrahepatic lipids is the result of an imbalance in the influx of lipids – through flux of free fatty acids from peripheral adipose tissue or diet, and de novo lipogenesis – and the efflux of lipids – through beta-oxidation of fatty acids or secretion of fatty acids in very-low-density lipoprotein (VLDL) particles (Figure 1.2)[23]. The contribution of each of these pathways to the intrahepatic lipid content is dependent on genetic, environmental and nutritional factors, among others. In obese individuals, the pathway of de novo lipogenesis has been found to be significantly upregulated[24].

De novo lipogenesis

The process of de novo lipogenesis occurs in the liver and, to a much lesser extent, in adipose tissue[19,24]. It encompasses the synthesis of palmitate, and other more complex fatty acids, from non-lipid precursors. Acetyl coenzyme A (acetyl-coA), which can be formed as a product of the glycolytic pathway, serves as a substrate for de novo lipogenesis[19,20]. Decarboxylative condensation reactions elongate acetyl-coA through the addition of acetyl units to form fatty acids[19]. This process is facilitated by several enzymes, including acetyl-coA carboxylase (ACC), the rate limiting enzyme facilitating the conversion of acetyl-coA to malonyl-coA, and fatty acid synthase (FAS), which contributes to several of the intermediate steps in the formation of palmitate from malonyl-coA[19].

De novo lipogenesis is tightly regulated by hormonal factors, including insulin and glucose. Insulin stimulates the expression of the transcription factors sterol regulatory element binding protein 1c (SREBP1c) and liver X receptor (LXR), which both increase transcription of lipogenic genes[21,25]. Glucose stimulates de novo lipogenesis in twofold; first, through increasing the expression of the transcription factor carbohydrate regulatory element binding protein (ChREBP), which in turn activates lipogenic enzymes including ACC and FAS, and second, by providing substrate for the glycolytic pathway and subsequently de novo lipogenesis[21,25].

The role of insulin and glucose in regulating rates of de novo lipogenesis highlights the importance of obesity and nutritional state in influencing de novo lipogenesis. The contribution of de novo lipogenesis to the total intrahepatic lipid content is ~10.9% in healthy individuals, but increases to ~19.4% in obese individuals[24]. This difference is even more pronounced in obese individuals with NAFLD, in whom the contribution of de novo lipogenesis to the total intrahepatic lipid content is ~38.5%[24,26]. Furthermore, de novo lipogenesis is very responsive to dietary changes and elevated rates of de novo lipogenesis are reported in individuals consuming a high-carbohydrate diet[27].

The association between non-alcoholic fatty liver disease and cardiometabolic disease

There is a well-established association between imaging or biopsy proven NAFLD and risk of PCOS, type 2 diabetes and coronary artery disease[28-30]. Although NAFLD and other cardiometabolic disease entities could simply co-exist, research thus far has suggested an active role for NAFLD in the pathogenesis of these disorders[31,32].

Mechanisms linking non-alcoholic fatty liver disease and polycystic ovary syndrome

Case-control studies report a significantly higher prevalence of NAFLD in women with PCOS (~50% to 65%) compared to controls (~25% to 35%)[33-35]. Nevertheless, the mechanisms that link NAFLD and PCOS are incompletely understood, and are likely complex and multifactorial. Insulin resistance and hyperinsulinemia, which has been associated with NAFLD (see "*Mechanisms linking non-alcoholic fatty liver disease and type 2 diabetes*") has been proposed as a potential mechanism linking NAFLD and PCOS. Hyperinsulinemia may induce ovarian follicular arrest by increasing follicle-stimulating hormone-induced differentiation of granulosa cells leading to premature follicle luteinization, and by acting in synergy with elevated luteinizing hormone levels to stimulate ovarian secretion of testosterone[36-38].

A second key link between NAFLD and PCOS may be sex hormone-binding globulin (SHBG), the liver-specific protein that binds testosterone and thereby regulates its bioavailable fraction. In vitro and animal studies have shown that monosaccharide-induced de novo lipogenesis downregulates the expression of hepatocyte nuclear factor 4 alpha, and consequently also serum SHBG levels (Figure 1.2)[39]. In addition, incubation of HepG2 cells with palmitate, the end-product of de novo lipogenesis, has likewise been shown to reduce levels of serum SHBG (Figure 1.2)[39]. In turn, genetically predicted serum SHBG levels have been causally linked to an increased risk of PCOS[40,41].

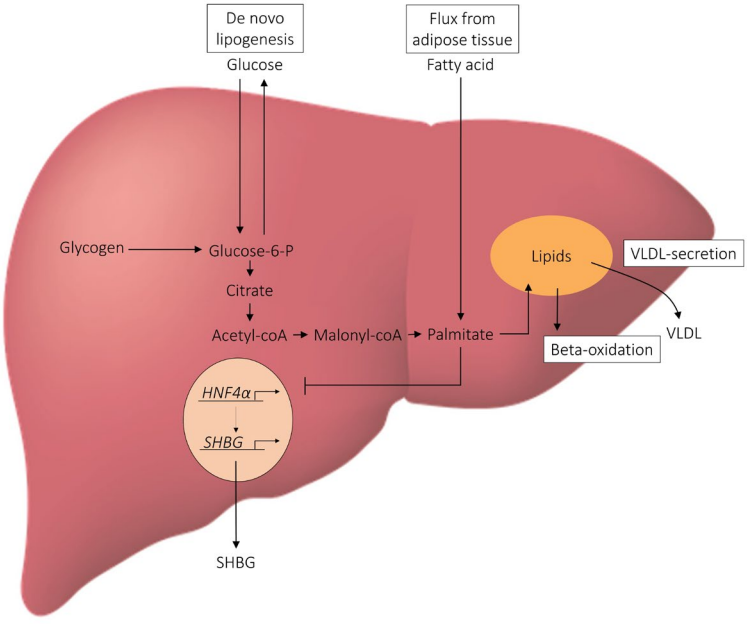

Figure 1.2 Pathways contributing to the accumulation of intrahepatic lipids and their possible consequences on sex hormone-binding globulin synthesis.
Schematic overview of different pathways that contribute to the accumulation of intrahepatic lipids. Experimental studies have shown that palmitate, the end-product of de novo lipogenesis, downregulates hepatocyte nuclear factor 4 alpha (HNF-4α) and subsequently sex hormone-binding globulin (SHBG)[39].
Abbreviations: Acetyl-coA acetyl coenzyme A; glucose-6-p glucose-6-phosphate.

Mechanisms linking non-alcoholic fatty liver disease and type 2 diabetes

Type 2 diabetes, characterized by hyperglycaemia, hyperinsulinemia and peripheral and hepatic insulin resistance, is a frequently observed complication of NAFLD[42,43]. The risk of type 2 diabetes is 2.19-fold higher in patients with biopsy proven NAFLD, independent of common confounders[44]. In addition, clinical studies have identified a significant inverse association of intrahepatic lipid content and de novo lipogenesis with insulin sensitivity, measured with hyperinsulinemic-euglycemic clamp[45,46].

The association between NAFLD and type 2 diabetes is likely the result of a bidirectional underlying pathophysiology. Hyperglycaemia can contribute to increased de novo lipogenesis by providing more substrate and by influencing the expression of ChREBP[21]. ChREBP in turn activates lipogenic enzymes and glucose-6-phosphatase, which converts glucose-6-phosphate into glucose, thereby contributing to increased de novo lipogenesis and hepatic glucose output, respectively[47]. Hyperinsulinemia is also a

stimulatory factor for de novo lipogenesis, by influencing the expression of SREBP-1c and LXR[21]. Moreover, in insulin resistant conditions, the hormone sensitive lipase and lipoprotein lipase activity are altered, thereby increasing the flux of free fatty acids from the adipose tissue towards the liver, increasing intrahepatic lipid content[48-50]. However, the reverse pathway likely also contributes; NAFLD may increase the risk of type 2 diabetes by contributing to the excess circulation of fatty acid metabolites in peripheral tissue resulting in insulin resistance[51,52], impaired suppression of hepatic glucose production (as a result of a reduced ability to suppress glycogenolysis and gluconeogenesis[53]) and elevation of inflammatory cytokines (including interleukin 6 [IL-6] and tumour necrosis factor alpha [TNF-α][54])[55,56].

More recently, it has been proposed that serum SHBG may play a role in mediating the association between NAFLD and type 2 diabetes[57]. A large-scale meta-analysis reported statistically significantly lower serum SHBG levels in women with type 2 diabetes compared to controls. Comparable results were found in men, albeit not statistically significant[58]. In addition, in prospective studies, women with higher serum SHBG levels (>60 nmol/l) had an 80% decreased relative risk of type 2 diabetes, while men with higher serum SHBG levels (>28.3 nmol/l) had a 52% reduced risk of type 2 diabetes[58]. Mendelian randomization analyses have suggested that this association may be the result of a causal effect of serum SHBG by itself on type 2 diabetes. This implicates SHBG as a protein with systemic metabolic effects, i.e. a hepatokine[40,59,60]. This finding has been corroborated by experimental studies[61-63]. Humanized transgenic *SHBG* mice fed a high-fat diet demonstrated improved glucose homeostasis compared to wild-type mice[63]. Nevertheless, it remains uncertain to what extent this pathway may contribute to the risk of type 2 diabetes, and whether the effect of SHBG on type 2 diabetes is direct or mediated by free testosterone.

Mechanisms linking non-alcoholic fatty liver disease and coronary artery disease

Cardiovascular disease has become the leading cause of death in patients with NAFLD[64,65]. A large-scale meta-analysis of 85,395 individuals has also reported an association between NAFLD and pre-clinical cardiovascular disease, including a risk of increased subclinical atherosclerosis, carotid artery intima media thickness, arterial stiffness, coronary artery calcification and endothelial dysfunction[66].

Several mechanisms have been identified that illustrate how NAFLD can actively contribute to the risk of coronary artery disease. First, atherogenic dyslipidaemia (i.e. elevated levels of VLDL, small dense low-density-lipoprotein [LDL] and reduced levels of high-density-lipoprotein [HDL]), a common driver of atherosclerosis and cardiovascular events, is in part the result of a dysregulation of de novo lipogenesis[67-69]. Second,

Mendelian randomization studies have identified an association between genetically predicted insulin resistance and type 2 diabetes with coronary artery disease[70,71]. Third, endothelial dysfunction and hypertension are important contributors to coronary artery disease risk. NAFLD increases the risk of endothelial dysfunction and hypertension through its effects on systemic inflammation (resulting from increased release of inflammatory cytokines [i.e. IL-6, TNF-α and CC-chemokine ligand 2]), oxidative stress (through release of homocysteine and advanced glycation end products from the liver), increased release of other hepatokines (i.e. fetuin-A and retinol-binding protein 4) and vasoactive substances (i.e. asymmetrical dimethylargine which consequently affects nitric oxide synthase)[67,72].

Markers of de novo lipogenesis

Thus far, epidemiological and intervention studies have established a strong relationship between NAFLD and cardiometabolic disease including PCOS, type 2 diabetes and coronary artery disease[31,32]. However, the role of de novo lipogenesis in mediating these relationships remains incompletely understood. In part, this can be attributed to the difficulties in accurately measuring de novo lipogenesis and the absence of a reliable and non-invasive biomarker.

Tracer methods are the most commonly used means to quantify de novo lipogenesis. Stable isotope tracer studies can use the synthesis of VLDL-triglyceride palmitate from tracers, such as deuterium oxide ($_2H_20$) or ^{13}C-acetate, as a marker of newly synthesized fatty acids from a non-lipid precursor, and hence a marker of de novo lipogenesis[73]. More recently, the hepatic saturated fatty acid fraction, measured with proton magnetic resonance imaging (^{1}H-MRS), has also been proposed as a marker of hepatic palmitate content, and hence de novo lipogenesis[45]. Nevertheless, both stable isotope methods and hepatic saturated fatty acid content are relatively invasive, time-consuming, and expensive means of quantifying de novo lipogenesis, and, consequently, cannot be used for large-scale studies.

Common genetic variants are an alternative avenue through which to study the role of de novo lipogenesis in cardiometabolic disease. Although genome-wide association studies of de novo lipogenesis are lacking, several genetic variants have been proposed to be involved in regulating levels of de novo lipogenesis[25]. One is a common and functional variant in the *GCKR* gene[74]. The minor allele of this common variant in *GCKR* encodes a variant of the liver-specific glucokinase regulatory protein which binds less effectively to glucokinase, the enzyme that facilitates the hepatic phosphorylation of glucose (Figure 1.3)[75,76]. The minor variant of *GCKR* is associated with increased hepatic glucose uptake, glycolysis and de novo lipogenesis (Figure 1.3)[74]. The identification of

such de novo lipogenesis susceptibility genes enables the exploration of causal associations between genetic variants of de novo lipogenesis and disease outcomes, i.e. with the use of a genetic risk score or Mendelian randomization analyses[77,78].

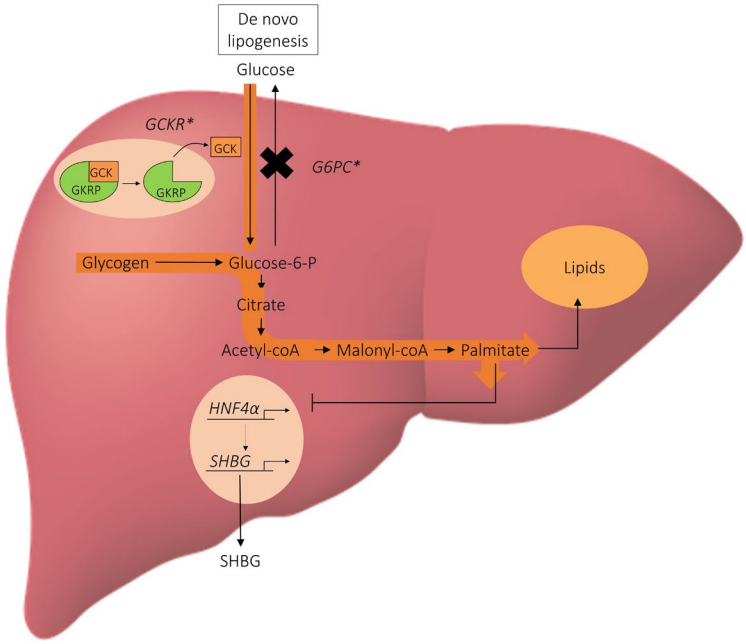

Figure 1.3 Genetic variants leading to increased rates of de novo lipogenesis.
Common variants in *GCKR* encode a variant in glucokinase regulatory protein (GKRP) that binds less efficiently to glucokinase (GCK). Consequently, this variant in *GCKR* is associated with increased hepatic glucose uptake and surplus of glucose-6-phosphate[82]. Rare genetic mutations in *G6PC* encoding glucose-6-phosphatase causes glycogen storage disease type 1a, which is associated with an inability to convert glucose-6-phosphate back to glucose (black cross)[81]. Similar to variants in *GCKR*, GSD1a is associated with a surplus of glucose-6-phosphate. Both genetic changes are associated with increased rates of glycolysis and de novo lipogenesis (orange arrow)[74,81].

Lastly, the downstream effects of de novo lipogenesis can be assessed by studying individuals with (rare) genetic mutations characterized by altered rates of de novo lipogenesis. Glycogen storage disease type 1a (GSD1a) is an inborn error of metabolism, resulting from a mutation in *G6PC* encoding glucose-6-phosphatase[79]. Individuals with GSD1a are characterized by a surplus of glucose-6-phosphate that can serve as a substrate for glycolysis and de novo lipogenesis (Figure 1.3)[80]. Indeed, similar to individuals with the minor allele of the common variant in *GCKR,* patients with GSD1a are characterized by increased rates of de novo lipogenesis (Figure 1.3)[45,81].

Aims and outline of this thesis

Despite the extensive research presented hitherto, it remains uncertain whether PCOS per se plays an active role in the development of cardiometabolic disorders. There is evidence that NAFLD, which is common in women with PCOS, may be the common denominator of PCOS and cardiometabolic disease. However, it remains uncertain to what extent de novo lipogenesis, one of the principal pathways leading to NAFLD, contributes to the development of PCOS and other cardiometabolic disease. In addition, experimental studies suggest that serum SHBG may be involved in the association between NAFLD and cardiometabolic disease, by acting as both a biomarker of de novo lipogenesis, as well as a hepatokine in cardiometabolic disease. Nevertheless, it remains uncertain whether these experimental data can be extrapolated to humans. Therefore, this thesis aims to investigate the association between PCOS and cardiometabolic disease, and to assess the role of de novo lipogenesis and serum SHBG herein. Figure 1.4 summarizes the outline of this thesis.

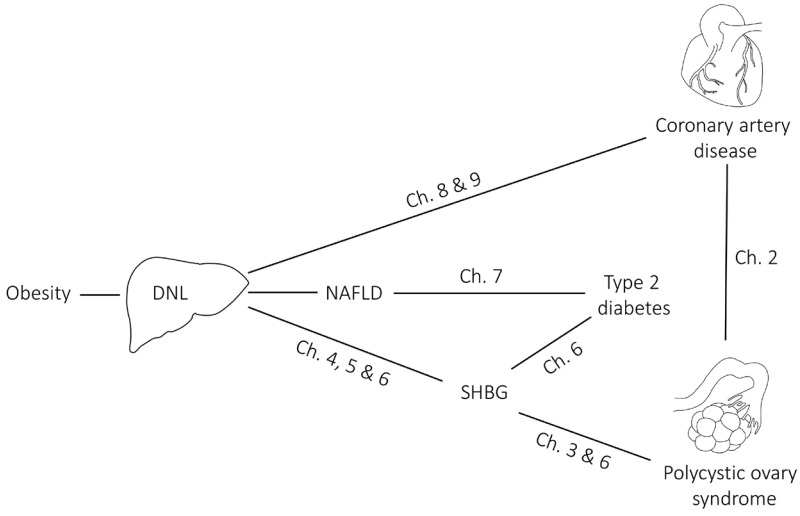

Figure 1.4 Schematic outline of the relationships studied in this thesis.

We first aimed to investigate the (causal) association between PCOS and (risk factors of) cardiometabolic disease.

In **chapter two,** we conducted Mendelian randomization analyses to assess the association between genetically predicted risk of PCOS and risk of coronary artery disease. Moreover, we aimed to assess whether obesity is a common denominator of the risk of both PCOS and coronary artery disease.

In chapter three, by using data of an extensively phenotyped group of PCOS patients, we assessed the associations between serum SHBG, androstenedione, total testosterone and free testosterone with several metabolic and reproductive features of PCOS. As such, we aimed to gain more insight into the association between different features of PCOS, in particular biochemical hyperandrogenism, and several cardiometabolic risk factors.

Second, we aimed to study whether de novo lipogenesis decreases serum SHBG levels in humans.
In chapter four, we studied the association between de novo lipogenesis, measured with stable isotopes, and serum SHBG levels in a sample of men and women with varying degrees of hepatic steatosis and obesity.
In chapter five, we compared the serum SHBG levels in individuals with GSD1a and controls. In this case-control study, we aimed to gain more insight into the effects of de novo lipogenesis on serum SHBG levels in humans.
In chapter six, we extensively studied the results of a genome-wide association study of serum SHBG. We provided our interpretation of these data and identify several SHBG susceptibility genes that are also known to be involved in the regulation of de novo lipogenesis.

Finally, we aimed to study the link between de novo lipogenesis, NAFLD and cardiometabolic disease, and to assess to what extent SHBG acts as a hepatokine in mediating these relationships.
Therefore, in chapter six, we have additionally provided our interpretation of the results of a Mendelian randomization study analysing the association between serum SHBG, total testosterone and free testosterone with PCOS and type 2 diabetes. We studied whether the association between SHBG with PCOS and type 2 diabetes are the result of a direct or indirect effect.
In chapter seven, using data of The Maastricht Study, we conducted mediation analyses to assess whether, and to what extent, serum SHBG has a role in mediating the relationship between intrahepatic lipid content and type 2 diabetes.
In chapter eight, we conducted a literature search and meta-analyses to explore the association between minor variants of *GCKR* – that predispose to higher rates of de novo lipogenesis – with coronary artery disease, estimated glomerular filtration rate and chronic kidney disease.
In chapter nine, using three independent approaches, we identified de novo lipogenesis susceptibility genes and clustered the genetic variants to study their association with coronary artery disease.

In chapter ten, the main findings of this thesis are discussed in view of other literature and methodological considerations.

References

1. Azziz R, Carmina E, Chen Z, et al. Polycystic ovary syndrome. *Nature Reviews Disease Primers.* 2016;2(1):16057.
2. Teede H, Deeks A, Moran L. Polycystic ovary syndrome: a complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC Med.* 2010;8:41-41.
3. Zawadski JK, Dunaif A. Diagnostic Criteria for Polycystic Ovary Syndrome: Towards a Rational Approach. In: Dunaif A, Givens JR and Haseltine F, Eds., Polycystic Ovary Syndrome, Blackwell Scientific, Boston, 377-384. 1992.
4. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril.* 2004;81(1):19-25.
5. Azziz R, Carmina E, Dewailly D, et al. The Androgen Excess and PCOS Society criteria for the polycystic ovary syndrome: the complete task force report. *Fertil Steril.* 2009;91(2):456-488.
6. Teede HJ, Misso ML, Costello MF, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Hum Reprod.* 2018;33(9):1602-1618.
7. Escobar-Morreale HF. Polycystic ovary syndrome: Definition, aetiology, diagnosis and treatment. *Nat Rev Endocrinol.* 2018;14(5):270-284.
8. Wolf WM, Wattick RA, Kinkade ON, Olfert MD. Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity. *International Journal of Environmental Research and Public Health.* 2018;15(11):2589.
9. Azziz R, Marin C, Hoq L, Badamgarav E, Song P. Health care-related economic burden of the polycystic ovary syndrome during the reproductive life span. *Journal of Clinical Endocrinology and Metabolism.* 2005;90(8):4650-4658.
10. de Groot PC, Dekkers OM, Romijn JA, Dieben SW, Helmerhorst FM. PCOS, coronary heart disease, stroke and the influence of obesity: A systematic review and meta-analysis. *Hum Reprod Update.* 2011;17(4):495-500.
11. Kakoly NS, Khomami MB, Joham AE, et al. Ethnicity, obesity and the prevalence of impaired glucose tolerance and type 2 diabetes in PCOS: A systematic review and meta-regression. *Hum Reprod Update.* 2018;24(4):455-467.
12. Lim SS, Davies MJ, Norman RJ, Moran LJ. Overweight, obesity and central obesity in women with polycystic ovary syndrome: A systematic review and meta-analysis. *Hum Reprod Update.* 2012;18(6):618-637.
13. Brower MA, Hai Y, Jones MR, et al. Bidirectional Mendelian randomization to explore the causal relationships between body mass index and polycystic ovary syndrome. *Hum Reprod.* 2018;34(1):127-136.
14. Zhao Y, Xu Y, Wang X, et al. Body mass index and polycystic ovary syndrome: A 2-sample bidirectional Mendelian randomization study. *Journal of Clinical Endocrinology and Metabolism.* 2020;105(6).
15. Powell-Wiley TM, Poirier P, Burke LE, et al. Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Circulation.* 2021;143(21):e984-e1010.
16. Blüher M. Obesity: global epidemiology and pathogenesis. *Nature Reviews Endocrinology.* 2019;15(5):288-298.
17. Younossi ZM. Non-alcoholic fatty liver disease - A global public health perspective. *J Hepatol.* 2019;70(3):531-544.
18. Tilg H, Targher G. NAFLD-related mortality: simple hepatic steatosis is not as 'benign' as thought. *Gut.* 2020:gutjnl-2020-323188.
19. Nguyen P, Leray V, Diez M, et al. Liver lipid metabolism. *Journal of Animal Physiology and Animal Nutrition.* 2008;92(3):272-283.
20. Ameer F, Scandiuzzi L, Hasnain S, Kalbacher H, Zaidi N. De novo lipogenesis in health and disease. *Metabolism.* 2014;63(7):895-902.
21. Sanders FWB, Griffin JL. De novo lipogenesis in the liver in health and disease: more than just a shunting yard for glucose. *Biological Reviews.* 2016;91(2):452-468.

22. EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol.* 2016;64(6):1388-1402.
23. Ipsen DH, Lykkesfeldt J, Tveden-Nyborg P. Molecular mechanisms of hepatic lipid accumulation in non-alcoholic fatty liver disease. *Cellular and Molecular Life Sciences.* 2018;75(18):3313-3327.
24. Smith GI, Shankaran M, Yoshino M, et al. Insulin resistance drives hepatic de novo lipogenesis in nonalcoholic fatty liver disease. *J Clin Invest.* 2020;130(3):1453-1460.
25. Strable MS, Ntambi JM. Genetic control of de novo lipogenesis: role in diet-induced obesity. *Critical reviews in biochemistry and molecular biology.* 2010;45(3):199-214.
26. Lambert JE, Ramos-Roman MA, Browning JD, Parks EJ. Increased de novo lipogenesis is a distinct characteristic of individuals with nonalcoholic fatty liver disease. *Gastroenterology.* 2014;146(3):726-735.
27. Schwarz JM, Linfoot P, Dare D, Aghajanian K. Hepatic de novo lipogenesis in normoinsulinemic and hyperinsulinemic subjects consuming high-fat, low-carbohydrate and low-fat, high-carbohydrate isoenergetic diets. *Am J Clin Nutr.* 2003;77(1):43-50.
28. Mantovani A, Byrne CD, Bonora E, Targher G. Nonalcoholic fatty liver disease and risk of incident type 2 diabetes: A meta-analysis. *Diabetes Care.* 2018;41(2):372-382.
29. Targher G, Byrne CD, Lonardo A, Zoppini G, Barbui C. Non-alcoholic fatty liver disease and risk of incident cardiovascular disease: A meta-analysis. *J Hepatol.* 2016;65(3):589-600.
30. Rocha ALL, Faria LC, Guimarães TCM, et al. Non-alcoholic fatty liver disease in women with polycystic ovary syndrome: Systematic review and meta-analysis. *J Endocrinol Invest.* 2017;40(12):1279-1288.
31. Adams LA, Anstee QM, Tilg H, Targher G. Non-alcoholic fatty liver disease and its relationship with cardiovascular disease and other extrahepatic diseases. *Gut.* 2017;66(6):1138-1153.
32. Vassilatou E. Nonalcoholic fatty liver disease and polycystic ovary syndrome. *World J Gastroenterol.* 2014;20(26):8351-8363.
33. Macut D, Tziomalos K, Božić-Antić I, et al. Non-alcoholic fatty liver disease is associated with insulin resistance and lipid accumulation product in women with polycystic ovary syndrome. *Hum Reprod.* 2016;31(6):1347-1353.
34. Gutierrez-Grobe Y, Ponciano-Rodríguez G, Ramos MH, Uribe M, Méndez-Sánchez N. Prevalence of non alcoholic fatty liver disease in premenopausal, posmenopausal and polycystic ovary syndrome women. . *Ann Hepatol.* 2010;9(4):402-409.
35. Karoli R, Fatima J, Chandra A, Gupta U, Islam FU, Singh G. Prevalence of hepatic steatosis in women with polycystic ovary syndrome. *J Hum Reprod Sci.* 2013;6(1):9-14.
36. Goodarzi MO, Dumesic DA, Chazenbalk G, Azziz R. Polycystic ovary syndrome: etiology, pathogenesis and diagnosis. *Nat Rev Endocrinol.* 2011;7(4):219-231.
37. Franks S, Gilling-Smith C, Watson H, Willis D. Insulin action in the normal and polycystic ovary. *Endocrinol Metab Clin North Am.* 1999;28(2):361-378.
38. Willis D, Mason H, Gilling-Smith C, Franks S. Modulation by insulin of follicle-stimulating hormone and luteinizing hormone actions in human granulosa cells of normal and polycystic ovaries. *Journal of Clinical Endocrinology and Metabolism.* 1996;81(1):302-309.
39. Selva DM, Hogeveen KN, Innis SM, Hammond GL. Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. *J Clin Invest.* 2007;117(12):3979-3987.
40. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat Med.* 2020;26(2):252-258.
41. Day FR, Hinds DA, Tung JY, et al. Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. *Nature Communications.* 2015;6:8464-8464.
42. Targher G, Byrne CD. Nonalcoholic fatty liver disease: A novel cardiometabolic risk factor for type 2 diabetes and its complications. *Journal of Clinical Endocrinology and Metabolism.* 2013;98(2):483-495.
43. DeFronzo RA, Ferrannini E, Groop L, et al. Type 2 diabetes mellitus. *Nature Reviews Disease Primers.* 2015;1(1):15019.
44. Mantovani A, Petracca G, Beatrice G, Tilg H, Byrne CD, Targher G. Non-alcoholic fatty liver disease and risk of incident diabetes mellitus: an updated meta-analysis of 501 022 adult individuals. *Gut.* 2021;70(5):962-969.

45. Roumans KHM, Lindeboom L, Veeraiah P et al. Hepatic saturated fatty acid fraction is associated with de novo lipogenesis and hepatic insulin resistance. *Nature Communications.* 2020;11(1):1891.
46. Korenblat KM, Fabbrini E, Mohammed BS, Klein S. Liver, muscle, and adipose tissue insulin action is directly related to intrahepatic triglyceride content in obese subjects. *Gastroenterology.* 2008;134(5):1369-1375.
47. Kim MS, Krawczyk SA, Doridot L, et al. ChREBP regulates fructose-induced glucose production independently of insulin signaling. *J Clin Invest.* 2019;126(11):4372-4386.
48. Watt MJ, Carey AL, Wolsk-Petersen E, Kraemer FB, Pedersen BK, Febbraio MA. Hormone-sensitive lipase is reduced in the adipose tissue of patients with type 2 diabetes mellitus: influence of IL-6 infusion. *Diabetologia.* 2005;48(1):105-112.
49. Costabile G, Annuzzi G, Di Marino L, et al. Fasting and post-prandial adipose tissue lipoprotein lipase and hormone-sensitive lipase in obesity and type 2 diabetes. *J Endocrinol Invest.* 2011;34(5):e110-114.
50. Ruge T, Sukonina V, Kroupa O, et al. Effects of hyperinsulinemia on lipoprotein lipase, angiopoietin-like protein 4, and glycosylphosphatidylinositol-anchored high-density lipoprotein binding protein 1 in subjects with and without type 2 diabetes mellitus. *Metabolism.* 2012;61(5):652-660.
51. Kitade H, Chen G, Ni Y, Ota T. Nonalcoholic Fatty Liver Disease and Insulin Resistance: New Insights and Potential New Treatments. *Nutrients.* 2017;9(4):387.
52. Petersen KF, Dufour S, Befroy D, Lehrke M, Hendler RE, Shulman GI. Reversal of Nonalcoholic Hepatic Steatosis, Hepatic Insulin Resistance, and Hyperglycemia by Moderate Weight Reduction in Patients With Type 2 Diabetes. *Diabetes.* 2005;54(3):603-608.
53. Sunny NE, Parks EJ, Browning JD, Burgess SC. Excessive hepatic mitochondrial TCA cycle and gluconeogenesis in humans with nonalcoholic fatty liver disease. *Cell Metab.* 2011;14(6):804-810.
54. Khan RS, Bril F, Cusi K, Newsome PN. Modulation of insulin resistance in nonalcoholic fatty liver disease. *Hepatology.* 2019;70(2):711-724.
55. Sharabi K, Tavares CDJ, Rines AK, Puigserver P. Molecular pathophysiology of hepatic glucose production. *Mol Aspects Med.* 2015;46:21-33.
56. Tilg H, Moschen AR, Roden M. NAFLD and diabetes mellitus. *Nature Reviews Gastroenterology & Hepatology.* 2017;14(1):32-42.
57. Le TN, Nestler JE, Strauss JF 3rd, Wickham EP 3rd. Sex hormone-binding globulin and type 2 diabetes mellitus. *Trends in Endocrinology and Metabolism.* 2012;23(1):32-40.
58. Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: A systematic review and meta-analysis. *JAMA.* 2006;295(11):1288-1299.
59. Perry JRB, Weedon MN, Langenberg C, et al. Genetic evidence that raised sex hormone binding globulin (SHBG) levels reduce the risk of type 2 diabetes. *Hum Mol Genet.* 2010;19(3):535-544.
60. Ding EL, Song Y, Manson JE, et al. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *New England Journal of Medicine.* 2009;361(12):1152-1163.
61. Saez-Lopez C, Barbosa-Desongles A, Hernandez C, et al. Sex hormone-binding globulin reduction in metabolic disorders may play a role in NAFLD development. *Endocrinology.* 2017;158(3):545-559.
62. Sáez-López C, Salcedo-Allende MT, Hernandez C, Simó-Servat O, Simó R, Selva DM. Sex hormone-binding globulin expression correlates with acetyl-coenzyme a carboxylase and triglyceride content in human liver. *J Clin Endocrinol Metab.* 2019;104(5):1500-1507.
63. Saez-Lopez C, Villena JA, Simó R, Selva DM. Sex hormone-binding globulin overexpression protects against high-fat diet-induced obesity in transgenic male mice. *J Nutr Biochem.* 2020;85:108480.
64. Adams LA, Lymp JF, St. Sauver J, et al. The natural history of nonalcoholic fatty liver disease: A population-based cohort study. *Gastroenterology.* 2005;129(1):113-121.
65. Ong JP, Pitts A, Younossi ZM. Increased overall mortality and liver-related mortality in non-alcoholic fatty liver disease. *J Hepatol.* 2008;49(4):608-612.
66. Zhou YY, Zhou XD, Wu SJ, et al. Nonalcoholic fatty liver disease contributes to subclinical atherosclerosis: A systematic review and meta-analysis. *Hepatology Communications.* 2018;2(4):376-392.
67. Targher G, Corey KE, Byrne CD. NAFLD, and cardiovascular and cardiac diseases: Factors influencing risk, prediction and treatment. *Diabetes Metab.* 2021;47(2):101215.

68. Brouwers MCGJ, Simons N, Stehouwer CDA, Isaacs A. Non-alcoholic fatty liver disease and cardiovascular disease: assessing the evidence for causality. *Diabetologia.* 2020;63(2):253-260.
69. Brouwers MCGJ, Simons N, Stehouwer CDA, Koek GH, Schaper NC, Isaacs A. Relationship between nonalcoholic fatty liver disease susceptibility genes and coronary artery disease. *Hepatology Communications.* 2019;3(4):587-596.
70. Ahmad OS, Morris JA, Mujammami M, et al. A Mendelian randomization study of the effect of type-2 diabetes on coronary heart disease. *Nature Communications.* 2015;6(1):7060.
71. Ross S, Gerstein HC, Eikelboom J, Anand SS, Yusuf S, Paré G. Mendelian randomization analysis supports the causal role of dysglycaemia and diabetes in the risk of coronary artery disease. *Eur Heart J.* 2015;36(23):1454-1462.
72. Zhao YC, Zhao GJ, Chen Z, She ZG, Cai J, Li H. Nonalcoholic fatty liver disease: An emerging driver of hypertension. *Hypertension.* 2020;75(2):275-284.
73. McCullough A, Previs S, Kasumov T. Stable isotope-based flux studies in nonalcoholic fatty liver disease. *Pharmacol Ther.* 2018;181:22-33.
74. Santoro N, Caprio S, Pierpont B, Van Name M, Savoye M, Parks EJ. Hepatic de novo lipogenesis in obese youth is modulated by a common variant in the GCKR gene. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(8):E1125-E1132.
75. Beer NL, Tribble ND, McCulloch LJ, et al. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum Mol Genet.* 2009;18(21):4081-4088.
76. Rees MG, Wincovitch S, Schultz J, et al. Cellular characterisation of the GCKR P446L variant associated with type 2 diabetes risk. *Diabetologia.* 2012;55(1):114-122.
77. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med.* 2020;12(1):44.
78. Smith GD, Ebrahim S. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003;32(1):1-22.
79. Rake JP, Visser G, Labrune P, Leonard JV, Ullrich K, Smit GP. Glycogen storage disease type I: Diagnosis, management, clinical course and outcome. Results of the European Study on Glycogen Storage Disease Type I (ESGSD I). *Eur J Pediatr.* 2002;161 Suppl 1:S20-34.
80. Bandsma RH, Smit GP, Kuipers F. Disturbed lipid metabolism in glycogen storage disease type 1. *Eur J Pediatr.* 2002;161 Suppl 1:S65-69.
81. Bandsma RH, Prinsen BH, van der Velden de Sain M, et al. Increased de novo lipogenesis and delayed conversion of large VLDL into intermediate density lipoprotein particles contribute to hyperlipidemia in glycogen storage disease type 1a. *Pediatr Res.* 2008;63(6):702-707.
82. Brouwers MCGJ, Jacobs C, Bast A, Stehouwer CDA, Schaper NC. Modulation of glucokinase regulatory protein: A double-edged sword? *Trends Mol Med.* 2015;21(10):583-594.

PART I

THE ASSOCIATION BETWEEN PCOS AND (RISK FACTORS) OF CARDIOMETABOLIC DISEASE

2

CHAPTER TWO

Causal relationship between polycystic ovary syndrome and coronary artery disease: a Mendelian randomization study

Pomme I.H.G. Simons, Merel E.B. Cornelissen, Olivier Valkenburg, N. Charlotte Onland-Moret, Yvonne T. van der Schouw, Coen D.A. Stehouwer, Stephen Burgess, Martijn C.G.J. Brouwers

Clinical Endocrinology (2022); 96 (4): 599-604

Summary

Objective: Polycystic ovary syndrome (PCOS) has been associated with an increased risk of coronary artery disease. However, it remains uncertain whether this increased risk is the result of PCOS per se or, alternatively, is explained by obesity, a common feature of PCOS. The aim of this study was to assess the causal association between PCOS and coronary artery disease, and the role of obesity herein.

Design and method: We conducted two-sample Mendelian randomization analyses in large-scale, female-specific datasets to study the association between genetically predicted 1) risk of PCOS and risk of coronary artery disease, 2) BMI and risk of PCOS, and 3) BMI and risk of coronary artery disease. Primary analyses were conducted with the inverse-variance weighted (IVW) method. Simple median, penalised weighted median and contamination mixture analyses were performed to assess the robustness of the outcomes.

Results: IVW analyses did not show a statistically significant association between PCOS and coronary artery disease (OR: 0.99, 95% CI: 0.89;1.11). In contrast, genetically predicted BMI was statistically significantly associated with an increased odds of PCOS (OR: 3.21, 95% CI: 2.26;4.56) and coronary artery disease (OR: 1.38, 95% CI: 1.14;1.67). Similar results were obtained when secondary analyses were performed.

Conclusion: These sex-specific analyses show that genetically predicted risk of PCOS is not associated with the risk of coronary artery disease. Instead, genetically predicted risk of obesity (and its downstream metabolic effects) is the common denominator of both PCOS and coronary artery disease risk.

Introduction

Polycystic ovary syndrome (PCOS) is the most common endocrine disorder in premenopausal women[1]. Epidemiological studies have shown that patients with PCOS are at increased risk of developing coronary artery disease[2]. There is, however, an ongoing discussion on whether this increased risk is explained by PCOS per se, or, alternatively, by other factors that are frequently observed in PCOS, such as obesity and its metabolic sequalae[3].

The Mendelian randomization approach may be helpful in resolving this conundrum. As individuals are randomized at conception to receive gene variants that either predispose to or protect from PCOS (or obesity), these gene variants can be used as instrumental variables to study the causal relationship between PCOS and coronary artery disease, and the role of obesity herein. A valid Mendelian randomization analysis is subject to three primary assumptions: 1) the genetic variants are associated with the exposure, 2) the genetic variants do not influence the outcome directly, other than through the exposure and 3) the genetic variants do not associate with any confounders[4].

Although a recent Mendelian randomization study failed to demonstrate an association between genetically predicted risk of PCOS and risk of coronary artery disease, the validity of the outcomes is limited by the use of a gene-outcome dataset that included women *and* men[5]. The importance of a sex-specific dataset is emphasized by the recognition of sexual dimorphism in gene-outcome associations[6].

Therefore, in the present study, we conducted a two-sample Mendelian randomization analysis to assess the association between genetically predicted risk of PCOS and the risk of coronary artery disease, using female-specific data. Furthermore, we performed two-sample Mendelian randomization analyses to determine the association between genetically predicted body mass index (BMI) and risk of PCOS and coronary artery disease (Figure 2.1).

Methods

All analyses were conducted with female-specific, summary-level data, which were derived from large-scale cohorts as described below.

Polycystic ovary syndrome

Gene-exposure data for the association between PCOS and coronary artery disease, and gene-outcome data for the association between BMI and PCOS, were retrieved from a meta-analysis of genome-wide association (GWA) studies of the PCOS trait, adjusted for age[7]. This database includes 10,074 PCOS cases and 103,164 controls, of European ancestry. Cases were defined according to the National Institutes of Health (NIH), the Rotterdam criteria for the diagnosis of PCOS, or self-reported history of PCOS (Table 2.1). Single nucleotide polymorphisms (SNPs) were selected as instrumental variables if they demonstrated genome-wide significance ($p<5*10^{-8}$) for the association with PCOS. SNPs were excluded if they were in linkage disequilibrium ($r^2>0.1$, the SNP with the largest absolute effect estimate was retained), had poor imputation quality ($R^2<0.3$ or INFO <0.4), or were palindromic (with a minor allele frequency >0.42). The mean F statistic (determined as the average F-statistic of all genetic variants, calculated as $\frac{\hat{\gamma}_j^2}{\sigma_{Xj}^2}$, where $\hat{\gamma}_j^2$ and σ_{Xj}^2 represent the effect estimate and standard error of the gene-exposure regression, respectively[8]) was calculated as a measure of instrumental variable strength, where a mean F statistic >10 is indicative of a strong set of instrumental variables[9].

Body mass index

Gene-exposure data for the association between BMI and PCOS, and between BMI and coronary artery disease were retrieved from a sex-stratified GWA study of BMI[10]. This GWA study was performed in 73,137 women primarily of European descent (~99.5%) (Table 2.1). Selection of female-specific, genome-wide significant SNPs was similar to the selection of the PCOS SNPs.

Coronary artery disease

Summary-level, gene-outcome data for the association between BMI and coronary artery disease, and between PCOS and coronary artery disease were retrieved from the UK Biobank (application #7439)[11]. This population-based cohort study includes 8,403 female coronary artery disease cases and 190,435 female controls of European descent, aged between 40 and 69 years. Coronary artery disease was defined according to ICD-9 codes (410.X-412.X, 414.X, 414.8, 414.9), ICD-10 or cause of death codes (I21.X-I24.X, I25.1, I25.2, I25.5, I25.6, I25.8, I25.9), or self-reported history of coronary artery disease (Table 2.1).

Table 2.1 Overview of databases used for gene-exposure and gene-outcome data.

GWA study	Trait	N controls	N cases	Definition of cases	Ethnicity	Reference
Day et al.	PCOS	103,164	10,074	– National Institutes of Health (NIH) criteria for PCOS (i.e. the presence of oligo- or amenorrhea and clinical or biochemical hyperandrogenism) or – Rotterdam criteria for PCOS (i.e. the presence of two out of three characteristics: oligo- or amenorrhea, clinical or biochemical hyperandrogenism and/or polycystic ovarian morphology), or – Self-reported history of PCOS	European	[7]
Locke et al.	BMI	73,137*		Not applicable	Primarily European (~99.5%)	[10]
UK Biobank	CAD	190,435	8,403	– ICD-9 codes: 410.X-412.X, 414.X, 414.8, 414.9, or – ICD-10 and cause of death codes: I21.X-I24.X, I25.1, I25.2, I25.5, I25.6, I25.8, I25.9, or – Self-reported history of CAD.	European	[11]

***** Total number of included individuals

Abbreviations: BMI body mass index; CAD coronary artery disease; GWA genome-wide association; ICD international classification of diseases; PCOS polycystic ovary syndrome

Statistical analyses

Inverse-variance weighted Mendelian randomization analyses with a random-effects model were performed as the primary analysis for all three associations (Figure 2.1). Cochran's Q statistic was calculated to identify heterogeneity of the effect estimates. Egger's regression analyses (MR-Egger) were conducted to assess potential directional pleiotropy. A statistically significant intercept is indicative of directional pleiotropy, which is a violation of one of the instrumental variable assumptions[12]. We additionally conducted: 1) simple median (which provides a consistent effect estimate if at least 50% of the genetic variants are valid instruments[13]), penalised weighted median (which downweighs the contribution of genetic variants with heterogeneous effect estimates, and is, therefore, less influenced by significant outliers[13]), and contamination mixture analyses (which assumes that the true effect estimate is represented by the largest number of genetic instruments, and, hence, only a minority of genetic variants need to be valid provided there is no larger group of invalid variants with similar estimates [i.e. the plurality assumption][14,15]), to assess the Mendelian randomization effect estimates under more stringent assumptions; 2) the Mendelian Randomization-Pleiotropy Residual Sum and Outlier (MR-PRESSO) method, which excludes any variant that shows significant heterogeneity for the effect estimates, and, hence, is more robust for

outliers[16]; and 3) Steiger-filtering analyses, which identifies any genetic variant that has a stronger association with the outcome than with the exposure, therefore accounting for potential reverse causality[17]. The effect estimates for all analyses are presented as an increase in odds of the outcome per unit increase in log(odds) of PCOS, or per standard deviation increase in BMI. All analyses were performed using R statistical software, version 4.0.1 (R Foundation for Statistical Computing, Vienna, Austria) with the *TwoSampleMR* and *MendelianRandomization* packages[18,19].

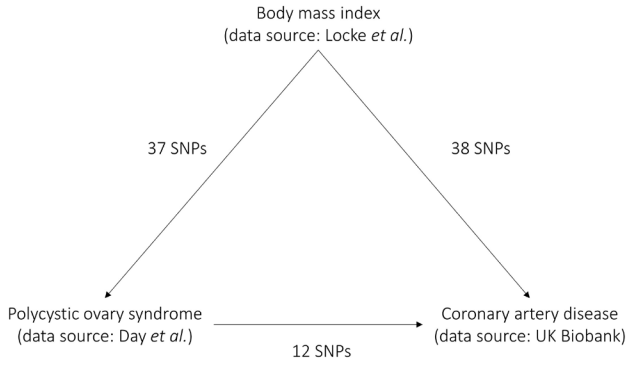

Figure 2.1 Overview of the Mendelian randomization analyses.
Three Mendelian randomization analyses were conducted to assess the association between 1) genetically predicted risk of polycystic ovary syndrome (PCOS) and risk of coronary artery disease (CAD), 2) genetically predicted body mass index (BMI) and risk of CAD and 3) genetically predicted BMI and risk of PCOS.

Results

Polycystic ovary syndrome and coronary artery disease

The GWA study of PCOS identified 19 SNPs that showed genome-wide significance[7]. Seven SNPs were excluded as they were in linkage disequilibrium (rs1351592, rs10993397, rs11031006, rs1795379), were palindromic (rs1351592; rs2271194), or had poor imputation quality (rs151212108). This resulted in 12 independent SNPs that were used as genetic instruments for PCOS (Supplementary Table S2.1), with a mean F statistic of 41.6. These SNPs were primarily associated with polycystic ovarian morphology and ovulatory dysfunction, but not with BMI (Supplementary Table S2.2; data obtained from Day et al.[7]). Only one SNP (rs9696009) reached nominal statistical significance with BMI (p=0.01), though it did not reach genome-wide significance.

Inverse-variance weighted Mendelian randomization analysis with a random effects model did not show a statistically significant association between genetically predicted risk of PCOS and risk of coronary artery disease (OR: 0.99, 95% CI: 0.89;1.11, Q: 18.5) (Figure 2.2 and Supplementary Figure S2.1). MR-Egger regression analysis showed a non-significant intercept (p=0.89). Similar associations were observed when the simple median, penalised weighted median and contamination mixture methods were applied (Figure 2.2). Furthermore, the MR-PRESSO method did not identify any genetic variants that showed significant heterogeneity. Finally, the Steiger-filtering method did not identify any genetic variants that explained significantly more of the variance in the outcome than the exposure trait. Repeat analyses after exclusion of rs9696009 yielded similar results (data not shown).

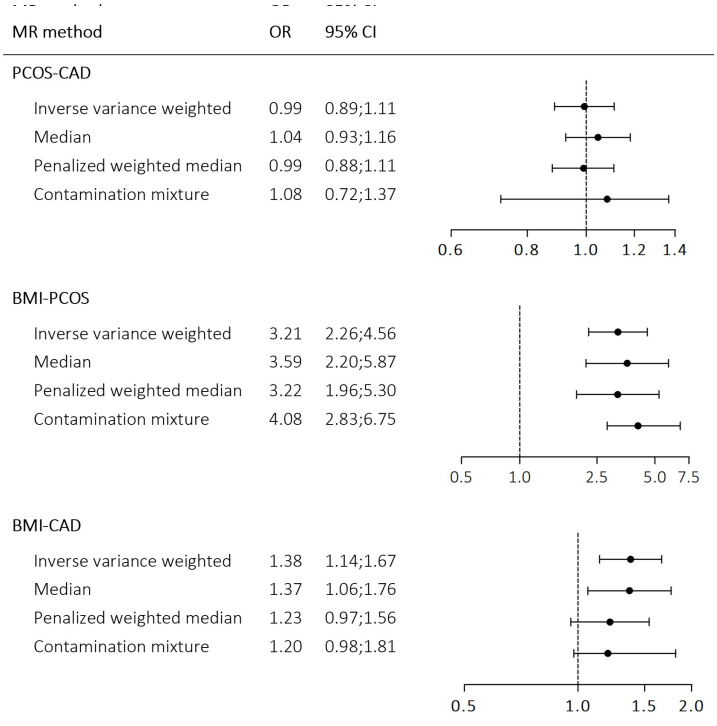

Figure 2.2 Effect estimates of the Mendelian randomization analyses for the association between 1) PCOS and coronary artery disease, 2) BMI and PCOS and 3) BMI and coronary artery disease. Effect estimates are presented as increase in odds of the outcome per unit increase in log(odds) of PCOS, or per standard deviation increase in BMI.
Abbreviations: BMI body mass index; PCOS polycystic ovary syndrome; CAD coronary artery disease; MR Mendelian randomization; OR odds ratio; CI confidence interval

In addition, we repeated the analyses using gene-exposure data for the individual diagnostic criteria of PCOS (i.e. NIH criteria, Rotterdam criteria or self-reported history of PCOS)[7]. The results remained similar for all three diagnostic criteria (IVW OR: 0.99, 95% CI: 0.93;1.07; OR: 1.00, 95% CI: 0.90;1.10; and OR: 1.03, 95% CI: 0.90;1.17, respectively).

Body mass index and polycystic ovary syndrome

The female-specific GWA study identified 38 SNPs that were robustly associated with BMI[10]. One SNP (rs1558902; *FTO*) was palindromic with a minor allele frequency of 0.43 in the gene-outcome data, and, therefore, excluded. The remaining 37 SNPs were used as genetic instruments for BMI used in the association between BMI and PCOS (Supplementary Table S2.3), with a mean F statistic of 55.4.

Inverse-variance weighted Mendelian randomization analysis with a random effects model showed a significant association between genetically predicted BMI and risk of PCOS (OR: 3.21, 95% CI: 2.26;4.56, Q: 38.4) (Figure 2.2 and Supplementary Figure S2.2). The intercept of the MR-Egger regression analysis was not statistically significant (p=0.97). The simple median, penalised weighted median and contamination mixture methods showed comparable effect sizes and were all statistically significant (Figure 2.2). Furthermore, the MR-PRESSO method did not identify any genetic variant that showed significant heterogeneity. The Steiger-filtering method identified one genetic variant (rs2287019) that explained significantly more of the variance in the outcome than the exposure trait, which is suggestive of reverse causality. However, repeat analyses after exclusion of this variant showed similar results (data not shown).

In addition, as the excluded palindromic variant (rs1558902) maps to *FTO*, a very well-known and important obesity gene[20], we repeated the analyses with a proxy, non-palindromic variant (rs1121980) that is in high linkage disequilibrium with the excluded variant (r^2=0.96). The strength and statistical significance of the association remained similar after inclusion of this proxy variant (IVW OR: 3.58, 95% CI: 2.57;4.92, Q: 40.6).

Body mass index and coronary artery disease

None of the 38 SNPs that showed genome-wide significance with BMI matched any of the exclusion criteria (Supplementary Table S2.3), and were, therefore, used as genetic instruments for BMI, with a mean F statistic of 63.8.

Inverse-variance weighted Mendelian randomization analysis with a random effects model showed a significant association between genetically predicted BMI and risk of coronary artery disease (OR: 1.38, 95% CI: 1.14;1.67, Q: 52.1) (Figure 2.2 and Supplementary Figure S2.3). MR-Egger regression analysis showed a non-significant

intercept (p=0.21). The simple median, penalised weighted median and contamination mixture methods resulted in similar effect estimates, although not statistically significant in the latter two (Figure 2.2). The MR-PRESSO method did not identify any genetic variant that showed significant heterogeneity. Finally, the Steiger-filtering method did not identify any genetic variant that explained significantly more of the variance in the outcome than the exposure trait.

Discussion

The aim of this Mendelian randomization study was to examine the triangular association between BMI, PCOS and coronary artery disease (Figure 2.1), all by using female-specific data. We found that genetically predicted risk of PCOS was not associated with an increased risk of coronary artery disease, suggesting that PCOS per se does not play a causal role in the pathogenesis of coronary artery disease. Instead, genetically predicted BMI was associated with an increased risk of both PCOS and coronary artery disease.

The results of our study corroborate with a recent Mendelian randomization study that also failed to show an association between genetically predicted risk of PCOS and risk of coronary artery disease[5]. A serious limitation of that study, however, was the use of publicly available gene-outcome data from the UK Biobank and the Coronary ARtery DIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) plus the Coronary Artery Disease (C4D) Genetics consortium (CARDIoGRAMplusC4D) that were not female-specific. In the present study, we were able to obtain female-specific gene outcome data from the UK Biobank, which allowed us to draw a more valid causal inference. Our findings appear to be in contrast with previous observational studies, which consistently reported that women with PCOS have an approximately twofold increased risk of developing coronary artery disease[2,21,22]. This discrepancy may be explained by the presence of (residual) confounding in the observational studies. Indeed, one meta-analysis reported that adjustment for BMI reduced the strength of the association between PCOS and coronary artery disease[2], suggesting that obesity, at least in part, accounts for both PCOS and coronary artery disease risk.

In support of this hypothesis, we found that genetically predicted BMI was associated with the risk of both PCOS and coronary artery disease. Although previous Mendelian randomization studies examining these associations were not conducted with sex-specific instrumental variables or sex-specific datasets, they reported similar findings[23-25]. Further studies are warranted to unravel the downstream effects of obesity that mediate these relationships. Intrahepatic lipid accumulation – which is a frequently observed phenomenon in obesity[26] – may be one of the denominators of

both PCOS and coronary artery disease risk. Intrahepatic lipid accumulation, more specifically de novo lipogenesis, has been associated with a decrease in serum sex hormone-binding globulin levels[27,28], which has been causally associated with PCOS risk[29]. Furthermore, we have previously shown that genetically predicted intrahepatic lipid accumulation is also associated with coronary artery disease risk, which appears to be mediated by serum lipid levels[30,31]. Intrahepatic lipid accumulation, reduced serum sex hormone-binding globulin levels, and dyslipidemia are commonly observed in patients with PCOS, in particular those who are obese[32,33].

Although both gene-BMI and gene-coronary artery disease data were derived from the general population, the findings in this study support a more personalized approach towards women with PCOS. Since BMI appears the common denominator of both PCOS and coronary artery disease, our results suggest that particularly obese patients with PCOS should be offered counseling about future risk of coronary artery disease upon which preventive measures can be undertaken. Although scientific evidence on the (cost) effectiveness of such a strategy is currently lacking (and will require decades of follow-up), we believe that the clinical presentation of this metabolic disorder relatively early in life offers opportunities to prevent cardiometabolic complications in the sixth decade and onwards.

This study has several strengths and limitations. First, as already mentioned, by using large-scale, female-specific datasets, we were able to draw valid conclusions on the relationships between PCOS and coronary artery disease, and the role of BMI herein. Second, an advantage of the current study design is that, where in traditional epidemiological research a long follow-up is required to obtain sufficient coronary artery disease cases in a cohort of PCOS patients, a two-sample Mendelian randomization allows for gene-outcome data to be retrieved from a cohort of older individuals with significantly more coronary artery disease cases. Of note, the participants from all currently used cohorts were primarily of European descent (Table 2.1). One limitation is the relatively small number of PCOS SNPs that have been identified and, hence, could be used as instrumental variables. This could have restricted statistical power, and it is, therefore, advisable to repeat the current analyses once additional PCOS SNPs have been identified. Furthermore, these SNPs have primarily been associated with two of the three PCOS features, i.e. polycystic ovarian morphology and ovulatory dysfunction, but to a lesser extent with hyperandrogenism[7], and, consequently, may represent only a subset of the PCOS phenotype. Although sensitivity analyses with the different diagnostic criteria of PCOS likewise did not identify a statistically significant association between genetically predicted PCOS and coronary artery disease, it would be relevant to further study the effects of PCOS sub-phenotypes on coronary artery disease if GWA studies for the different features of PCOS become in available in the future.

In conclusion, in this female-specific Mendelian randomization study we did not observe an association between genetically predicted PCOS and risk of coronary artery disease, suggesting that PCOS per se is not causal in the pathogenesis of coronary artery disease. Rather, obesity appears to be the common denominator of both PCOS and coronary artery disease.

References

1. Wolf WM, Wattick RA, Kinkade ON, Olfert MD. Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity. *International Journal of Environmental Research and Public Health.* 2018;15(11):2589.
2. de Groot PC, Dekkers OM, Romijn JA, Dieben SW, Helmerhorst FM. PCOS, coronary heart disease, stroke and the influence of obesity: A systematic review and meta-analysis. *Human Reproduction Update.* 2011;17(4):495-500.
3. Lim SS, Davies MJ, Norman RJ, Moran LJ. Overweight, obesity and central obesity in women with polycystic ovary syndrome: A systematic review and meta-analysis. *Human Reproduction Update.* 2012;18(6):618-637.
4. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Human Molecular Genetics.* 2018;27(R2):R195-R208.
5. Zhu T, Cui J, Goodarzi MO. Polycystic ovary syndrome and risk of type 2 diabetes, coronary heart disease, and stroke. *Diabetes.* 2020.
6. Liu LY, Schaub MA, Sirota M, Butte AJ. Sex differences in disease risk from reported genome-wide association study findings. *Human Genetics.* 2012;131(3):353-364.
7. Day F, Karaderi T, Jones MR, et al. Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. *Plos Genetics.* 2018;14(12):e1007813.
8. Bowden J, Del Greco M F, Minelli C, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *International Journal of Epidemiology.* 2019;48(3):728-742.
9. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. *International Journal of Epidemiology.* 2011;40(3):755-764.
10. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature.* 2015;518(7538):197-206.
11. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine.* 2015;12(3):e1001779.
12. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiology.* 2017;28(1):30-42.
13. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genetic Epidemiology.* 2016;40(4):304-314.
14. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nature Communications.* 2020;11(1):376.
15. Slob EAW, Burgess S. A comparison of robust Mendelian randomization methods using summary data. *Genetic Epidemiology.* 2020;44(4):313-329.
16. 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.* 2018;50(5):693-698.
17. Hemani G, Bowden J, Haycock P, et al. Automating Mendelian randomization through machine learning to construct a putative causal map of the human phenome. *bioRxiv.* 2017.
18. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife.* 2018;7(e34408).
19. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *International Journal of Epidemiology.* 2017;46(6):1734-1739.
20. Dina C, Meyre D, Gallina S, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nature Genetics.* 2007;39(6):724-726.

21. Zhang J, Xu J-H, Qu Q-Q, Zhong G-Q. Risk of cardiovascular and cerebrovascular events in polycystic ovarian syndrome women: A meta-analysis of cohort studies. *Frontiers in Cardiovascular Medicine.* 2020;7:552421-552421.
22. Ramezani Tehrani F, Amiri M, Behboudi-Gandevani S, Bidhendi-Yarandi R, Carmina E. Cardiovascular events among reproductive and menopausal age women with polycystic ovary syndrome: a systematic review and meta-analysis. *Gynecological Endocrinology.* 2020;36(1):12-23.
23. Brower MA, Hai Y, Jones MR, et al. Bidirectional Mendelian randomization to explore the causal relationships between body mass index and polycystic ovary syndrome. *Human Reproduction.* 2018;34(1):127-136.
24. Larsson SC, Bäck M, Rees JMB, Mason AM, Burgess S. Body mass index and body composition in relation to 14 cardiovascular conditions in UK Biobank: a Mendelian randomization study. *European Heart Journal.* 2020;41(2):221-226.
25. Zhao Y, Xu Y, Wang X, et al. Body mass index and polycystic ovary syndrome: A 2-sample bidirectional Mendelian randomization study. *Journal of Clinical Endocrinology and Metabolism.* 2020;105(6).
26. Fabbrini E, Sullivan S, Klein S. Obesity and nonalcoholic fatty liver disease: Biochemical, metabolic, and clinical implications. *Hepatology.* 2010;51(2):679-689.
27. Selva DM, Hogeveen KN, Innis SM, Hammond GL. Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. *Journal of Clinical Investigation.* 2007;117(12):3979-3987.
28. Telgenkamp I, Kusters Y, Schalkwijk CG, et al. Contribution of liver fat to weight loss-induced changes in serum hepatokines: A randomized controlled trial. *Journal of Clinical Endocrinology and Metabolism*. 2019;104(7):2719-2727.
29. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nature Medicine.* 2020;26(2):252-258.
30. Brouwers MCGJ, Simons N, Stehouwer CDA, et al. Relationship between nonalcoholic fatty liver disease susceptibility genes and coronary artery disease. *Hepatology Communications.* 2019;3(4):587-596.
31. Brouwers MCGJ, Simons N, Stehouwer CDA, Isaacs A. Non-alcoholic fatty liver disease and cardiovascular disease: assessing the evidence for causality. *Diabetologia.* 2020;63(2):253-260.
32. Cerda C, Perez-Ayuso RM, Riquelme A, et al. Nonalcoholic fatty liver disease in women with polycystic ovary syndrome. *Journal of Hepatology.* 2007;47(3):412-417.
33. Lim SS, Norman RJ, Davies MJ, Moran LJ. The effect of obesity on polycystic ovary syndrome: a systematic review and meta-analysis. *Obesity Reviews.* 2013;14(2):95-109.

Supplementary materials

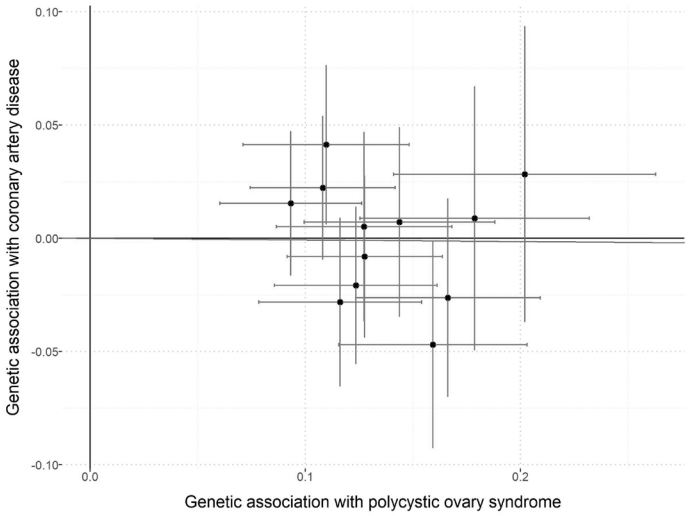

Figure S2.1 Associations between genetic risk of polycystic ovary syndrome (PCOS) and coronary artery disease. The fitted line represents the effect estimate of the inverse variance weighted method.

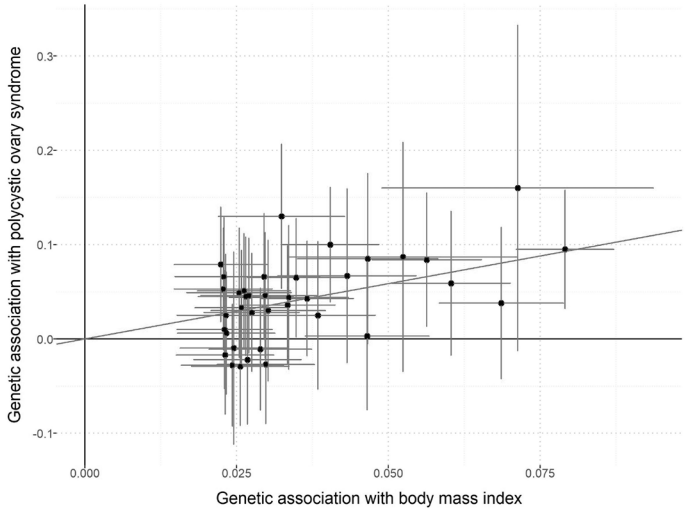

Figure S2.2 Associations between genetic risk of body mass index (BMI) and polycystic ovary syndrome (PCOS). The fitted line represents the effect estimate of the inverse variance weighted method.

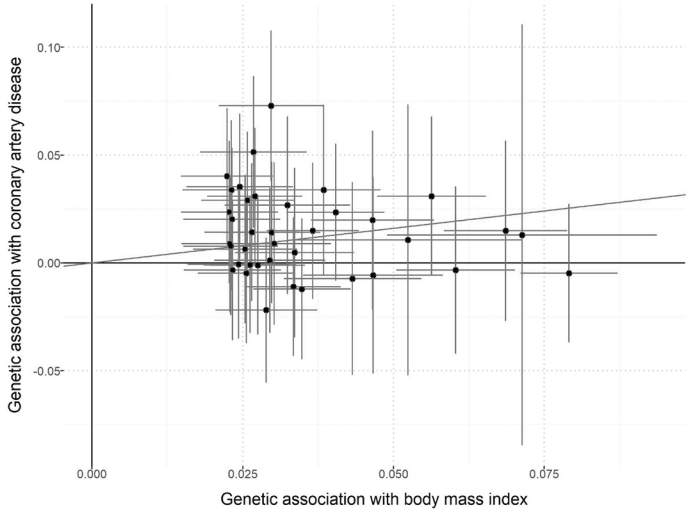

Figure S2.3 Associations between genetic risk of body mass index (BMI) and coronary artery disease. The fitted line represents the effect estimate of the inverse variance weighted method.

Table S2.1 PCOS SNPs that were used as instrumental variables.

SNP	Chr:position*	Nearest gene	Effect allele	Other allele	EAF	β	SE	P-value
rs7563201	2:43334641	*THADA*	G	A	0.5493	0.1081	0.0172	3.678E-10
rs2178575	2:212527042	*ERBB4*	A	G	0.1512	0.1663	0.0219	3.344E-14
rs13164856	5:132477512	*IRF1/RAD50*	T	C	0.7291	0.1235	0.0193	1.453E-10
rs804279	8:11766380	*GATA4/NEIL2*	A	T	0.2616	0.1276	0.0184	3.761E-12
rs10739076	9:5440589	*PLGRKT*	A	C	0.3078	0.1097	0.0197	2.510E-08
rs7864171	9:94960984	*C9orf3*	G	A	0.5716	0.0933	0.0168	2.946E-08
rs9696009	9:123856954	*DENND1A*	A	G	0.0679	0.202	0.0311	7.958E-11
rs11031005	11:30204809	*ARL14EP/FSHB*	C	T	0.1463	0.1593	0.0223	8.664E-13
rs11225154	11:102172509	*YAP1*	A	G	0.0941	0.1787	0.0272	5.438E-11
rs1784692	11:114078510	*ZBTB16*	T	C	0.8237	0.1438	0.0226	1.876E-10
rs1275468	12:75541377	*KRR1*	C	T	0.7554	0.1162	0.0193	1.868E-09
rs8043701	16:52341865	*TOX3*	T	A	0.8150	0.1273	0.0208	9.610E-10

Data retrieved from [1]. Beta is presented as the log(odds) risk of polycystic ovary syndrome per additional copy of the effect allele. * Chromosome:position in GRCh38.p13 reference genome assembly.
Abbreviations: EAF effect allele frequency; SE standard error; SNP single nucleotide polymorphism

Table S2.2 Effect of PCOS SNPs on polycystic ovarian morphology, ovulatory dysfunction and BMI.

SNP	Nearest gene	Effect allele	Other allele	PCOM		OD		HA		BMI	
				β	P-value	β	P-value	β	P-value	β	P-value
rs7563201	*THADA*	G	A	0.16	3.7E-04	0.07	1.5E-03	0.061	8.0E-02	*	*
rs2178575	*ERBB4*	A	G	0.24	1.4E-05	0.23	1.2E-11	-0.126	4.3E-03	-0.0054	4.1E-01
rs13164856	*IRF1/RAD50*	T	C	0.16	1.4E-03	0.08	5.6E-03	0.092	1.8E-02	-0.0037	4.9E-01
rs804279	*GATA4/NEIL2*	A	T	0.22	1.5E-06	0.16	9.9E-09	0.126	8.7E-04	-0.0041	4.6E-01
rs10739076	*PLGRKT*	A	C	0.10	5.9E-02	0.00	8.9E-01	0.026	5.3E-01	*	*
rs7864171	*C9orf3*	G	A	0.19	1.3E-05	0.10	2.3E-04	0.124	3.8E-04	*	*
rs9696009	*DENND1A*	A	G	0.32	4.0E-05	0.36	4.4E-15	-0.330	2.9E-07	0.0239	1.2E-02
rs11031005	*ARL14EP/FSHB*	C	T	0.18	1.3E-03	0.13	2.8E-04	-0.079	8.2E-02	0.0013	8.5E-01
rs11225154	*YAP1*	A	G	0.24	3.5E-04	0.23	5.7E-08	-0.144	1.4E-02	*	*
rs1784692	*ZBTB16*	T	C	0.30	2.8E-06	0.21	6.6E-09	0.146	4.6E-03	0.0049	4.5E-01
rs1275468	*KRR1*	C	T	0.16	1.5E-03	0.11	1.8E-04	-0.104	8.0E-02	0.0014	8.0E-01
rs8043701	*TOX3*	T	A	0.17	1.5E-03	0.08	9.2E-03	-0.166	1.4E-04	*	*

Data for the association between SNP and polycystic ovarian morphology, ovulatory dysfunction and hyperandrogenism retrieved from [7] and for body mass index from [2]. * SNP was not reported in the genome wide association study of BMI.

Abbreviations: BMI body mass index; HA hyperandrogenism; OD ovulatory dysfunction; PCOM polycystic ovarian morphology; SNP single nucleotide polymorphism

Table S2.3 BMI SNPs that were used as instrumental variables

SNP	Chr:Position*	Nearest gene	Effect allele	Other allele	EAF	β	SE	P-value
rs543874	1:77920345	*SEC16B*	G	A	0.1939	0.0603	0.0050	4.03E-84
rs3101336	1:72285502	*NEGR1*	C	T	0.6108	0.0334	0.0040	8.91E-17
rs12566985	1:74536509	*FPGT*	G	A	0.4444	0.0270	0.0040	1.05E-11
rs17024393	1:109612066	*GNAT2*	C	T	0.0390	0.0713	0.0114	3.53E-10
rs11165643	1:96458541	*PTBP2*	T	C	0.5831	0.0230	0.0040	5.76E-09
rs12401738	1:77981077	*FUBP1*	A	G	0.3548	0.0256	0.0041	6.09E-10
rs657452	1:49124175	*AGBL4*	A	G	0.3940	0.0233	0.0041	1.70E-08
rs13021737	2:632348	*TMEM18*	G	A	0.8280	0.0686	0.0052	6.99E-40
rs10182181	2:24927427	*ADCY3*	G	A	0.4604	0.0366	0.0039	3.91E-21
rs1016287	2:59078490	*LINC01122*	T	C	0.2853	0.0254	0.0044	5.87E-09
rs1516725	3:186106215	*ETV5*	C	T	0.8703	0.0466	0.0059	1.91E-15
rs13078960	3:85758440	*CADM2*	G	T	0.1973	0.0336	0.0050	1.41E-11
rs16851483	3:141556594	*RASA2*	T	G	0.0654	0.0524	0.0096	4.81E-08
rs10938397	4:45180510	*GNPDA2*	G	A	0.4321	0.0404	0.0041	2.98E-23
rs2112347	5:75719417	*POC5*	T	G	0.6306	0.0298	0.0041	3.15E-13
rs2207139	6:50877777	*TFAP2B*	G	A	0.1769	0.0465	0.0052	2.40E-19
rs205262	6:34595387	*C6orf106*	G	A	0.2722	0.0268	0.0045	2.04E-09
rs6465468	7:95540202	*ASB4*	T	G	0.3061	0.0245	0.0045	4.98E-08
rs17405819	8:75894349	*HNF4G*	T	C	0.6994	0.0243	0.0043	1.45E-08
rs10968576	9:28414341	*LINGO2*	G	A	0.3194	0.0289	0.0043	1.04E-11
rs1928295	9:117616205	*TLR4*	T	C	0.5488	0.0258	0.0039	3.42E-11
rs10733682	9:126698635	*LMX1B*	A	G	0.4772	0.0229	0.0041	1.67E-08
rs11030104	11:27662970	*BDNF*	A	G	0.7920	0.0384	0.0048	2.32E-15
rs3817334	11:47629441	*MTCH2*	T	C	0.4088	0.0265	0.0040	2.47E-11
rs4256980	11:8652392	*TRIM66*	G	C	0.6470	0.0232	0.0041	1.65E-08
rs7138803	12:49853685	*BCDIN3D*	A	G	0.3838	0.0348	0.0041	1.80E-17
rs7141420	14:79433111	*NRXN3*	T	C	0.5255	0.0262	0.0039	1.45E-11
rs12885454	14:29267632	*PRKD1*	C	A	0.6391	0.0228	0.0041	2.89E-08
rs16951275	15:67784830	*MAP2K5*	T	C	0.7827	0.0302	0.0048	2.96E-10
rs3888190	16:28878165	*ATP2A1*	A	C	0.4029	0.0275	0.0040	3.51E-12
rs12446632	16:19924067	*GPRC5B*	G	A	0.8648	0.0432	0.0058	1.45E-13
rs1558902†	16:53769662	*FTO*	A	T	0.4127	0.0791	0.0041	4.03E-84
rs6567160	18:60161902	*MC4R*	C	T	0.2372	0.0563	0.0046	5.08E-34
rs1808579	18:23524924	*C18orf8*	C	T	0.5337	0.0224	0.0039	1.23E-08
rs7239883	18:42567706	*LOC284260*	G	A	0.3911	0.0231	0.0041	1.51E-08
rs2287019	19:45698914	*QPCTL*	C	T	0.8036	0.0324	0.0053	8.60E-10
rs3810291	19:47065746	*ZC3H4*	A	G	0.6656	0.0295	0.0047	2.60E-10
rs6091540	20:52471323	*ZFP64*	C	T	0.7205	0.0297	0.0044	2.15E-11

Data retrieved from [2]. Beta is presented as standard deviation change in body mass index per additional copy of the effect allele. * Chromosome:position in GRCh38.p13 reference genome assembly. † This SNP was not included as an instrumental variable in the association between BMI and PCOS.
Abbreviations: EAF effect allele frequency; SE standard error; SNP single nucleotide polymorphism

References

1. Day F., Karaderi T., Jones M.R., et al. Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. *PLoS Genet.* 2018;14(12):e1007813.
2. Locke A.E., Kahali B., Berndt S.I., et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature.* 2015;518(7538):197-206.

3

CHAPTER THREE

The relationships of sex hormone-binding globulin, total testosterone, androstenedione and free testosterone with metabolic and reproductive features of polycystic ovary syndrome

Pomme I.H.G. Simons, Olivier Valkenburg, Judith A.P. Bons, Coen D.A. Stehouwer, Martijn C.G.J. Brouwers

Endocrinology, Diabetes & Metabolism (2021); 4(3): e00267

Summary

Objective: A recent Mendelian randomization study has suggested a causal role for sex hormone-binding globulin (SHBG), total testosterone and free testosterone in the pathogenesis of polycystic ovary syndrome (PCOS). The aim of this study was to assess the relationships of SHBG, androstenedione, total and free testosterone with the individual metabolic and reproductive features of PCOS.

Design: Cross-sectional data in PCOS patients (n=96) prospectively collected in a secondary/tertiary clinic for menstrual cycle disorders.

Methods: Multivariable regression analyses were conducted to study the associations between SHBG, androstenedione, total and free testosterone with metabolic (BMI, waist circumference, systolic and diastolic blood pressure, total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, and homeostatic model assessment for insulin resistance [HOMA2-IR]), and reproductive features (menstrual cycle length, antral follicle count, anti-Müllerian hormone, luteinizing hormone, follicle-stimulating hormone and Ferriman-Gallwey score) of PCOS.

Results: Serum SHBG and free testosterone, but not total testosterone or androstenedione, were significantly associated with BMI, waist circumference, serum triglycerides, HDL cholesterol, LDL cholesterol and HOMA2-IR. The strength of the associations with serum lipids was reduced after adjustment for BMI, but not for HOMA2-IR. Total testosterone was significantly associated with antral follicle count. SHBG, total testosterone, and androstenedione were significantly associated with serum AMH. Only the strength of the association for SHBG was reduced after adjustment for BMI.

Conclusions: Serum SHBG is associated with primarily metabolic features, whereas total testosterone and androstenedione are associated with reproductive features of PCOS. These results suggest a differential underlying pathophysiology for the metabolic and reproductive features of PCOS.

Introduction

Polycystic ovary syndrome (PCOS) is the most common endocrine disorder among premenopausal women, with an estimated prevalence of 10%[1]. Many, but not all women with PCOS exhibit metabolic disturbances, including obesity, insulin resistance, hypertension and dyslipidaemia, and endocrine abnormalities such as an increased ratio of luteinizing hormone (LH) to follicle-stimulating hormone (FSH), and increased anti-Müllerian hormone (AMH) levels[2].

Androgen excess is a fundamental, diagnostic feature of PCOS that is present in approximately half to three quarters of PCOS patients[3,4]. A recent Mendelian randomization study has shown that genetically predicted SHBG, total testosterone and free testosterone levels are associated with the risk of PCOS[5]. This establishes a potential causal role of free testosterone, and its determinants, in the pathogenesis of PCOS. Nevertheless, there is an ongoing discussion on the role of free testosterone in the development of the individual features of PCOS.

The role of androgens in the pathophysiology of PCOS is likely a complex, multifactorial process, driven by genetics, hormonal imbalance and lifestyle factors[5-7]. Some have speculated that free testosterone plays a central role in the pathogenesis of all PCOS features, by actively contributing to the arrest of follicular development, theca cell dysfunction, ovarian stromal hyperplasia, abnormal gonadotrophin-releasing hormone (GnRH) secretion, and insulin resistance[2,8-10]. Others have argued that hyperandrogenism is merely a consequence of metabolic dysfunction or ovarian and endocrine changes, and does not, in itself, contribute to the pathophysiology of PCOS[11-14]. SHBG may reflect primarily metabolic changes, while total testosterone and androstenedione may reflect reproductive dysfunction[14].

The aim of the present study is, therefore, to study the associations of serum SHBG, androstenedione, total testosterone and free testosterone with the individual metabolic and reproductive features of PCOS.

Methods

Study population

Data were prospectively collected at the outpatient clinic for menstrual cycle disorders of the Maastricht University Medical Centre (Maastricht, the Netherlands) between March 2017 and February 2020. PCOS was retrospectively diagnosed according to the Rotterdam criteria, which requires the presence of at least two of the following three

characteristics: irregular menstrual cycle, hyperandrogenism and polycystic ovarian morphology (PCOM)[15]. Irregular menstrual cycle was defined as a menstrual cycle length ≥35 days. Hyperandrogenism was defined as a free androgen index (total testosterone (nmol/l) * 100 / SHBG (nmol/l)) >4.5[16], total testosterone >1.9 nmol/l, androstenedione >9.6 nmol/l, or a Ferriman-Gallwey score ≥4 for women of Caucasian, Black or Mixed ethnicity, and ≥6 for women of Middle-Eastern and Asian ethnicity[17-19]. PCOM was defined as the presence of ≥20 follicles (2-9 mm in diameter) in either ovary or an ovarian volume ≥10 ml, according to the revised international evidence-based PCOS guidelines (European Society of Human Reproduction and Embryology [ESHRE] guidelines, 2018)[17]. Women who were pregnant, used hormonal contraceptives at the time of the clinical assessment, had abnormal thyroid stimulating hormone (TSH) levels (Table 3.1), elevated prolactin levels (Table 3.1), or individuals diagnosed with non-classic congenital adrenal hyperplasia, were excluded from the current study.

This study was approved by the Medical Ethics Committee of Maastricht University Medical Centre.

Table 3.1 General characteristics of the study population.

	PCOS population (n=96)	Reference interval[†]
Age,years	28.4 ± 4.2	
Ethnicity, n (%)		
Caucasian	88 (92)	
Black	2 (2)	
Middle-Eastern	3 (3)	
Asian	2 (2)	
Mixed	1 (1)	
Smoking, cigarettes/day	0.0 (0.0-0.0)	
Alcohol, units/week	0.0 (0.0-2.0)	
Fasting, yes (%)	89 (93)	
TSH, mU/l	1.8 (1.5-2.4)	0.4-4.3
Prolactin, U/l	0.23 ± 0.09	0.10-0.64[‡] 0.01-0.50[‡]
Metabolic features		
BMI, kg/m^2	26.0 (22.3-33.1)	
Waist circumference, cm	84.0 (75.0-101.0)	
Systolic blood pressure, mmHg	122 ± 13	
Diastolic blood pressure, mmHg	75 ± 10	
Glucose, mmol/l	4.9 ± 0.5	3.1-6.1
Insulin, pmol/l	39.8 (16.8-64.3)	12-150
HOMA2-IR	0.7 (0.3-1.2)	
Triglycerides, mmol/l	0.8 (0.6-1.2)	0.9-1.94
Total cholesterol, mmol/l	4.6 ± 0.9	<5.0
HDL cholesterol. mmol/l	1.6 ± 0.4	>0.9
LDL cholesterol, mmol/l	2.9 ± 0.9	<2.5
Metabolic syndrome, yes (%)	15 (16)	

Table 3.1 (continued)

	PCOS population (n=96)	Reference interval†
Reproductive features		
Average length of menstrual cycle, days	51 (40-96)	
Regularity of menstrual cycle, n (%)		
Regular	3 (3)	
Oligomenorrhoe	72 (75)	
Amenorrhoe	19 (20)	
Metrorrhagia	2 (2)	
AMH, µg/l	7.6 (4.8-11.1)	<6.9
Antral follicle count§	20.8 ± 7.8	
Ovarian volume§	8.0 (6.5-10.7)	
PCOM, yes (%)	73 (76)	
FSH, U/l	5.6 ± 2.3	Follicular phase: 2.8-14.4 Ovulatory phase: 5.8-21.0 Luteal phase: 1.2-9.0
LH, U/l	7.7 (5.0-11.4)	Follicular phase: 1.1-11.6 Ovulatory phase: 17.0-77.0 Luteal phase: <0.05-14.7
Total testosterone, nmol/l	1.7 ± 0.7	0.3-1.9
SHBG, nmol/l	42 (28-63)	40-120
Free testosterone. pmol/l	21.1 (14.3-29.8)	3.5-24
Free androgen index	3.5 (2.4-6.0)	
Androstenedione, nmol/l	11.8 ± 4.1	3.0-9.6
Hyperandrogenism, yes (%)¶	77 (80)	
Ferriman Gallwey score	5 (1-9)	
Hirsutism, yes (%)	27 (28)	
Self-reported acne, yes (%)	44 (46)	

† Reference intervals according to the Central Diagnostic Laboratory at the Maastricht University Medical Centre (the Netherlands). ‡ Reference intervals prior to November 2018 and after November 2018, respectively; see methods section. § Average of both ovaries. ¶ Biochemical or clinical (according to the Ferriman-Gallwey score) hyperandrogenism

Abbreviations: BMI body mass index; HOMA2-IR homeostatic model assessment for insulin resistance; HDL high-density lipoprotein; LDL low-density lipoprotein; AMH anti-Müllerian Hormone; PCOM polycystic ovarian morphology; FSH follicle stimulating hormone; LH luteinizing hormone; SHBG sex hormone-binding globulin; TSH thyroid stimulating hormone

Clinical assessment

All patients filled out questionnaires regarding demographics (age and ethnicity), lifestyle (smoking status and alcohol consumption), self-reported history of acne and hirsutism (defined according to the aforementioned Ferriman-Gallwey score cut-off values) and gynaecological history (length and regularity of menstrual cycle). A regular menstrual cycle was defined as a menstrual cycle <35 days, oligomenorrhea as a menstrual cycle ≥35 days, amenorrhea as no menstrual period during the prior six months, and metrorrhagia as vaginal bleeding at irregular intervals.

Physical examination was performed to determine body mass index (BMI; calculated as body weight [kilograms] divided by length [meters] squared), waist circumference (at the level of the umbilicus), and systolic and diastolic blood pressure measured in semi-seated position after 10 minutes of rest with an Omron 705IT automated measuring device. A transvaginal ultrasound was performed to count the total number of antral follicles (2-9 mm in diameter) in each ovary and calculate the ovarian volume (as 0.523*length*width*depth for each ovary[20]), which were subsequently expressed as the average of two ovaries. In four cases, an abdominal ultrasound was performed instead and, where possible, ovarian volume and antral follicle count were assessed.

Blood was drawn in the morning. Patients were asked to visit the outpatient clinic after an overnight fast. Laboratory analyses were performed by the Central Diagnostic Laboratory at the Maastricht University Medical Centre (the Netherlands). All reference intervals were locally established by the Central Diagnostic Laboratory. Total testosterone and TSH were measured with an electrochemiluminescence immunoassay (Cobas 8000 instrument, Roche Diagnostics, Mannheim, Germany); FSH, LH, SHBG, and insulin with an chemiluminescent immunometric assay (Immulite XPi instrument, Siemens Healthcare Diagnostics, New Orleans, LA, USA); serum glucose with an enzymatic spectrophotometric assay (Cobas 8000 instrument, Roche Diagnostics, Mannheim, Germany); triglycerides, total cholesterol and HDL cholesterol with an enzymatic colorimetric assay (Cobas 8000 instrument, Roche Diagnostics, Mannheim, Germany); and androstenedione with a radio immunoassay (IBL International, Hamburg, Germany). Prolactin was measured with electrochemiluminescence immunoassay (Cobas 8000 instrument, Roche Diagnostics, Mannheim, Germany) until November 2018, and with immunoassay (AutoDelfia, Perkin Elmer, Turku, Finland) after this date. AMH was measured with an enzyme-linked immunosorbent assay (Gen II, Beckman Coulter, Brea, CA, USA) until July 2019, and with a chemiluminescent immunometric assay (Lumipulse G1200, Fujirebio, Tokyo, Japan) after this date. AMH levels determined by the enzyme-linked immunosorbent assay were multiplied with a correction factor of 0.88 to obtain chemiluminescent immunometric assay calibrated AMH values[21]. Free testosterone was calculated using the Ross algorithm[22]. LDL cholesterol was calculated using the Friedewald formula[23]. The homeostasis model assessment 2 (HOMA2-IR) was calculated as a measure of insulin resistance (available at http://www.dtu.ox.ac.uk/homacalculator/). The metabolic syndrome was defined as the presence of at least three of the following five characteristics: a waist circumference ≥88 cm, triglycerides ≥1.7 mmol/l, HDL cholesterol <1.3 mmol/l, blood pressure ≥130 / ≥85 mmHg, and a fasting glucose ≥6.1 mmol/l[24].

Statistical analyses

Continuous data are presented as mean ± standard deviation (SD) or as median (interquartile range [IQR]) in case of non-normal distribution. Categorical data are presented as frequencies. Non-normally distributed variables were log-transformed before further analyses. Multivariable linear regression analyses were performed to study the associations of SHBG, androstenedione, total testosterone and free testosterone with metabolic (BMI, waist circumference, systolic and diastolic blood pressure, total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, and HOMA2-IR), and reproductive features (length of menstrual cycle, antral follicle count, AMH, LH, FSH, and Ferriman-Gallwey score) of PCOS, independent of potential confounders. Z-scores (= individual value minus population mean, divided by population SD) were calculated for SHBG, androstenedione, total testosterone and free testosterone before entry into the model to allow comparison of the strength of association between these variables. Since not all patients visited the outpatient clinic in the fasting state, analyses for most metabolic characteristics were adjusted for fasting (yes/no). Additional adjustments were made for age, BMI and HOMA2-IR, for those metabolic and reproductive features that showed a statistically significant association with any of the androgen markers. Sensitivity analyses were conducted in fasted individuals only. All results were considered statistically significant at $p<0.05$. All statistical analyses were performed using IBM Statistical Package of Social Science (SPSS) version 25.0 for Windows (IBM Corp., Armonk, N.Y., USA).

Results

Study population

Between March 2017 and February 2020, we retrospectively identified 111 women who fulfilled the diagnostic criteria for PCOS. Fifteen individuals were excluded because they were pregnant (n=3), used hormonal contraceptives (n=9), or had elevated prolactin levels (n=3) at the time of the clinical assessment. The general characteristics of the patients with PCOS (n=96) are presented in Table 3.1. On average, the study population was young (mean age: 28.4 ± 4.2 years) and overweight (median BMI: 26.0, IQR: 22.3-33.1 kg/m^2). Due to the specialized outpatient setting, the majority of women experienced oligomenorrhea (75%) or amenorrhea (20%). Additionally, the majority of women (76%) were found to have PCOM on ultrasound examination. Finally, 28% of women suffered from hirsutism and 46% reported a history of acne. Only a small percentage of women (16%) met the criteria for the metabolic syndrome.

Associations of serum SHBG, androstenedione, total testosterone and free testosterone with metabolic features of PCOS

Figure 3.1 shows the associations of SHBG, androstenedione, total testosterone and free testosterone with nine metabolic features of PCOS. Serum SHBG and free testosterone, but not total testosterone or androstenedione, were associated with BMI (Figure 3.1, panel A). Although similar patterns were observed for all other metabolic characteristics (Figure 3.1, panel B-I), statistical significance was reached for the relationship of both SHBG and free testosterone with waist circumference, LDL cholesterol, HDL cholesterol, serum triglycerides and HOMA2-IR (Figure 3.1, panel B and panel F-I). Adjustment for age did not materially alter the strength of the statistically significant associations (Table 3.2). In contrast, additional adjustment for BMI reduced the strengths of all associations, whereas addition of HOMA2-IR to the regression models did not have a substantial effect (Table 3.2). The strengths of associations did not substantially change when repeating the analyses in fasted individuals only (n=89) (Supplementary Figure S3.1 and Supplementary Table S3.1).

Associations of serum SHBG, androstenedione, total testosterone and free testosterone with reproductive features of PCOS

Figure 3.2 shows the relationships of SHBG, androstenedione, total testosterone, and free testosterone with six reproductive features of PCOS. None of these were associated with menstrual cycle length (Figure 3.2, panel A). Total testosterone, but not SHBG, was statistically significantly associated with antral follicle count (Figure 3.2, panel B), which was not affected by adjustment for age and BMI (Table 3.2). The strength of association was reduced and no longer statistically significant after further adjustment for HOMA2-IR (Table 3.2). Androstenedione and SHBG were significantly associated with serum AMH (Figure 3.2, panel C). The significant association between SHBG and serum AMH was lost after adjustment for age and BMI (Table 3.2). Total testosterone was significantly associated with serum AMH after adjustment for age and BMI (Table 3.2). No significant associations were observed for serum LH and FSH (Figure 3.2, panel D and E, respectively). Finally, although the direction of the associations of serum SHBG, androstenedione, total testosterone and free testosterone with the Ferriman-Gallwey score were as anticipated, i.e. inverse for SHBG and positive for androstenedione, total testosterone, and free testosterone, none of these associations were statistically significant (Figure 3.2, panel F). The strengths of associations did not substantially change when repeating the analyses in fasted individuals only (n=89) (Supplementary Figure S3.2 and Supplementary Table S3.1).

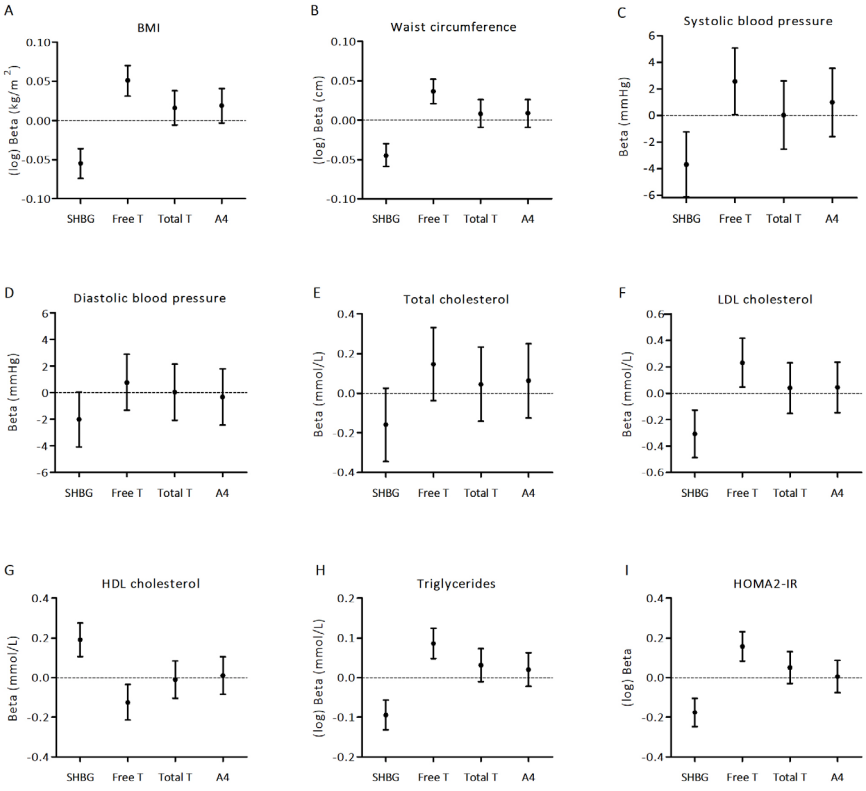

Figure 3.1 Associations of serum sex hormone-binding globulin (SHBG), free testosterone (Free T), total testosterone (Total T), and androstenedione (A4) with metabolic features of polycystic ovary syndrome: BMI (n=96) (A) waist circumference (n=95) (B) systolic blood pressure (n=96) (C) diastolic blood pressure (n=96) (D) total cholesterol (n=93) (E) LDL cholesterol (n=93) (F) HDL cholesterol (n=93) (G) triglycerides (n=93) (H) and homeostatic model assessment of insulin resistance (HOMA2-IR) (n=92) (I). Analyses were conducted with Z-scores to allow comparison. Regression coefficients should therefore be interpreted as the increase in the dependent variable per standard deviation increase in serum SHBG, free testosterone, total testosterone or androstenedione (after adjustment for fasting (yes/no), panel C-I). See methods section.

Table 3.2 Associations of serum SHBG, free testosterone, total testosterone and androstenedione with metabolic and reproductive features of PCOS.

Independent variables	SHBG B (95% CI)	Free testosterone B (95% CI)	Total testosterone B (95% CI)	Androstenedione B (95% CI)
Metabolic features				
(log) BMI				
Crude	**-0.06 (-0.08;-0.04)**	**0.05 (0.03;0.07)**	0.02 (0.00;0.04)	0.02 (0.00;0.04)
Age	**-0.06 (-0.07;-0.04)**	**0.05 (0.03;0.07)**	0.02 (-0.01;0.04)	0.02 (-0.01;0.04)
(log) Waist circumference				
Crude	**-0.05 (-0.06;-0.03)**	**0.04 (0.02;0.05)**	0.01 (-0.01;0.03)	0.01 (-0.01;0.03)
Age	**-0.05 (-0.06;-0.03)**	**0.04 (0.02;0.05)**	0.01 (-0.01;0.03)	0.01 (-0.01;0.03)
Systolic blood pressure				
Crude†	**-3.44 (-6.06;-0.82)**	2.62 (-0.08;5.31)	0.46 (-2.30;3.23)	1.12 (-1.60;3.84)
Age	**-4.00 (-6.65;-1.34)**	**3.29 (0.54;6.05)**	0.71 (-2.07;3.49)	1.86 (-0.99;4.71)
Age, BMI	-0.99 (-3.79;1.82)	0.10 (-2.75;2.94)	-0.49 (-2.97;1.98)	0.74 (-1.81;3.28)
Age, BMI, HOMA2-IR	-0.93 (-3.73;1.88)	-0.30 (-3.23;2.63)	-0.90 (-3.46;1.65)	0.91 (-1.59;3.56)
Diastolic blood pressure				
Crude†	**-2.16 (-4.29;-0.04)**	1.00 (-1.19;3.20)	0.39 (-1.82;2.60)	-0.13 (-2.32;2.05)
Age	**-2.77 (-4.89;-0.66)**	1.70 (-0.51;3.92)	0.71 (-1.49;2.90)	0.63 (-1.63;2.90)
Age, BMI	-1.12 (-3.48;1.24)	-0.28 (-2.67;2.12)	0.00 (-2.08;2.09)	-0.05 (-2.20;2.10)
Age, BMI, HOMA2-IR	-1.03 (-3.36;1.30)	0.86 (-3.30;1.57)	-0.53 (-2.66;1.61)	0.27 (-1.88;2.42)
LDL cholesterol				
Crude†	**-0.31 (-0.49;-0.13)**	**0.25 (0.06;0.44)**	0.07 (-0.13;0.27)	0.07 (-0.13;0.26)
Age	**-0.35 (-0.53;-0.17)**	**0.30 (0.11;0.49)**	0.09 (-0.10;0.28)	0.11 (-0.09;0.31)
Age, BMI	**-0.24 (-0.44;-0.03)**	0.17 (-0.04;0.38)	0.03 (-0.16;0.21)	0.06 (-0.16;0.25)
Age, BMI, HOMA2-IR	**-0.24 (-0.44;-0.03)**	0.16 (-0.05;0.38)	0.01 (-0.18;0.21)	0.07 (-0.13;0.26)
HDL cholesterol				
Crude†	**0.19 (0.11;0.27)**	**-0.13 (-0.22;-0.04)**	-0.02 (-0.11;0.07)	0.00 (-0.09;0.09)
Age	**0.18 (0.09;0.26)**	**-0.11 (-0.20;-0.02)**	-0.01 (-0.10;0.08)	0.04 (-0.06;0.13)
Age, BMI	0.08 (-0.01;0.17)	0.00 (-0.09;0.09)	0.04 (-0.04;0.12)	0.08 (0.00;0.16)
Age, BMI, HOMA2-IR	0.08 (-0.01;0.17)	0.02 (-0.07;0.12)	0.06 (-0.02;0.14)	0.07 (-0.01;0.15)
(log) Triglycerides				
Crude†	**-0.10 (-0.13;-0.06)**	**0.09 (0.05;0.13)**	0.04 (-0.01;0.08)	0.02 (-0.02;0.07)
Age	**-0.10 (-0.14;-0.06)**	**0.09 (0.05;0.13)**	0.04 (-0.01;0.08)	0.02 (-0.02;0.07)
Age, BMI	**-0.07 (-0.11;-0.02)**	**0.06 (0.01;0.10)**	0.02 (-0.02;0.06)	0.00 (-0.04;0.05)
Age, BMI, HOMA2-IR	**-0.07 (-0.11;-0.02)**	**0.06 (0.01;0.10)**	0.02 (-0.03;0.06)	0.01 (-0.04;0.05)
(log) HOMA2-IR				
Crude†	**-0.17 (-0.24;-0.10)**	**0.16 (0.08;0.23)**	0.05 (-0.03;0.13)	0.00 (-0.08;0.08)
Age	**-0.17 (-0.24;-0.10)**	**0.16 (0.08;0.23)**	0.05 (-0.03;0.13)	-0.01 (-0.09;0.07)
Age, BMI	-0.07 (-0.14;0.00)	0.05 (-0.02;0.12)	0.00 (-0.06;0.07)	-0.05 (-0.12;0.01)
Reproductive features				
Antral follicle count				
Crude	0.51 (-1.15;2.17)	1.38 (-0.26;3.02)	**1.99 (0.38;3.59)**	1.12 (-0.53;2.78)
Age	0.83 (-0.85;2.51)	1.10 (-0.60;2.81)	**1.82 (0.19;3.45)**	0.81 (-0.94;2.56)
Age, BMI	0.78 (-1.18;2.74)	1.75 (-0.19;3.68)	**1.97 (0.32;3.62)**	0.93 (-0.85;2.70)
Age, BMI, HOMA2-IR‡	0.86 (-1.07;2.79)	1.28 (-0.72;3.27)	1.58 (-0.15;3.31)	1.25 (-0.52;3.01)
(log) AMH				
Crude	**0.07 (0.01;0.12)**	-0.01 (-0.07;0.04)	0.05 (0.00;0.11)	**0.06 (0.00;0.11)**
Age	**0.08 (0.02;0.13)**	-0.02 (-0.08;0.04)	0.05 (-0.01;0.10)	0.05 (-0.01;0.11)
Age, BMI	0.04 (-0.03;0.10)	0.03 (-0.03;0.09)	**0.07 (0.01;0.12)**	**0.07 (0.02;0.13)**
Age, BMI, HOMA2-IR‡	0.04 (-0.03;0.10)	0.03 (-0.03;0.10)	**0.07 (0.01;0.13)**	**0.07 (0.02;0.13)**

Analyses were conducted with Z-scores. Beta coefficients should therefore be interpreted as per standard deviation increase in serum SHBG, free testosterone, total testosterone or androstenedione. See methods section. Bold values indicate statistical significance ($p<0.05$). † Adjusted for fasting (yes/no) in all models. ‡ Model additionally adjusted for fasting (yes/no).

Abbreviations: AMH anti-Müllerian hormone; BMI body mass index; HOMA2-IR homeostatic model assessment for insulin resistance; HDL high-density lipoprotein; LDL low-density lipoprotein; AMH anti-Müllerian hormone; SHBG sex hormone-binding globulin.

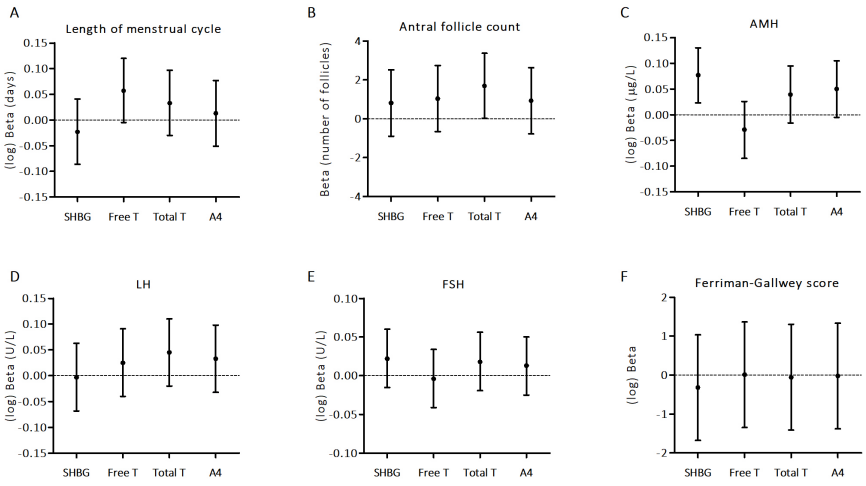

Figure 3.2 Associations of serum sex hormone-binding globulin (SHBG), free testosterone (Free T), total testosterone (Total T), and androstenedione (A4) with reproductive features of polycystic ovary syndrome: length of menstrual cycle (n=89) (A) antral follicle count (n=92) (B) anti-Müllerian hormone (AMH) (n=92) (C) luteinizing hormone (LH) (n=95) (D) follicle stimulating hormone (FSH) (n=96) (E) and Ferriman-Gallwey score (n=96) (F). Analyses were conducted with Z-scores to allow comparison. Regression coefficients should, therefore, be interpreted as the increase in the dependent variable per standard deviation increase in serum SHBG, free testosterone, total testosterone or androstenedione. See methods section.

Discussion

The aim of this study was to examine the associations of SHBG, androstenedione, total testosterone and free testosterone with the individual metabolic and reproductive features of PCOS. Serum SHBG and free testosterone, but not total testosterone or androstenedione, were significantly associated with BMI, waist circumference, serum triglycerides, HDL cholesterol, LDL cholesterol and HOMA2-IR. In contrast, in the adjusted models total testosterone was significantly associated with antral follicle count and serum AMH, while androstenedione was significantly associated with serum AMH. Adjustment for BMI substantially reduced the strength of association of free testosterone and SHBG with the metabolic features of PCOS, but hardly affected the associations of total testosterone or androstenedione with the reproductive features of PCOS.

The observed patterns of associations, i.e. SHBG mainly associates with metabolic features whereas total testosterone and androstenedione associate with reproductive abnormalities of PCOS, are in line with previous observational studies in PCOS[11-14]. A

recent Mendelian randomization study showed that genetically predicted SHBG, total testosterone and free testosterone levels were associated with PCOS risk[5], which is not surprising given the adoption of hyperandrogenism as a diagnostic criterion of PCOS[15]. However, PCOS is a complex disorder, which comprises several metabolic and ovarian sub-phenotypes[2,25]. The patterns of associations seen in this study, support that different features of PCOS could have a unique aetiology with diverse, though potentially intertwining, pathophysiological pathways.

Experimental studies have shown that hepatic de novo lipogenesis, which is increased in obesity and insulin resistance[26], impairs SHBG synthesis in the liver[27]. We recently demonstrated that de novo lipogenesis, assessed by stable isotopes, is inversely associated with serum SHBG levels in women[28]. Hepatic de novo lipogenesis has also been associated with a disadvantageous lipid profile[29]. The reduction of the strength of the association between SHBG and serum lipids after adjustment for BMI in the current study is in line with these previous observations and suggests that BMI is an important driver of the metabolic features and low serum SHBG levels that characterize PCOS[30]. Indeed, a recent bidirectional Mendelian randomization study suggested that BMI is causal in the development of PCOS, but not vice versa[31]. Furthermore, other Mendelian randomization studies have demonstrated that free testosterone, SHBG and total testosterone do not appear to exert a causal effect on BMI, serum lipids and blood pressure[5,32,33]. In contrast, Mendelian randomization studies have suggested that SHBG is actively involved in the pathogenesis of type 2 diabetes, either directly or via free testosterone[5,34,35]. This is supported by experimental studies in humanized transgenic *SHBG* mice that were fed a high-fat diet and demonstrated an improved glucose homeostasis compared to wild-type mice[36].

The pathophysiology of the reproductive features of PCOS is not as well understood. Elevated AMH levels – a result of impaired follicle development and an increased number of antral follicles – has been linked to ovarian androgen hypersecretion by inhibition of the aromatase-induced conversion of androgens to oestrogens and stimulation of GnRH-dependent LH secretion[2,10,37,38]. Simultaneously, hyperandrogenism may increase AMH levels through its proposed disruptive effects on follicular development[2]. Insulin resistance has also been indicated as an important contributor to ovarian androgen secretion[39,40], yet we did not find a significant association between total testosterone or androstenedione with HOMA2-IR. The mechanism by which insulin resistance influences reproductive features of PCOS therefore deserves further investigation.

In the current study there was no association between any of the androgen markers and the Ferriman-Gallwey score. Although this may be the result of insufficient statistical power, it also corroborates a recent meta-regression analysis in 6593 women

with PCOS demonstrating that free testosterone levels were not associated with clinical hyperandrogenism[41]. Hirsutism is a phenotypic expression of several factors, including the androgen concentration, androgen receptor activity, and 5-α reductase activity at the pilosebaceous units[41]. Furthermore, the Ferriman-Gallwey score is a subjective measure with significant interobserver variability[42]. Both facets may contribute to the lack of an association between free testosterone levels and clinical hyperandrogenism.

The differential patterns of associations as observed in the present study indicates that PCOS constitutes a heterogeneous phenotype. The recent ESHRE guideline relies primarily on markers of free testosterone as a diagnostic tool for biochemical hyperandrogenism[17]. This allows the diagnosis of a broad range of PCOS phenotypes, as is also supported by the current findings. However, SHBG and total testosterone levels may guide clinicians in determining the primary contributing pathway (i.e. metabolic or reproductive) in individual patients. The extent to which either pathway is involved may vary greatly between individuals. Clinical follow-up and management of patients with PCOS might benefit from a more targeted approach, based on the primary underlying pathway, which warrants further study.

This study has several strengths and limitations. The PCOS cohort examined in this study has been systematically screened in a hospital setting, allowing us to gather information on a wide range of PCOS features. This setting, i.e. an outpatient clinic for menstrual cycle disorders, could have resulted in a selection of a particular subtype of PCOS. The interpretation of the results provides insight into the underlying pathways using observational data, but warrant further studies to unravel the exact role of free testosterone in the pathogenesis of the individual PCOS features. Finally, the relatively small sample size could have resulted in a lack of statistical power and, hence, type 2 errors. Indeed, several associations approached, but did not reach, statistical significance (Figure 3.1 and 3.2).

In conclusion, the current observational study shows differential associations of SHBG, androstenedione, total testosterone and free testosterone levels with metabolic and reproductive features of PCOS. These differential associations highlight the heterogeneous nature of PCOS, and suggest that the underlying pathways contributing to the features of PCOS are diverse. The combination of SHBG, total testosterone and androstenedione levels may provide information on the primary underlying pathophysiological pathway in women with PCOS.

References

1. Wolf WM, Wattick RA, Kinkade ON, Olfert MD. Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity. *International Journal of Environmental Research and Public Health.* 2018;15(11):2589.
2. Azziz R, Carmina E, Chen Z, et al. Polycystic ovary syndrome. *Nature Reviews Disease Primers.* 2016;2(1):16057.
3. Azziz R, Sanchez LA, Knochenhauer ES, et al. Androgen excess in women: Experience with over 1000 consecutive patients. *Journal of Clinical Endocrinology and Metabolism.* 2004;89(2):453-462.
4. Livadas S, Pappas C, Karachalios A, et al. Prevalence and impact of hyperandrogenemia in 1,218 women with polycystic ovary syndrome. *Endocrine.* 2014;47(2):631-638.
5. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nature Medicine.* 2020;26(2):252-258.
6. Hall JE, Taylor AE, Hayes FJ, Crowley WF, Jr. Insights into hypothalamic-pituitary dysfunction in polycystic ovary syndrome. *Journal of Clinical Investigation.* 1998;21(9):602-611.
7. Norman RJ, Davies MJ, Lord J, Moran LJ. The role of lifestyle modification in polycystic ovary syndrome. *Trends in Endocrinology and Metabolism.* 2002;13(6):251-257.
8. Sanchez-Garrido MA, Tena-Sempere M. Metabolic dysfunction in polycystic ovary syndrome: Pathogenic role of androgen excess and potential therapeutic strategies. *Molecular Metabolism.* 2020;35:100937.
9. Lebbe M, Woodruff TK. Involvement of androgens in ovarian health and disease. *Molecular Human Reproduction.* 2013;19(12):828-837.
10. Walters KA, Gilchrist RB, Ledger WL, et al. New Perspectives on the Pathogenesis of PCOS: Neuroendocrine Origins. *Trends in Endocrinology and Metabolism.* 2018;29(12):841-852.
11. Deswal R, Yadav A, Dang AS. Sex hormone binding globulin - an important biomarker for predicting PCOS risk: A systematic review and meta-analysis. *Systems Biology in Reproductive Medicine.* 2018;64(1):12-24.
12. Moran LJ, Teede HJ, Noakes M, et al. Sex hormone binding globulin, but not testosterone, is associated with the metabolic syndrome in overweight and obese women with polycystic ovary syndrome. *Journal of Clinical Investigation.* 2013;36(11):1004-1010.
13. Kajaia N, Binder H, Dittrich R, et al. Low sex hormone-binding globulin as a predictive marker for insulin resistance in women with hyperandrogenic syndrome. *European Journal of Endocrinology.* 2007;157(4):499-507.
14. Antonio L, Pauwels S, Laurent MR, et al. Free testosterone reflects metabolic as well as ovarian disturbances in subfertile oligomenorrheic women. *International Journal of Endocrinology.* 2018;2018:7956951-7956951.
15. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertility and Sterility.* 2004;81(1):19-25.
16. Laven JS, Imani B, Eijkemans MJ, Fauser BC. New approach to polycystic ovary syndrome and other forms of anovulatory infertility. *Obstetrical and Gynecological Survey.* 2002;57(11):755-767.
17. Teede HJ, Misso ML, Costello MF, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Human Reproduction.* 2018;33(9):1602-1618.
18. DeUgarte CM, Woods KS, Bartolucci AA, Azziz R. Degree of facial and body terminal hair growth in unselected black and white women: toward a populational definition of hirsutism. *Journal of Clinical Endocrinology and Metabolism.* 2006;91(4):1345-1350.
19. Zhao X, Ni R, Li L, et al. Defining hirsutism in Chinese women: a cross-sectional study. *Fertility and Sterility.* 2011;96(3):792-796.
20. Jonard S, Robert Y, Dewailly D. Revisiting the ovarian volume as a diagnostic criterion for polycystic ovaries. *Human Reproduction.* 2005;20(10):2893-2898.
21. Li HW, Wong BP, Ip WK, et al. Comparative evaluation of three new commercial immunoassays for anti-müllerian hormone measurement. *Human Reproduction.* 2016;31(12):2796-2802.

22. Ross HA, Meuleman EJ, Sweep FCGJ. A simple method for estimating equilibrium constants for serum testosterone binding resulting in an optimal free testosterone index for use in elderly men. *Clinical Chemistry and Laboratory Medicine.* 2005;43(6):613-616.
23. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clinical Chemistry.* 1972;18(6):499-502.
24. Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: An American Heart Association/National Heart, Lung, and Blood Institute scientific statement. *Circulation.* 2005;112(17):2735-2752.
25. Dapas M, Lin FTJ, Nadkarni GN, et al. Distinct subtypes of polycystic ovary syndrome with novel genetic associations: An unsupervised, phenotypic clustering analysis. *PLoS Medicine.* 2020;17(6):e1003132.
26. Smith GI, Shankaran M, Yoshino M, et al. Insulin resistance drives hepatic de novo lipogenesis in nonalcoholic fatty liver disease. *Journal of Clinical Investigation.* 2020;130(3):1453-1460.
27. Selva DM, Hogeveen KN, Innis SM, Hammond GL. Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. *Journal of Clinical Investigation.* 2007;117(12):3979-3987.
28. Simons PIHG, Valkenburg O, Telgenkamp I, et al. Relationship between de novo lipogenesis and serum sex hormone-binding globulin in humans. *Clinical Endocrinology.* 2021.
29. Schwarz J-M, Linfoot P, Dare D, Aghajanian K. Hepatic de novo lipogenesis in normoinsulinemic and hyperinsulinemic subjects consuming high-fat, low-carbohydrate and low-fat, high-carbohydrate isoenergetic diets. *American Journal of Clinical Nutrition.* 2003;77(1):43-50.
30. Zhu JL, Chen Z, Feng WJ, Long SL, Mo ZC. Sex hormone-binding globulin and polycystic ovary syndrome. *Clinica Chimica Acta.* 2019;499:142-148.
31. Zhao Y, Xu Y, Wang X, et al. Body mass index and polycystic ovary syndrome: A 2-sample bidirectional Mendelian randomization study. *Journal of Clinical Endocrinology and Metabolism.* 2020;105(6).
32. Wang Q, Kangas AJ, Soininen P, et al. Sex hormone-binding globulin associations with circulating lipids and metabolites and the risk for type 2 diabetes: Observational and causal effect estimates. *International Journal of Epidemiology.* 2015;44(2):623-637.
33. Haring R, Teumer A, Völker U, et al. Mendelian randomization suggests non-causal associations of testosterone with cardiometabolic risk factors and mortality. *Andrology.* 2013;1(1):17-23.
34. Ding EL, Song Y, Manson JE, et al. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *New England Journal of Medicine.* 2009;361(12):1152-1163.
35. Perry JRB, Weedon MN, Langenberg C, et al. Genetic evidence that raised sex hormone binding globulin (SHBG) levels reduce the risk of type 2 diabetes. *Human Molecular Genetics.* 2010;19(3): 535-544.
36. Saez-Lopez C, Villena JA, Simó R, Selva DM. Sex hormone-binding globulin overexpression protects against high-fat diet-induced obesity in transgenic male mice. *The Journal of Nutritional Biochemistry.* 2020;85:108480.
37. Chang HM, Klausen C, Leung PC. Antimüllerian hormone inhibits follicle-stimulating hormone-induced adenylyl cyclase activation, aromatase expression, and estradiol production in human granulosa-lutein cells. *Fertility and Sterility.* 2013;100(2):585-592.e581.
38. Broekmans FJ, Visser JA, Laven JS, et al. Anti-Müllerian hormone and ovarian dysfunction. *Trends in Endocrinology and Metabolism.* 2008;19(9):340-347.
39. Tosi F, Negri C, Brun E, et al. Insulin enhances ACTH-stimulated androgen and glucocorticoid metabolism in hyperandrogenic women. *Eur J Endocrinol.* 2011;164(2):197-203.
40. Nestler JE, Jakubowicz DJ, de Vargas AF, et al. Insulin stimulates testosterone biosynthesis by human thecal cells from women with polycystic ovary syndrome by activating its own receptor and using inositolglycan mediators as the signal transduction system. *Journal of Clinical Endocrinology and Metabolism.* 1998;83(6):2001-2005.
41. Amiri M, Ramezani Tehrani F, Nahidi F, et al. Association between biochemical hyperandrogenism parameters and Ferriman-Gallwey score in patients with polycystic ovary syndrome: A systematic review and meta-regression analysis. *Clinical Endocrinology.* 2017;87(3):217-230.

42. Wild RA, Vesely S, Beebe L, Whitsett T, Owen W. Ferriman Gallwey Self-Scoring I: Performance Assessment in Women with Polycystic Ovary Syndrome. *Journal of Clinical Endocrinology and Metabolism.* 2005;90(7):4112-4114.

Supplementary materials

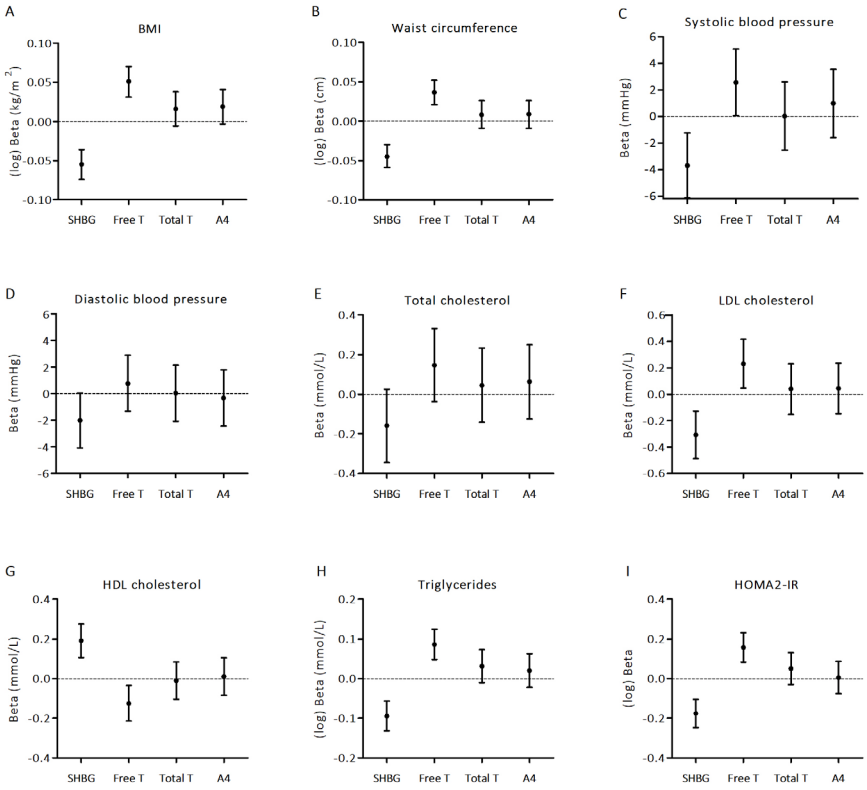

Figure S3.1 Associations of serum sex hormone-binding globulin (SHBG), free testosterone (Free T), total testosterone (Total T), and androstenedione (A4) with metabolic features of polycystic ovary syndrome in fasted individuals only: BMI (n=89) (A) waist circumference (n=88) (B) systolic blood pressure (n=89) (C) diastolic blood pressure (n=89) (D) total cholesterol (n=88) (E) LDL cholesterol (n=88) (F) HDL cholesterol (n=88) (G) triglycerides (n=88) (H) and homeostatic model assessment of insulin resistance (HOMA2-IR) (n=87) (I). Analyses were conducted with Z-scores to allow comparison. Regression coefficients should therefore be interpreted as the increase in the dependent variable per standard deviation increase in serum SHBG, free testosterone, total testosterone or androstenedione. See methods section.

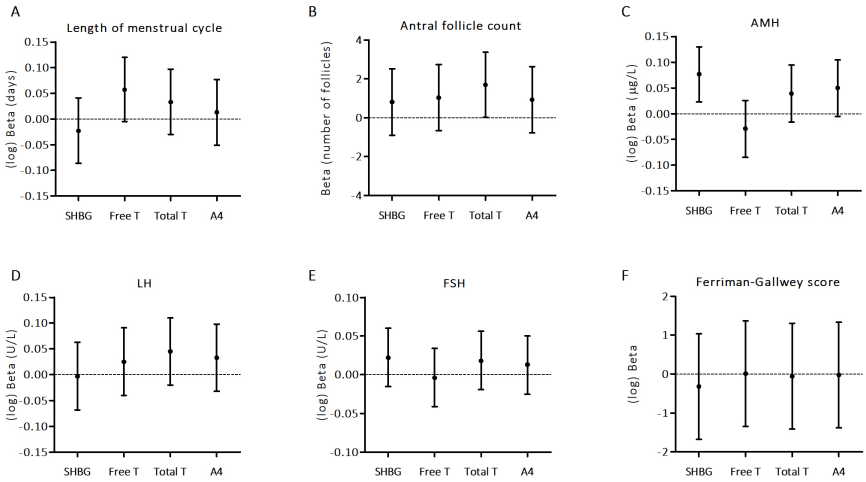

Figure S3.2 Associations of serum sex hormone-binding globulin (SHBG), free testosterone (Free T), total testosterone (Total T), and androstenedione (A4) with reproductive features of polycystic ovary syndrome in fasted individuals only: length of menstrual cycle (n=83) (A) antral follicle count (n=85) (B) anti-Müllerian hormone (AMH) (n=86) (C) luteinizing hormone (LH) (n=89) (D) follicle stimulating hormone (FSH) (n=89) (E) and Ferriman-Gallwey score (n=89) (F). Analyses were conducted with Z-scores to allow comparison. Regression coefficients should, therefore, be interpreted as the increase in the dependent variable per standard deviation increase in serum SHBG, free testosterone, total testosterone or androstenedione. See methods section.

Table S3.1 Associations of serum SHBG, free testosterone, total testosterone and androstenedione with metabolic and reproductive features of PCOS in fasted individuals only.

Independent variables	SHBG B (95% CI)	Free testosterone B (95% CI)	Total testosterone β (95% CI)	Androstenedione B (95% CI)
Metabolic features				
(log) BMI				
Crude	**-0.06 (-0.07;-0.04)**	**0.05 (0.03;0.07)**	0.02 (-0.01;0.04)	0.02 (0.00;0.04)
Age	**-0.05 (-0.07;-0.04)**	**0.05 (0.03;0.07)**	0.01 (-0.01;0.04)	0.02 (-0.01;0.04)
(log) Waist circumference				
Crude	**-0.05 (-0.06;-0.03)**	**0.04 (0.02;0.05)**	0.01 (-0.01;0.03)	0.01 (-0.01;0.03)
Age	**-0.05 (-0.06;-0.03)**	**0.04 (0.02;0.05)**	0.01 (-0.01;0.03)	0.01 (-0.01;0.03)
Systolic blood pressure				
Crude	**-3.69 (-6.45;-1.24)**	**2.56 (0.04;5.08)**	0.03 (-2.55;2.60)	0.98 (-1.59;3.55)
Age	**-4.00 (-6.49;-1.50)**	**2.91 (0.32;5.50)**	0.12 (-2.49;2.72)	1.33 (-1.37;4.02)
Age, BMI	-1.67 (-3.76;1.42)	-0.28 (-2.90;2.35)	-1.02 (-3.28;1.25)	0.13 (-2.24;2.50)
Age, BMI, HOMA2-IR	-1.12 (-3.72;1.48)	-0.63 (-3.33;2.07)	-1.43 (-3.77;0.91)	0.35 (-2.05;2.76)
Diastolic blood pressure				
Crude	-2.02 (-4.09;0.05)	0.76 (-1.35;2.87)	0.03 (-2.08;2.15)	-0.33 (-2.45;1.78)
Age	**-2.52 (-4.57;-0.47)**	1.33 (-0.79;3.46)	0.27 (-1.82;2.37)	0.30 (-1.88;2.48)
Age, BMI	-1.21 (-3.50;1.09)	-0.36 (-2.69;1.97)	-0.31 (-2.32;1.71)	-0.33 (-2.43;1.77)
Age, BMI, HOMA2-IR	-1.11 (-3.36;1.14)	-1.01 (-3.35;1.33)	-0.95 (-3.00;1.09)	0.07 (-2.02;2.16)
LDL cholesterol				
Crude	**-0.31 (-0.49;-0.13)**	**0.23 (0.05;0.42)**	0.04 (-0.15;0.23)	0.05 (-0.15;0.24)
Age	**-0.34 (-0.52;-0.16)**	**0.27 (0.08;0.46)**	0.05 (-0.14;0.28)	0.08 (-0.12;0.28)
Age, BMI	**-0.23 (-0.43;-0.03)**	0.14 (-0.08;0.35)	-0.01 (-0.19;0.18)	0.01 (-0.18;0.21)
Age, BMI, HOMA2-IR	**-0.23 (-0.43;-0.02)**	0.13 (-0.09;0.35)	-0.02 (-0.22;0.17)	0.02 (-0.17;0.22)
HDL cholesterol				
Crude	**0.19 (0.11;0.28)**	**-0.13 (-0.22;-0.04)**	-0.01 (-0.10;0.08)	0.01 (-0.08;0.11)
Age	**0.18 (0.09;0.26)**	**-0.11 (-0.20;-0.01)**	0.00 (-0.09;0.10)	0.05 (-0.05;0.14)
Age, BMI	0.08 (-0.01;0.17)	0.02 (-0.08;0.11)	0.05 (-0.03;0.12)	**0.09 (0.01;0.17)**
Age, BMI, HOMA2-IR	0.08 (-0.01;0.17)	0.03 (-0.06;0.13)	0.06 (-0.02;0.14)	**0.09 (0.06;0.17)**
(log) Triglycerides				
Crude	**-0.10 (-0.13;-0.06)**	**0.09 (0.05;0.12)**	0.03 (-0.01;0.07)	0.02 (-0.02;0.06)
Age	**-0.09 (-0.13;-0.06)**	**0.09 (0.05;0.13)**	0.03 (-0.01;0.07)	0.02 (-0.03;0.06)
Age, BMI	**-0.07 (-0.11;-0.02)**	**0.05 (0.01;0.10)**	0.01 (-0.02;0.05)	0.00 (-0.04;0.04)
Age, BMI, HOMA2-IR	**-0.06 (-0.11;-0.02)**	**0.05 (0.01;0.10)**	0.01 (-0.03;0.05)	0.01 (-0.04;0.04)
(log) HOMA2-IR				
Crude	**-0.18 (-0.25;-0.11)**	**0.16 (0.08;0.23)**	0.05 (-0.03;0.13)	0.01 (-0.08;0.09)
Age	**-0.17 (-0.25;-0.10)**	**0.16 (0.08;0.23)**	0.05 (-0.04;0.13)	-0.01 (-0.10;0.08)
Age, BMI	-0.07 (-0.14;0.00)	0.04 (-0.03;0.12)	0.00 (-0.06;0.07)	-0.06 (-0.12;0.01)
Reproductive features				
Antral follicle count				
Crude	0.81 (-0.89;2.52)	1.04 (-0.66;2.74)	**1.70 (0.22;3.37)**	0.93 (-0.78;2.63)
Age	1.07 (-0.65;2.80)	0.81 (-0.94;2.56)	1.59 (-0.10;3.27)	0.65 (-1.13;2.43)
Age, BMI	1.11 (-0.88;3.10)	1.37 (-0.63;3.37)	**1.72 (0.01;3.43)**	0.76 (-1.06;2.58)
Age, BMI, HOMA2-IR	1.19 (-0.77;3.14)	0.95 (-1.09;3.00)	1.35 (-0.41;3.11)	1.12 (-0.69;2.92)
(log) AMH				
Crude	**0.08 (0.02;0.13)**	-0.03 (-0.09;0.03)	0.04 (-0.02;0.10)	0.05 (-0.01;0.11)
Age	**0.08 (0.03;0.14)**	-0.04 (-0.09;0.02)	0.04 (-0.02;0.09)	0.05 (-0.01;0.11)
Age, BMI	0.05 (-0.01;0.11)	0.01 (-0.05;0.07)	**0.06 (0.00;0.11)**	**0.07 (0.01;0.12)**
Age, BMI, HOMA2-IR	0.05 (-0.01;0.11)	0.02 (-0.05;0.08)	**0.06 (0.01;0.13)**	**0.07 (0.01;0.12)**

Analyses were conducted with Z-scores. Beta coefficients should therefore be interpreted as per standard deviation increase in serum SHBG, free testosterone, total testosterone or androstenedione. See methods section. Bold values indicate statistical significance (p<0.05).

Abbreviations: AMH anti-Mullerian hormone; BMI body mass index; HOMA2-IR homeostatic model assessment for insulin resistance; HDL high-density lipoprotein; LDL low-density lipoprotein; AMH anti-Mullerian hormone; SHBG sex hormone-binding globulin.

PART II

THE ASSOCIATION BETWEEN DE NOVO LIPOGENESIS AND SEX HORMONE-BINDING GLOBULIN

4

CHAPTER FOUR

Relationship between de novo lipogenesis and serum sex hormone-binding globulin in humans

Pomme I.H.G. Simons, Olivier Valkenburg, Ine Telgenkamp, Koen M. van der Waaij, David M. de Groot, Pandichelvam Veeraiah, Judith A.P. Bons, Marja-Riitta Taskinen, Jan Borén, Patrick Schrauwen, Joost H.W. Rutten, David Cassiman, Casper G. Schalkwijk, Coen D.A. Stehouwer, Vera B. Schrauwen-Hinderling, Leanne Hodson, Martijn C.G.J. Brouwers

Clinical Endocrinology (2021); 95 (1): 101-106

Summary

Objective: Obesity and liver fat are associated with decreased levels of serum sex hormone-binding globulin (SHBG). Laboratory studies suggest that hepatic de novo lipogenesis is involved in the downregulation of SHBG synthesis. The aim of the present study was to address the role of de novo lipogenesis on serum SHBG in humans.

Design: A cross-sectional study examining the association between de novo lipogenesis, measured by stable isotopes, and serum SHBG, stratified by sex.

Participants: Healthy men (n=34) and women (n=21) were combined from two cross-sectional studies. Forty-two percent of participants had hepatic steatosis, and the majority were overweight (62%) or obese (27%).

Results: De novo lipogenesis was inversely associated with SHBG in women (β: -0.015, 95% CI: -0.030;0.000), but not in men (β: 0.007, 95% CI: -0.005;0.019) (p for interaction=0.068). Adjustment for study population, age and BMI did not materially change these results, although statistical significance was lost after adjustment for serum insulin.

Conclusions: An inverse association between de novo lipogenesis and SHBG may explain the decreased SHBG levels that are observed in obesity, at least in women.

Introduction

Obesity has become a worldwide health burden that is associated with many health concerns including hypertension, dyslipidaemia, type 2 diabetes, non-alcoholic fatty liver disease, cardiovascular disease, gout, osteoarthritis, fractures and gall bladder disease[1-3]. Individuals with obesity are commonly characterized by low serum sex hormone-binding globulin (SHBG) levels[4]. SHBG is a liver-specific glycoprotein that binds sex hormones in blood and thereby regulates their bioavailability[5]. The exact mechanism by which obesity leads to a decrease in serum SHBG levels in humans is not fully understood, although it is likely to be multifactorial. We and others recently showed that a weight reduction program was associated with an increase in serum SHBG levels[4,6]. Furthermore, the change in intrahepatic lipid (IHL) content was inversely associated with serum SHBG levels[4].

Excessive accumulation of IHL in obesity can be explained by an increased conversion of glucose to fat (i.e. de novo lipogenesis) and an increased flux of free fatty acids from insulin-resistant adipose tissue to the liver[7]. Of interest, previous in vitro studies and mice studies have demonstrated that monosaccharide-induced de novo lipogenesis reduced serum SHBG levels[8]. Furthermore, palmitate – a saturated fatty acid that is the principal end product of de novo lipogenesis – directly reduced SHBG expression in HepG2 cells[8].

The aim of the present study was to extrapolate these experimental data to the human situation. For this we determined the relationship between de novo lipogenesis, assessed by stable isotopes (the gold standard), and serum SHBG, corrected for potential confounding factors.

Methods

Study design

In this cross-sectional study, data from two previously published cohorts in Oxford (UK) and Maastricht (the Netherlands) were combined[9,10]. Both studies were performed according to the Declaration of Helsinki[11] and approved by the Medical Ethical Committee of Maastricht University Medical Centre or the Portsmouth Clinical Research Ethics Committee. All participants gave written informed consent prior to participation.

Oxford study population

This study originally aimed to assess the effect of insulin resistance on the synthesis and partitioning of intrahepatic fatty acids. For this purpose, healthy individuals were included when they had a body mass index (BMI) <30 kg/m^2, did not use medication affecting lipid or glucose metabolism and did not excessively smoke or consume alcohol[9]. Individuals with high serum triglyceride levels (>4 mmol/l) were excluded from the present study, as high triglycerides affect the reliability of de novo lipogenesis assessment[10,12]. All measurements were performed after an overnight fast, and individuals were asked not to consume foods rich in ^{13}C or alcohol and to avoid strenuous exercise[9].

Anthropometrics, measurements of serum lipids, insulin and glucose and quantification of IHL content by proton magnetic resonance spectroscopy (^{1}H-MRS) were done as described previously[9]. IHL content was expressed as the ratio CH_2/H_20. SHBG in the Oxford study population and the Maastricht study population was measured with an automated chemiluminescent immunometric assay on the Immulite XPi instrument (Siemens Healthcare Diagnostics, New Orleans, LA) in heparinized plasma and serum, respectively (for practical reasons SHBG is further referred to as 'serum SHBG').

De novo lipogenesis was quantified by oral ingestion of deuterated water (^2H_2O) (3 g/kg body water) the evening prior to the measurements and throughout the measurement day[9]. The incorporation of deuterium from $^2H_20$ in plasma water into very-low-density lipoprotein (VLDL)-triglyceride palmitate is representative of newly synthesized fatty acids from a non-lipid precursor, and, hence, a marker of de novo lipogenesis. This was measured with gas chromatography-mass spectrometry (Finnigan GasBench II Thermo Fisher Scientific, Paisley, UK)[9].

Maastricht study population

This study was primarily conducted to establish a ^{1}H-MRS methodology to distinguish intrahepatic saturated, mono- and polyunsaturated fatty acids in vivo and to assess the relationship between hepatic lipid composition and de novo lipogenesis[10]. Participants were excluded if they had an active illness, participated in an exercise program for more than two hours per week, had significant weight change prior to enrolment, consumed more than two units of alcohol per day or smoked more than five cigarettes per day, used anti-coagulants or other medication that interferes with hepatic lipid composition, or had high serum triglyceride levels (>4 mmol/l). Participants were instructed to refrain from alcohol consumption or physical exercise for two days prior to the measurements and to consume a standardized high carbohydrate dinner the

evening prior to the measurements. They visited the metabolic research ward after an overnight fast[10].

Anthropometrics, measurements of serum lipids and glucose and quantification of IHL content (expressed as the ratio CH_2/H_20) by ^{1}H-MRS were performed as previously described[10]. Serum insulin was measured with an automated chemiluminescent immunometric assay on the Immulite XPi instrument (Siemens Healthcare Diagnostics, New Orleans, LA).

DNL was quantified by oral ingestion of deuterated water (2.86 g/kg body weight; 70% $^2H_20$, Cambridge Isotope laboratories) the evening prior to the measurement. De novo lipogenesis was quantified by the isotopic enrichment ratio of VLDL-triglyceride palmitate, measured with gas chromatography-mass spectrometry (Agilent, Santa Rosa, CA; Model 6890N/5975B)[13].

Statistical analyses

Continuous data are presented as mean ± standard deviation (SD) or as median (interquartile range) in case of non-normal distribution. Categorical data are presented as frequencies. Non-normally distributed variables were log-transformed before further analyses. Multivariable regression analyses were performed to study the association between de novo lipogenesis and serum SHBG, adjusted for study population (Oxford or Maastricht), age, BMI and fasting insulin levels. Given the well-known sex differences in SHBG levels, all primary analyses were stratified by sex. A potential interaction between sex and de novo lipogenesis on serum SHBG was formally tested by adding an interaction term (sex*de novo lipogenesis) to the regression model in the overall population, i.e. men and women combined.

All results were considered statistically significant at $p<0.05$, except for interaction terms ($p<0.10$). All statistical analyses were performed using IBM Statistical Package of Social Science (SPSS) version 25.0 for Windows (IBM Corp., Armonk, N.Y., USA).

Results

Study population characteristics

Three participants of the original Oxford (n=41) and Maastricht (n=17) cohorts were excluded from further analyses due to insufficient serum to determine SHBG (n=1, Oxford) or serum triglycerides >4 mmol/l (n=1, Maastricht; n=1, Oxford). The general characteristics of the Oxford and Maastricht study populations are presented in Table 4.1. Maastricht participants were older and more overweight compared to the Oxford

cohort. BMI distribution in the combined cohort ranged from lean (BMI <25 kg/m^2: 6/55 [11%]), overweight (BMI ≥25 and <30 kg/m^2: 34/55 [62%]) to obese (BMI ≥30 kg/m^2: 15/55 [27%]). Twenty-two out of 55 (40%) individuals had an IHL content above the cut-off value for hepatic steatosis (i.e. >5.56% IHL content)[7]. None of the female participants used oral contraceptives. Serum insulin levels were substantially higher in the Oxford cohort, which is most likely explained by a difference in assay. There was no statistically significant association between de novo lipogenesis and IHL content (β: 0.008, 95% CI: -0.016;0.032; adjusted for study population; Figure 4.1). Stratification by sex showed similar results in men (β: 0.000, 95% CI: -0.029;0.029) and women (β: 0.014, 95% CI: -0.030;0.059).

Table 4.1 Characteristics of Oxford and Maastricht study populations.

	Oxford (n=39)	Maastricht (n=16)
Male/Female,n	28/11	6/10
Age, years	44.2 ± 6.4	59.3 ± 7.0
BMI, kg/m^2	27.9 ± 2.9	29.6 ± 2.2
Glucose, mmol/l	5.3 ± 0.5	5.5 ± 0.6
Insulin, pmol/l	83.9 (60.3-99.2)	54.8 (32.3-85.0)
Total cholesterol, mmol/l	5.4 ± 0.9	5.6 ± 1.1
LDL-cholesterol, mmol/l	3.9 ± 0.8	4.0 ± 1.1
HDL-cholesterol, mmol/l	1.1 ± 0.3	1.3 ± 0.4
Triglycerides, mmol/l	1.8 (1.1-2.2)	1.6 (1.3-2.5)
Intrahepatic lipids, %	4.1 (1.7-6.8)	4.1 (1.3-12.2)
De novo lipogenesis,, %	7.6 (5.3-11.6)	10.3 (6.9-12.2)
SHBG, nmol/l	28.1 (22.8-36.1)	38.2 (31.6-57.3)
Use of oral contraceptives, n	0	0

Data are expressed as mean ± SD or as median (interquartile range).
Abbreviations: BMI body mass index; SHBG sex hormone-binding globulin.

Relationship between de novo lipogenesis and serum SHBG levels

In the combined cohort, serum SHBG levels were not significantly different between men with and without obesity, whereas SHBG levels were lower in women who were obese (β: -0.083, 95% CI: -0.212;0.046, Figure 4.2, panel A; and β: -0.183, 95% CI: -0.361;-0.005, Figure 4.2, panel B, respectively; adjusted for study population).

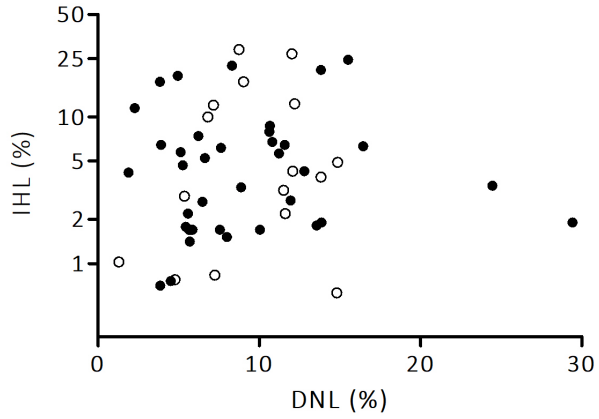

Figure 4.1 Association between de novo lipogenesis (DNL) and intrahepatic lipid (IHL) content stratified by study population, i.e. Oxford (closed circles) and Maastricht (open circles).

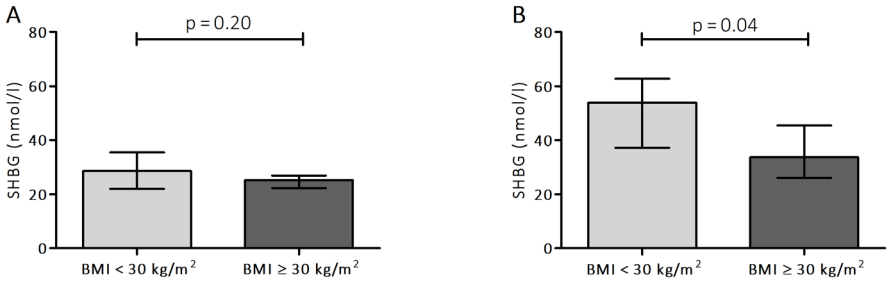

Figure 4.2 Sex hormone-binding globulin (SHBG) levels in obese (BMI ≥30 kg/m^2) and non-obese (BMI <30 kg/m^2) men (A) and women (B). Data are expressed as median with interquartile range. Differences between groups were analysed with linear regression analyses, adjusted for study population.

In men, there was no statistically significant association between de novo lipogenesis and serum SHBG levels (β: 0.007, 95% CI: -0.005;0.019; Table 4.2, Figure 4.3, panel A). Similar trends were observed when the study populations, i.e. Oxford and Maastricht were analysed separately (β: 0.003, 95% CI: -0.009;0.015 and β: 0.038; 95% CI: -0.009;0.085, respectively; Figure 4.3, panel A). Further adjustment for study population, age, BMI and serum insulin did not alter the results (β: 0.002, 95% CI: -0.009;0.014; Table 4.2). In women, there was a statistically significant, inverse association between de novo lipogenesis and serum SHBG (β: -0.015, 95% CI: -0.030;0.000; Table 4.2, Figure 4.3, panel B). The strength of association was statistically significantly different from men (p for interaction =0.068). Similar trends were observed when the study populations were analysed separately (β: -0.019, 95%

CI: -0.043;0.004 and β: -0.006, 95% CI: -0.003;0.019 for Oxford and Maastricht study population, respectively; Figure 4.3, panel B). The strength of the association did not materially change after further adjustment for study population, age and BMI, although statistical significance was lost after further adjustment for serum insulin (β: -0.013, 95% CI: -0.028;0.003; Table 4.2). Of note, serum insulin was not an independent determinant of serum SHBG in this fully adjusted model (p=0.219).

Table 4.2 Association of de novo lipogenesis with (log) sex hormone-binding globulin in men and women.

	Men (n=34)		Women (n=21)	
Model, independent variables	**β**	**95% CI**	**β**	**95% CI**
Crude	0.007	-0.005;0.019	-0.015	-0.030;0.000
Model 1: study population (Oxford/Maastricht)	0.005	-0.006;0.016	-0.015	-0.031;0.000
Model 2: model 1 + age	0.001	-0.011;0.012	-0.015	-0.031;0.001
Model 3: model 2 + BMI	0.002	-0.010;0.014	-0.018	-0.031;-0.006
Model 4: model 3 + serum insulin	0.002	-0.009;0.014	-0.013	-0.028;0.003

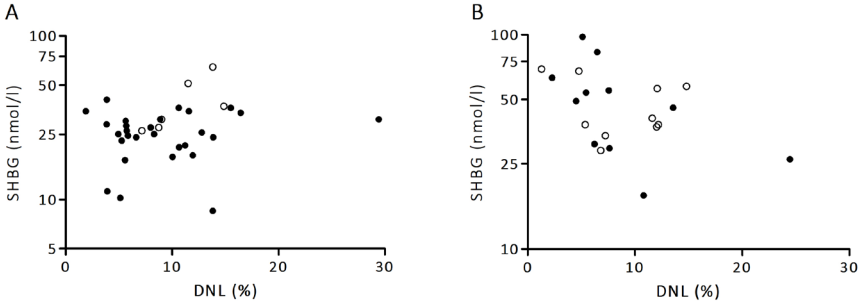

Figure 4.3 Associations between de novo lipogenesis (DNL) and sex hormone-binding globulin (SHBG) in men (A) and women (B). Data are stratified by study cohort, i.e. Oxford cohort (closed circles) and Maastricht cohort (open circles).

Discussion

The aim of this study was to examine the relationship between de novo lipogenesis and serum SHBG in humans. We found an inverse association between de novo lipogenesis, measured with stable isotopes, and serum SHBG in women but not in men.

The current findings support and extend previous in vitro and animal studies, showing that de novo lipogenesis is involved in SHBG regulation. In vitro studies have demonstrated that monosaccharide-induced de novo lipogenesis in HepG2 cells resulted in downregulation of hepatocyte nuclear factor-4alpha (HNF-4α) and, consequently, reduced expression of SHBG[8]. Similar results were obtained after

incubation with palmitate, a saturated fatty acid that is the end-product of de novo lipogenesis[8]. In the present study, we found that incorporation of deuterium into VLDL-triglyceride palmitate, a measure of de novo lipogenesis, was inversely associated with serum SHBG levels in women.

We observed a statistically significant interaction between sex and de novo lipogenesis on serum SHBG levels. The inverse relationship between de novo lipogenesis and serum SHBG was observed in women, but not in men. Strikingly, the inverse association between obesity and SHBG was also more pronounced in women. These sex differences may be accidental and, hence, deserve further replication. Alternatively, the difference between men and women may be the result of biological differences in transcriptional regulation of SHBG between men and women[14]. It has been suggested that HNF-4α, the oestrogen receptor alpha, and PPARG compete for binding to the SHBG promotor, with the former two stimulating and the latter inhibiting *SHBG* gene expression[15]. The net effect of this competition on serum SHBG levels is difficult to predict and deserves further investigation.

Nevertheless, we postulate that the relationship between de novo lipogenesis and serum SHBG in women is of particular interest as it may provide a mechanistic link between obesity, more specifically hepatic fat accumulation and polycystic ovary syndrome (PCOS). Previous observational studies have shown that patients with PCOS have a high IHL content[16]. A recent Mendelian randomization studies has inferred a causal relationship between low serum SHBG levels and PCOS risk[17]. Of note, it is likely that factors other than de novo lipogenesis, such as tumor necrosis factor α and interleukin 1β, also contribute to the decreased serum SHBG levels in obesity and related disorders[18,19].

In this study, insulin did not appear to be a major contributor of serum SHBG levels. To date, a large body of literature has reported an inverse association between serum insulin and serum SHBG levels in humans[20]. It is, however, virtually impossible to distinguish a potential direct effect of insulin on SHBG expression from confounding in an observational study design, particularly because insulin also affects de novo lipogenesis[21,22]. Although statistical significance was lost when insulin was added to the model as a potential confounder of the relationship between de novo lipogenesis and serum SHBG in women, the effect size for that relationship was hardly affected (the beta coefficient decreased from -0.018 to -0.013), which indicates that insulin is not a major contributor. A lack of statistical power, as a result of adjustment for multiple variables, is more likely. Indeed, serum insulin was not an independent determinant of serum SHBG in this cohort.

In the present study we did not observe an association between de novo lipogenesis and IHL content. IHL content is the net result of the influx of lipids – via de novo lipogenesis and free fatty acids from adipose tissue – and the efflux of lipids – via beta-oxidation and VLDL secretion[7]. Each pathway is regulated by many genetic, environmental and hormonal factors[7]. The original Oxford study showed that, as a result of differential partitioning of fatty acids in the liver, higher rates of de novo lipogenesis are not necessarily reflected by an increased IHL content[9]. The authors speculated that this may be the result of preferential channelling of de novo synthesized fatty acids towards VLDL secretion rather than hepatic storage[9].

This study has several strengths and limitations. First, by combining data from two study populations, i.e. Oxford and Maastricht, we were able to create a relatively large cohort to study the sex-specific relationship between de novo lipogenesis, assessed with stable isotopes, and serum SHBG. Although differences between the cohorts may exist, regression analyses were adjusted for study population, which did not affect the strength of the association. In addition, stratified analyses in the Oxford and Maastricht cohort yielded similar trends. Second, although none of the included women used oral contraceptives, which are known to significantly affect SHBG levels[23], we did not have information on postmenopausal status or phase of menstrual cycle. Previous studies have shown that menopausal status does not seem to have an independent effect on de novo lipogenesis or SHBG levels[24-26]. Furthermore other studies have shown that de novo lipogenesis varies significantly throughout the menstrual cycle, while levels of SHBG remain constant[27,28]. Despite these scattering effects, a significant, inverse association was observed between de novo lipogenesis and SHBG in women.

In conclusion, in the present study we corroborate and extrapolate findings from previous in vitro and animal studies by showing that de novo lipogenesis is inversely associated with serum SHBG in women.

References

1. Guh DP, Zhang W, Bansback N, et al. The incidence of co-morbidities related to obesity and overweight: A systematic review and meta-analysis. *BMC Public Health.* 2009;9:88.
2. Williams EP, Mesidor M, Winters K, Dubbert PM, Wyatt SB. Overweight and Obesity: Prevalence, Consequences, and Causes of a Growing Public Health Problem. *Curr Obes Rep.* 2015;4(3):363-370.
3. Shanbhogue VV, Mitchell DM, Rosen CJ, Bouxsein ML. Type 2 diabetes and the skeleton: new insights into sweet bones. *Lancet Diabetes Endocrinol.* 2016;4(2):159-173.
4. Telgenkamp I, Kusters YHAM, Schalkwijk CG, et al. Contribution of Liver Fat to Weight Loss–Induced Changes in Serum Hepatokines: A Randomized Controlled Trial. *The Journal of Clinical Endocrinology & Metabolism.* 2019;104(7):2719-2727.
5. Hammond GL, Wu TS, Simard M. Evolving utility of sex hormone-binding globulin measurements in clinical medicine. *Curr Opin Endocrinol Diabetes Obes.* 2012;19(3):183-189.
6. Niskanen L, Laaksonen DE, Punnonen K, et al. Changes in sex hormone-binding globulin and testosterone during weight loss and weight maintenance in abdominally obese men with the metabolic syndrome. *Diabetes Obes Metab.* 2004;6(3):208-215.
7. Ipsen DH, Lykkesfeldt J, Tveden-Nyborg P. Molecular mechanisms of hepatic lipid accumulation in non-alcoholic fatty liver disease. *Cellular and Molecular Life Sciences.* 2018;75(18):3313-3327.
8. Selva DM, Hogeveen KN, Innis SM, Hammond GL. Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. *J Clin Invest.* 2007;117(12):3979-3987.
9. Pramfalk C, Pavlides M, Banerjee R, et al. Fasting plasma insulin concentrations are associated with changes in hepatic fatty acid synthesis and partitioning prior to changes in liver fat content in healthy adults. *Diabetes.* 2016;65(7):1858-1867.
10. Roumans K H M., L L, Veeraiah P, et al. Hepatic saturated fatty acid fraction is associated with de novo lipogenesis and hepatic insulin resistance. *Nature Communications.* 2020;11(1):1891.
11. Association WM. World medical association declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA.* 2013;310(20):2191-2194.
12. Moen CJ, Tholens AP, Voshol PJ, et al. The Hyplip2 locus causes hypertriglyceridemia by decreased clearance of triglycerides. *J Lipid Res.* 2007;48(10):2182-2192.
13. Matikainen N, Adiels M, Söderlund S, et al. Hepatic lipogenesis and a marker of hepatic lipid oxidation, predict postprandial responses of triglyceride-rich lipoproteins. *Obesity (Silver Spring, Md).* 2014;22(8):1854-1859.
14. Jaruvongvanich V, Sanguankeo A, Riangwiwat T, Upala S. Testosterone, sex hormone-binding globulin and nonalcoholic fatty liver disease: A systematic review and meta-analysis. *Annals of Hepatology.* 2017;16(3):382-394.
15. Simo R, Saez-Lopez C, Barbosa-Desongles A, Hernandez C, Selva DM. Novel insights in SHBG regulation and clinical implications. *Trends Endocrinol Metab.* 2015;26(7):376-383.
16. Rocha ALL, Faria LC, Guimarães TCM, et al. Non-alcoholic fatty liver disease in women with polycystic ovary syndrome: systematic review and meta-analysis. *J Endocrinol Invest.* 2017;40(12):1279-1288.
17. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat Med.* 2020;26(2):252-258.
18. Simó R, Barbosa-Desongles A, Lecube A, Hernandez C, Selva DM. Potential role of tumor necrosis factor-α in downregulating sex hormone-binding globulin. *Diabetes.* 2012;61(2):372-382.
19. Simó R, Barbosa-Desongles A, Hernandez C, Selva DM. IL1β down-regulation of sex hormone-binding globulin production by decreasing HNF-4α via MEK-1/2 and JNK MAPK pathways. *Molecular Endocrinology.* 2012;26(11):1917-1927.
20. Wallace IR, McKinley MC, Bell PM, Hunter SJ. Sex hormone binding globulin and insulin resistance. *Clin Endocrinol (Oxf).* 2013;78(3):321-329.
21. Koo SH, Dutcher AK, Towle HC. Glucose and insulin function through two distinct transcription factors to stimulate expression of lipogenic enzyme genes in liver. *Journal of Biological Chemistry.* 2001;276(12):9437-9445.
22. Smith GI, Shankaran M, Yoshino M, et al. Insulin resistance drives hepatic de novo lipogenesis in nonalcoholic fatty liver disease. *Journal of Clinical Investigation.* 2020;130(3):1453-1460.

23. Zimmerman Y, Eijkemans MJC, Coelingh Bennink HJT, Blankenstein MA, Fauser BCJM. The effect of combined oral contraception on testosterone levels in healthy women: A systematic review and meta-analysis. *Human Reproduction Update.* 2014;20(1):76-105.
24. Pasquali R, Vicennati V, Bertazzo D, et al. Determinants of sex hormone-binding globulin blood concentrations in premenopausal and postmenopausal women with different estrogen status. Virgilio-Menopause-Health Group. *Metabolism*. 1997;46(1):5-9.
25. Davison SL, Bell R, Donath S, Montalto JG, Davis SR. Androgen levels in adult females: changes with age, menopause, and oophorectomy. *Journal of Clinical Endocrinology and Metabolism.* 2005;90(7):3847-3853.
26. Maggio M, Lauretani F, Basaria S, et al. Sex hormone binding globulin levels across the adult lifespan in women - the role of body mass index and fasting insulin. *Journal of Endocrinological Investigation.* 2008;31(7):597-601.
27. Jia MC, Zhou LY, Ren S, Dong L, Xiao B. Serum SHBG levels during normal menstrual cycle and after insertion of levonorgestrel-releasing IUD. *Adv Contracept.* 1992;8(1):33-40.
28. Faix D, Neese R, Kletke C, et al. Quantification of menstrual and diurnal periodicities in rates of cholesterol and fat synthesis in humans. *J Lipid Res.* 1993;34(12):2063-2075.

5

CHAPTER FIVE

Serum sex hormone-binding globulin levels are reduced and inversely associated with intrahepatic lipid content and saturated fatty acid fraction in adult patients with glycogen storage disease type 1a

Pomme I.H.G. Simons, Olivier Valkenburg, Ine Telgenkamp, Koen M. van der Waaij, David M. de Groot, Pandichelvam Veeraiah, Judith A.P. Bons, Terry G.J. Derks, Casper G. Schalkwijk, Vera B. Schrauwen-Hinderling, Coen D.A. Stehouwer, Martijn C.G.J. Brouwers

Journal of Endocrinological Investigation; in press

Summary

Purpose: De novo lipogenesis has been inversely associated with serum sex hormone-binding globulin (SHBG) levels. However, the directionality of this association has remained uncertain. We, therefore, studied individuals with glycogen storage disease type 1a (GSD1a), who are characterized by a genetic defect in glucose-6-phosphatase resulting in increased rates of de novo lipogenesis, to assess the downstream effect on serum SHBG levels.

Methods: A case-control study comparing serum SHBG levels in patients with GSD1a (n=10) and controls matched for age, sex, and BMI (n=10). Intrahepatic lipid content and saturated fatty acid fraction were quantified by proton magnetic resonance spectroscopy.

Results: Serum SHBG levels were statistically significantly lower in patients with GSD1a compared to the controls (p=0.041), while intrahepatic lipid content and intrahepatic saturated fatty acid fraction – a marker of de novo lipogenesis – were significantly higher in patients with GSD1a (p=0.001 and p=0.019, respectively). In addition, there was a statistically significant, inverse association of intrahepatic lipid content and saturated fatty acid fraction with serum SHBG levels in patients and controls combined (β: -0.28, 95% CI: -0.47;-0.09 and β: -0.02, 95% CI: -0.04;-0.01, respectively).

Conclusion: Patients with GSD1a, who are characterized by genetically-determined higher rates of de novo lipogenesis, have lower serum SHBG levels than controls.

Introduction

It has long been thought that sex hormone-binding globulin (SHBG) acts only as a simple carrier protein that regulates the bioavailable fraction of testosterone and other sex hormones[1]. In the past decades, however, serum SHBG has also been inversely associated with several metabolic disorders, including obesity, non-alcoholic fatty liver disease, and type 2 diabetes[2-4]. Even more recently, SHBG has been identified as a hepatokine that protects from type 2 diabetes[5,6]. These observations stress the need for a better understanding of the regulation of serum SHBG levels in humans.

In vitro and animal studies have shown that carbohydrate-induced de novo lipogenesis is one of the mechanisms involved in the downregulation of SHBG levels[7]. We have recently extrapolated these findings to humans, by showing that de novo lipogenesis, measured with stable isotopes, is inversely associated with serum SHBG levels[8]. However, given the observational nature of that study, we were unable to assess whether the effect of de novo lipogenesis on serum SHBG is causal. This is of importance, since previous in vitro and animal studies have shown that the association between de novo lipogenesis and SHBG appears to be bidirectional, i.e. serum SHBG may also directly reduce the rates of de novo lipogenesis[9,10].

Monogenetic disorders that affect de novo lipogenesis can be used to unravel whether there is a causal effect of de novo lipogenesis on serum SHBG, in humans. Glycogen storage disease type 1a (GSD1a) is an inborn error of metabolism, caused by a mutation in the *G6PC* gene encoding glucose-6-phosphatase[11]. As a consequence, there is an intrahepatic surplus of glucose-6-phoshate that can serve as a substrate for glycolysis and de novo lipogenesis. Previous studies have shown that patients with GSD1a are indeed characterized by higher rates of de novo lipogenesis and intrahepatic saturated fatty acid fraction, the product of de novo lipogenesis[12,13].

The aim of this study was, therefore, to examine serum SHBG levels in patients with GSD1a and controls matched for age, sex, and BMI, and to study the relationship of intrahepatic lipid (IHL) and saturated fatty acid content with serum SHBG.

Methods

Study design

For this case-control study, we recruited homozygous carriers of a mutation in the gene encoding glucose-6-phosphatase (*G6PC*), causing GSD1a, from outpatient metabolic clinics in the Netherlands and Belgium. Cases were matched to controls

based on factors that are known to affect serum SHBG levels, i.e. age, sex, and BMI[14,15]. To accomplish an adequate matching, we retrieved data for controls from 1) the effects of fructose restriction on liver steatosis [FRUITLESS] study[16], 2) the aldolase B deficiency study[17], and 3) prospective recruitment through local advertisement. All study protocols were similar, with the exception of differences in- and exclusion criteria. The rationale and design of the FRUITLESS and aldolase B deficiency studies have been published previously. In short, the FRUITLESS study was originally conducted to assess the effects of fructose restriction on IHL content[16]. Participants were included if they had a fatty liver index ≥60, and excluded in case of a history of liver disease, excessive alcohol consumption, change in weight or physical activity three months prior to participation, use of glucose-lowering drugs, recent illness, pregnancy and/or lactation[16]. For the current study, only data from baseline measurements were used. The aldolase B deficiency case-control study was originally conducted to compare IHL content in patients with hereditary fructose intolerance and controls. Only data from controls was used in this study[17]. All studies were performed according to the Declaration of Helsinki and approved by the Medical Ethical Committee of Maastricht University Medical Centre[18]. All participants gave written informed consent prior to participation.

All participants had to be at least 18 years of age and were excluded from participation if they had any contraindications for magnetic resonance imaging (MRI) or were unable to give informed consent. Furthermore, as oestrogen-containing medication, which substantially affects serum SHBG levels, are relatively contra-indicated in GSD1a[19-21], controls were excluded if they used oestrogen-containing medication. All controls were asked to visit the metabolic ward after an overnight fast. Since patients with GSD1a develop hypoglycaemia and lactic acidosis upon prolonged fasting, they are treated with either (modified) uncooked cornstarch at night or continuous nocturnal feeding with dextrose[22]. They were, therefore, asked to visit the metabolic ward before having breakfast.

Measurements

For all participants, anthropometrics and quantification of IHL content by proton magnetic resonance spectroscopy (^{1}H-MRS) using a 3T clinical MR scanner (Achieva 3T-X, Philips Healthcare, Best, the Netherlands), were performed as described previously[17]. In a subset of GSD1a patients and controls, the hepatic saturated fatty acid fraction (expressed as the ratio saturated fatty acid:IHL * 100%) was quantified with a newly-developed ^{1}H-MRS method. We previously showed that this fraction correlates well with de novo lipogenesis assessed by stable isotopes[13]. Serum uric acid, total cholesterol, HDL cholesterol, and triglycerides were measured with an enzymatic colorimetric assay (Cobas 8000 instrument, Roche Diagnostics, Mannheim, Germany).

Serum glucose was measured with an enzymatic spectrophotometric assay in patients with GSD1a and a subset of the controls (Cobas 8000 instrument, Roche Diagnostics, Mannheim, Germany). In the controls originating from the aldolase B deficiency and FRUITLESS studies, serum glucose was measured with YSI2300 STAT Plus Glucose Lactate Analyzer (YSI, Yellow Springs, OH, USA)[16,17]. Serum insulin and SHBG were measured with a chemiluminescent immunometric assay in all participants (Immulite XPi instrument, Siemens Healthcare Diagnostics, New Orleans, LA, USA). Serum SHBG measurements demonstrated an analytical variability of 5-7%, with an intra-individual biological variability of 9%[23]. In one patient with GSD1a, blood was drawn directly after the consumption of food, and hence this individual was not included in the analyses of fasting-sensitive measures (i.e. serum lipids, glucose and insulin).

Statistical analyses

Continuous data are presented as median (interquartile range) and categorical data are presented as frequencies. Continuous variables were compared between GSD1a patients and controls by means of Mann-Whitney U test. Univariate regression analyses were performed to study the association of BMI, serum insulin, IHL content and saturated fatty acid fraction with serum SHBG levels. The univariate regression analyses were additionally stratified by condition (i.e. GSD1a or controls). All results were considered statistically significant at $p<0.05$. All statistical analyses were performed using IBM Statistical Package of Social Science (SPSS) version 25.0 for Windows (IBM Corp., Armonk, N.Y., USA).

Results

Population characteristics

The general characteristics of the patients with GSD1a (n=10) and matched controls (n=10; n=5 from FRUITLESS, n=3 from the aldolase B deficiency study, and n=2 prospectively recruited) are shown in Table 5.1. The majority of the study population was female (70%), young (median age: 31.0, IQR: 20.5-47.0 years) and overweight (median BMI: 25.5, IQR: 23.6-29.7 kg/m^2). Age, sex distribution and BMI were, by design, comparable between GSD1a patients and controls (Table 5.1). Serum triglycerides were statistically significantly higher in individuals with GSD1a, while HDL-cholesterol was significantly lower in patients with GSD1a compared to controls (Table 5.1).

Table 5.1 Characteristics of controls and patients with glycogen storage disease type 1a (GSD1a).

	Control (n=10)	GSD1a (n=10)
Sex (male/female), n/n	3/7	3/7
Age, years	32.5 (20.0-56.3)	29.5 (21.5-47.0)
BMI, kg/m^2	26.0 (22.2-31.5)	25.5 (24.5-28.5)
Waist circumference, cm	93.6 (80.5-110.0)	93.5 (82.8-100.4)
Systolic blood pressure, mmHg	129 (112-138)	125 (105-134)
Diastolic blood pressure, mmHg	79 (65-84)	76 (64-81)
Glucose, mmol/l[#]	4.5 (4.2-5.0)	4.4 (3.4-4.8)
Insulin, pmol/l	28.7 (23.0-67.2)	16.5 (10.0-26.6)*
Uric acid, mmol/l	0.34 (0.30-0.35)	0.37 (0.26-0.47)
Total cholesterol, mmol/l	5.1 (4.1-5.4)	5.5 (4.8-6.9)
HDL-cholesterol, mmol/l	1.2 (1.1-1.5)	0.9 (0.7-1.0)*
Triglycerides, mmol/l	1.3 (1.0-1.6)	4.9 (3.7-6.3)*
Alcohol consumption, units/week	0.2 (0.0-5.0)	0.0 (0.0-0.3)

Unless otherwise noted, data are expressed as median (interquartile range). * $p<0.05$ compared to controls, analysed with Mann-Whitney U test. [#] Serum glucose was measured with enzymatic spectrophotometric assay in GSD1a and with enzymatic spectrophotometric assay YSI2300. STAT Plus Glucose Lactate Analyzer in controls, see methods section.
Abbreviations: BMI body mass index; GSD1a glycogen storage disease type 1a; HDL high density lipoprotein; LDL low density lipoprotein.

Patients with GSD1a had statistically significantly lower levels of serum insulin compared to controls (p=0.009) (Figure 5.1, panel A), while the IHL content and saturated fatty acid fraction in patients with GSD1a were statistically significantly higher compared to controls (p=0.001 and p=0.019, respectively) (Figure 5.1, panel B and C, respectively). Serum SHBG levels were statistically significantly lower in GSD1a patients compared to the controls (13.0 nmol/l [IQR: 10.8-25.3] versus 29.5 nmol/l [IQR: 18.5-34.8], respectively; p=0.041; Figure 5.1, panel D). As one male patient used pregnyl (i.e. human chorionic gonadotropin), which may potentially influence serum SHBG levels[24], we repeated the analyses after exclusion of this patient and the matched control, which did not affect the results (p=0.050).

Strikingly, two female patients with GSD1a were found to have high serum SHBG levels relative to the other (female) GSD1a patients, and were notable outliers (Figure 5.1, panel D, open circles). Of interest, these two patients also seemed metabolically healthier, as indicated by relatively low serum insulin levels, IHL content, and saturated fatty acid fraction compared to the values seen for the other patients with GSD1a (Figure 5.1, panel A, B, and C; open circles), whereas serum triglycerides and urate did not appear to differ (data not shown).

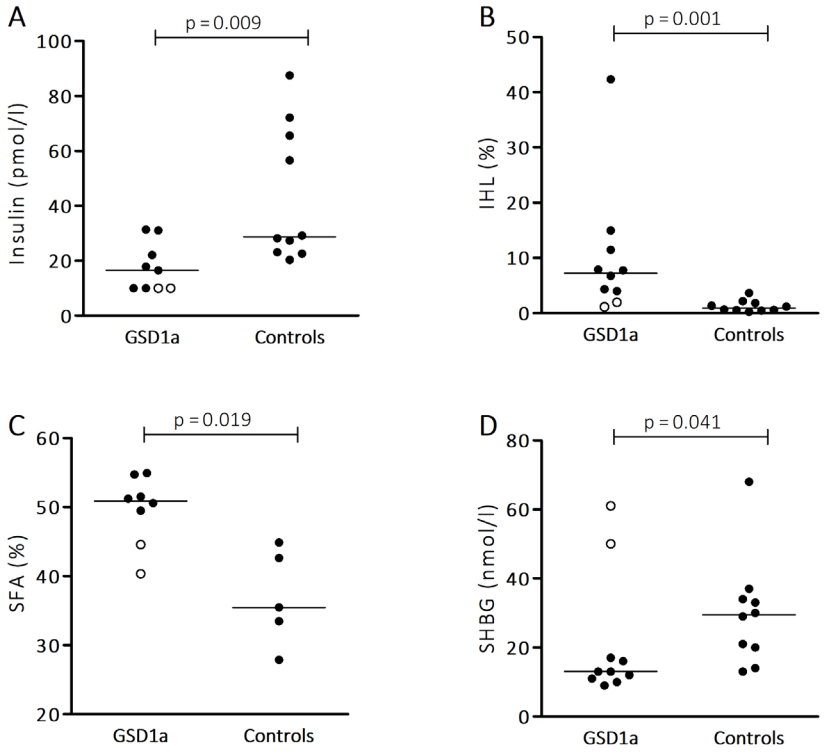

Figure 5.1 Serum insulin (A), intrahepatic lipid (IHL) content (B), saturated fatty acid (SFA) fraction (C) and serum sex hormone-binding globulin (SHBG) (D) in patients with glycogen storage disease type 1a (GSD1a) and controls matched for age, sex, and BMI. Data expressed as individual values with median. Open circles represent GSD1a patients with exceptionally high serum SHBG levels. Differences between the groups were analysed with a Mann-Whitney U test. Of note, fasting serum insulin was unavailable in one GSD1a patient, and SFA fraction was only measured in a subset of cases (n=8) and controls (n=5) (see methods).

Determinants of serum SHBG

In the overall population there was no statistically significant association between BMI or insulin with serum SHBG levels (β: 0.00, 95% CI: -0.03;0.04 and β: 0.00, 95% CI: -0.01;0.01, respectively), or after stratification by condition in the GSD1a group (β: 0.02, 95% CI: -0.06;0.10 and β: -0.02, 95% CI: -0.04;0.01, respectively), or in controls (β: -0.01, 95% CI: -0.04;0.03 and β: 0.00, 95% CI: -0.01;0.01, respectively) (Figure 5.2, panel A and B, respectively). In contrast, there was a statistically significant, inverse association between IHL content and serum SHBG levels in the overall population (β: -0.28, 95% CI: -0.47;-0.09, Figure 5.2, panel C). After stratification for condition, the association remained statistically significant only in patients with GSD1a (GSD1a:

β: -0.47, 95% CI: -0.83;-0.10, controls: β: -0.07, 95% CI: -0.57;0.44). The saturated fatty acid fraction was also statistically significantly inversely associated with lower serum SHBG levels in the overall population (β: -0.02, 95% CI: -0.04;-0.01; Figure 5.2, panel D). Stratified analyses showed a significant, inverse association in GSD1a but not in the controls (β: -0.06, 95% CI: -0.08;-0.03 and β: -0.01, 95% CI: -0.05;0.03, respectively).

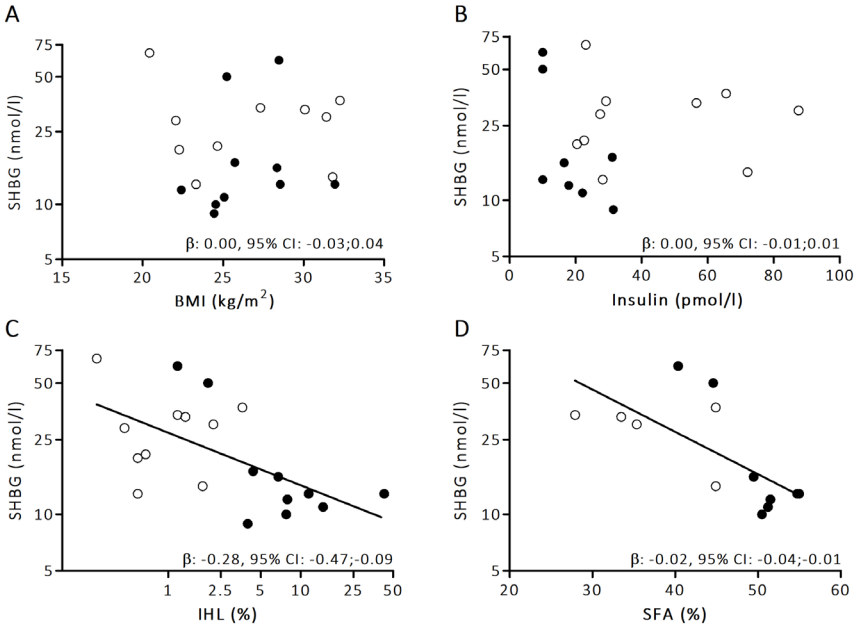

Figure 5.2 Association of body mass index (BMI) (A) serum insulin (B) intrahepatic lipid (IHL) content (C), and saturated fatty acid (SFA) fraction (D) with serum sex hormone-binding globulin (SHBG) levels stratified by glycogen storage disease type 1a (closed circles) and controls matched for age, sex, and BMI (open circles). Regression coefficients (β) and 95% confidence intervals (CI) are reported for the associations in the two groups combined. Of note, fasting serum insulin was unavailable in one GSD1a patient, and SFA was only measured in a subset of patients (n=8) and controls (n=5) (see methods).

Discussion

This study shows that serum SHBG levels are statistically significantly lower in adult patients with GSD1a when compared to controls matched for age, sex, and BMI. In addition, serum SHBG levels varied noticeably within the GSD1a group and appeared to correspond with the variation of other metabolic variables, i.e. IHL content, saturated fatty acid fraction and serum insulin levels. Indeed, there was a statistically significant,

inverse association of IHL content and saturated fatty acid fraction with serum SHBG levels in patients with GSD1a and controls combined.

The current findings support and elaborate on previous experimental and observational studies. In vitro studies have shown that monosaccharide-induced de novo lipogenesis downregulates hepatocyte nuclear factor 4 alpha, thereby decreasing SHBG synthesis in HepG2 cells[7]. In addition, incubation of HepG2 cells with palmitate, the product of de novo lipogenesis, was also found to decrease SHBG levels[7]. Although we have previously extrapolated these in vitro findings to humans by showing that de novo lipogenesis, assessed with stable isotopes, was inversely associated with serum SHBG levels, we could not assess the direction of the association[8]. This is of relevance, since in vitro experiments have also shown that SHBG can affect de novo lipogenesis[9]. By studying patients with GSD1a, who are characterized by a primary genetic defect resulting in high intrahepatic glucose-6-phosphate levels and, consequently, higher rates of de novo lipogenesis[11,12,25], our current observations support the concept that de novo lipogenesis results in lower serum SHBG levels in humans. This conclusion, however, deserves some caution given the small, predominantly female population that was studied, which limits the generalisability of the our findings. Furthermore, our observations do not exclude the reverse, i.e. SHBG affects de novo lipogenesis in humans.

The current findings are in line with large population studies that assessed the effect of a common variant in the glucokinase regulatory protein gene (*GCKR*), which also results in higher intrahepatic glucose-6-phosphate levels and rates of de novo lipogenesis, albeit by a different mechanism. Similar to patients with GSD1a, individuals carrying the *GCKR* minor allele are characterized by lower fasting glucose and insulin levels, higher serum triglycerides, higher rates of de novo lipogenesis and a higher IHL content[26-30]. Previous genome-wide association studies have reported that the *GCKR* minor allele is also associated with lower serum SHBG levels[5,31].

Besides a better knowledge on the causal role of de novo lipogenesis on serum SHBG in humans, the results of this study may also have several clinical implications for patients with GSD1a. First, as recent Mendelian randomization studies have shown that SHBG is causal in the pathogenesis of polycystic ovary syndrome (PCOS), the low serum SHBG levels may contribute to the higher prevalence of PCOS that has been observed in GSD1a[32,33]. Second, we noted that two female patients with GSD1a with relatively high levels of serum SHBG were also characterized by a better metabolic control with respect to IHL content and saturated fatty acid fraction. Serum SHBG could, therefore, serve as a biomarker of metabolic control in patients with GSD1a. In comparison to current measures of metabolic control (e.g. serum triglycerides), serum SHBG has several advantages. The half-life of serum SHBG is relatively long (7 days) and can

therefore reflect metabolic control over the past days[34]. In addition, serum SHBG levels are not directly affected by the nutritional state, i.e. fasted or fed, in contrast to serum triglycerides. Further studies are needed to assess whether intra-individual variations in serum SHBG adequately reflect changes in metabolic control in GSD1a patients.

This study has several strengths and limitations. The use of GSD1a as a model for higher rates of de novo lipogenesis is a unique approach to study the direct effects of de novo lipogenesis on serum SHBG levels in an observational setting. However, the rare nature of this disease (prevalence ~1 in 100,000 births[35]) results in small numbers and, consequently, low statistical power, in particular for the stratified univariate analyses. In addition, as a result of the small sample size, we were unable to explore the associations in men and women separately, or adjust for potential confounders. We, therefore, decided to match the GSD1a patients with controls based on factors that are known to have a substantial effect on serum SHBG levels. Despite adequate matching, however, it cannot be excluded that there may be residual confounding. Another limitation of this study is that we were unable to account for phase of menstrual cycle or endogenous sex hormone levels. Nevertheless, the phase of menstrual cycle is likely to have scattering effects, as de novo lipogenesis varies throughout the menstrual cycle while serum SHBG levels remain constant[36,37]. Despite this scattering effect, we did observe a statistically significant difference in serum SHBG levels between GSD1a patients and controls. A possible effect of endogenous oestrogens or androgens on de novo lipogenesis and serum SHBG deserves further investigation. Furthermore, as a result of the extensive matching, data for healthy controls had to be retrieved from several studies. The in- and exclusion criteria varied between the studies, and, consequently, controls originating from the FRUITLESS study are metabolically less healthy when compared to the general population[16]. This could explain the relatively low serum SHBG levels in the control group, and, hence, could have mitigated the difference in serum SHBG levels between GSD1a and controls. Finally, because of the extreme phenotype of GSD1a, there may have been pleiotropic effects that have contributed to the current findings. For instance, given the fasting intolerance in GSD1a, all measures in patients with GSD1a were conducted after (modified) uncooked cornstarch at night or continuous nocturnal feeding with dextrose (but before breakfast). This may have affected some of the fasting-sensitive outcome measures including serum triglycerides, insulin and glucose. Of note, previous studies have shown that a recent meal only mildly affects IHL content[38]. The IHL content in GSD1a patients in the current study was much higher than what can be expected from a recent meal. Furthermore, previous studies have shown that serum SHBG levels are not affected by a recent meal[39].

In conclusion, in the present study we found that patients with GSD1a, who are characterized by genetically-determined higher rates of de novo lipogenesis, have statistically significantly lower levels of serum SHBG than controls.

References

1. Hammond GL, Wu TS, Simard M. Evolving utility of sex hormone-binding globulin measurements in clinical medicine. *Current Opinion in Endocrinology, Diabetes, and Obesity.* 2012;19(3):183-189.
2. Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: A systematic review and meta-analysis. *JAMA.* 2006;295(11):1288-1299.
3. Cooper LA, Page ST, Amory JK, Anawalt BD, Matsumoto AM. The association of obesity with sex hormone-binding globulin is stronger than the association with ageing--implications for the interpretation of total testosterone measurements. *Clinical Endocrinology.* 2015;83(6):828-833.
4. Jaruvongvanich V, Sanguankeo A, Riangwiwat T, Upala S. Testosterone, sex hormone-binding globulin and nonalcoholic fatty liver disease: A systematic review and meta-analysis. *Annals of Hepatology.* 2017;16(3):382-394.
5. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nature Medicine.* 2020;26(2):252-258.
6. Stefan N, Häring HU. The role of hepatokines in metabolism. *Nature Reviews: Endocrinology.* 2013;9(3):144-152.
7. Selva DM, Hogeveen KN, Innis SM, Hammond GL. Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. *Journal of Clinical Investigation.* 2007;117(12):3979-3987.
8. Simons PIHG, Valkenburg O, Telgenkamp I, et al. Relationship between de novo lipogenesis and serum sex hormone binding globulin in humans. *Clinical Endocrinology.* 2021;95(1):101-106.
9. Saez-Lopez C, Barbosa-Desongles A, Hernandez C, et al. Sex hormone-binding globulin reduction in metabolic disorders may play a role in NAFLD development. *Endocrinology.* 2017;158(3):545-559.
10. Sáez-López C, Salcedo-Allende MT, Hernandez C, et al. Sex hormone-binding globulin expression correlates with acetyl-coenzyme a carboxylase and triglyceride content in human liver. *Journal of Clinical Endocrinology and Metabolism.* 2019;104(5):1500-1507.
11. Rake JP, Visser G, Labrune P, et al. Glycogen storage disease type I: Diagnosis, management, clinical course and outcome. Results of the European Study on Glycogen Storage Disease Type I (ESGSD I). *European Journal of Pediatrics.* 2002;161 Suppl 1:S20-34.
12. Bandsma RH, Prinsen BH, van Der Velden Mde S, et al. Increased de novo lipogenesis and delayed conversion of large VLDL into intermediate density lipoprotein particles contribute to hyperlipidemia in glycogen storage disease type 1a. *Pediatric Research.* 2008;63(6):702-707.
13. Roumans K H M., L L, Veeraiah P, et al. Hepatic saturated fatty acid fraction is associated with de novo lipogenesis and hepatic insulin resistance. *Nature Communications.* 2020;11(1):1891.
14. Cooper LA, Page ST, Amory JK, Anawalt BD, Matsumoto AM. The association of obesity with sex hormone-binding globulin is stronger than the association with ageing - implications for the interpretation of total testosterone measurements. *Clinical Endocrinology.* 2015;83(6):828-833.
15. Maggio M, Lauretani F, Basaria S, et al. Sex hormone binding globulin levels across the adult lifespan in women - the role of body mass index and fasting insulin. *Journal of Endocrinological Investigation.* 2008;31(7):597-601.
16. Simons N, Veeraiah P, Simons PIHG, et al. Effects of fructose restriction on liver steatosis (FRUITLESS); a double-blind randomized controlled trial. *The American Journal of Clinical Nutrition.* 2020;113(2): 91-400.
17. Simons N, Debray F-G, Schaper NC, et al. Patients with aldolase b deficiency are characterized by increased intrahepatic triglyceride content. *Journal of Clinical Endocrinology and Metabolism.* 2019;104(11):5056-5064.
18. Association WM. World medical association declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA.* 2013;310(20):2191-2194.
19. Kishnani PS, Austin SL, Abdenur JE, et al. Diagnosis and management of glycogen storage disease type I: a practice guideline of the American College of Medical Genetics and Genomics. *Genetics in Medicine.* 2014;16(11):e1-e1.
20. Mairovitz V, Labrune P, Fernandez H, Audibert F, Frydman R. Contraception and pregnancy in women affected by glycogen storage diseases. *European Journal of Pediatrics.* 2002;161(1):S97-S101.

21. Zimmerman Y, Eijkemans MJC, Coelingh Bennink HJT, Blankenstein MA, Fauser BCJM. The effect of combined oral contraception on testosterone levels in healthy women: A systematic review and meta-analysis. *Human Reproduction Update.* 2014;20(1):76-105.
22. Rake JP, Visser G, Labrune P, et al. Glycogen storage disease type I: diagnosis, management, clinical course and outcome. Results of the European Study on Glycogen Storage Disease Type I (ESGSD I). *Eur J Pediatr.* 2002;161 Suppl 1:S20-34.
23. Aarsand A, Fernandez-Calle P, Webster C, et al. The EFLM Biological Variation Database. https://biologicalvariation.eu/. Accessed2021.
24. Belgorosky A, Rivarola MA. Sex hormone-binding globulin response to human chorionic gonadotropin stimulation in children with cryptorchidism, anorchia, male pseudohermaphroditism, and micropenis. *Journal of Clinical Endocrinology and Metabolism*. 1982;54(4):698-704.
25. Hoogerland JA, Peeks F, Hijmans BS, et al. Impaired very-low-density lipoprotein catabolism links hypoglycemia to hypertriglyceridemia in Glycogen Storage Disease type Ia. *Journal of inherited metabolic disease*. 2021;44(4):879-892.
26. Santoro N, Caprio S, Pierpont B, et al. Hepatic de novo lipogenesis in obese youth is modulated by a common variant in the GCKR gene. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(8):E1125-E1132.
27. Rees MG, Wincovitch S, Schultz J, et al. Cellular characterisation of the GCKR P446L variant associated with type 2 diabetes risk. *Diabetologia.* 2012;55(1):114-122.
28. Beer NL, Tribble ND, McCulloch LJ, et al. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Human Molecular Genetics.* 2009;18(21):4081-4088.
29. Lin YC, Chang PF, Chang MH, Ni YH. Genetic variants in GCKR and PNPLA3 confer susceptibility to nonalcoholic fatty liver disease in obese individuals. *American Journal of Clinical Nutrition.* 2014;99(4):869-874.
30. Brouwers MCGJ, Jacobs C, Bast A, Stehouwer CDA, Schaper NC. Modulation of glucokinase regulatory protein: A double-edged sword? *Trends in Molecular Medicine*. 2015;21(10):583-594.
31. Coviello AD, Haring R, Wellons M, et al. A genome-wide association meta-analysis of circulating sex hormone–binding globulin reveals multiple loci implicated in sex steroid hormone regulation. *Plos Genetics.* 2012;8(7):e1002805.
32. Lee PJ, Patel A, Hindmarsh PC, Mowat AP, Leonard JV. The prevalence of polycystic ovaries in the hepatic glycogen storage diseases: Its association with hyperinsulinism. *Clinical Endocrinology.* 1995;42(6):601-606.
33. Sechi A, Deroma L, Lapolla A, et al. Fertility and pregnancy in women affected by glycogen storage disease type I, results of a multicenter Italian study. *Journal of Inherited Metabolic Disease*. 2013;36(1):83-89.
34. Selby C. Sex hormone binding globulin: origin, function and clinical significance. *Annals of Clinical Biochemistry.* 1990;27 (Pt 6):532-541.
35. Froissart R, Piraud M, Boudjemline AM, et al. Glucose-6-phosphatase deficiency. *Orphanet Journal of Rare Diseases.* 2011;6(1):27.
36. Jia MC, Zhou LY, Ren S, Dong L, Xiao B. Serum SHBG levels during normal menstrual cycle and after insertion of levonorgestrel-releasing IUD. *Advances in Contraception.* 1992;8(1):33-40.
37. Faix D, Neese R, Kletke C, et al. Quantification of menstrual and diurnal periodicities in rates of cholesterol and fat synthesis in humans. *Journal of Lipid Research.* 1993;34(12):2063-2075.
38. Lindeboom L, Nabuurs CI, Hesselink MK, et al. Proton magnetic resonance spectroscopy reveals increased hepatic lipid content after a single high-fat meal with no additional modulation by added protein. *American Journal of Clinical Nutrition.* 2015;101(1):65-71.
39. Habito RC, Ball MJ. Postprandial changes in sex hormones after meals of different composition. *Metabolism.* 2001;50(5):505-511.

6

CHAPTER SIX

Sex hormone-binding globulin: biomarker *and* hepatokine?

Pomme I.H.G. Simons, Olivier Valkenburg, Coen D.A. Stehouwer, Martijn C.G.J. Brouwers

Trends in Endocrinology and Metabolism (2021); 32(8): 544-553

Summary

Over the past decade, there have been important breakthroughs in our understanding of the regulation and function of sex hormone-binding globulin (SHBG). A recent genome-wide association and Mendelian randomization study has provided new insights at the population level. Thorough study of genetic variants affecting serum SHBG has identified de novo lipogenesis as one of the mechanistic links between the metabolic syndrome and reduced serum SHBG levels in humans. Furthermore, careful deduction of the Mendelian randomization results suggests a direct, causal role for SHBG in the pathogenesis of type 2 diabetes, as a hepatokine, in women. These findings prompt the development of SHBG-raising therapies as a means to prevent or treat disorders such as type 2 diabetes and polycystic ovary syndrome.

Highlights

- A recent genome-wide association study has identified a cluster of de novo lipogenesis genes that affect serum SHBG levels.
- A recent Mendelian randomization study has shown that SHBG functions as a true hepatokine, at least in women, by reducing the risk of type 2 diabetes.
- SHBG-raising therapies, such as inhibitors of de novo lipogenesis or thyroid hormone receptor beta agonists, may provide promising future directions for the treatment and prevention of type 2 diabetes and polycystic ovary syndrome.

The traditional view on sex hormone-binding globulin

Since its discovery by Mercier et al. in 1966[1], sex hormone-binding globulin (SHBG) has been viewed as the principal protein that binds circulating sex hormones with high affinity, primarily 5-alpha-dihydrotestosterone, testosterone, and 17-beta-estradiol[2,3]. As such, according to the **free hormone hypothesis** (see Glossary), it regulates the bioavailability of these sex hormones at the target site[4,5].

A recent genome-wide association study (GWAS) and **Mendelian randomization** analysis have provided new insights into the regulation and function of SHBG in humans[6]. On the basis of these outcomes, we postulate that SHBG is both a biomarker of metabolic derangements, such as **de novo lipogenesis**, and a protein that exerts systemic metabolic effects by itself, a so-called **hepatokine**. In this opinion, we provide an extensive description of this landmark study and elaborate on the interpretation and clinical implications of the outcomes.

Current knowledge on the regulation and function of SHBG

SHBG is synthesized primarily in the liver as a homodimeric glycoprotein[7-10]. The expression of the *SHBG* gene, located on chromosome 17p13.1[11], is under transcriptional control of hepatocyte nuclear factor 4 alpha (HNF-4α) and constitutive androstane receptor (both stimulatory), as well as peroxisome proliferator-activated receptor gamma and chicken ovalbumin upstream promotor transcription factor (both inhibitory)[12-14]. These transcription factors are affected by several hormonal, metabolic, nutritional, and inflammatory factors, among others thyroid hormone[15], adiponectin[16], and several cytokines, including tumor necrosis factor alpha, which downregulates HNF-4α through nuclear factor-κB[17], and interleukin-1β, which downregulates HNF-4α through mitogen-activated protein kinase (MAPK)/extracellular signaling regulated kinase-1/2 and c-Jun N terminal kinase MAPK signaling[18]. The complex regulation of SHBG synthesis has been extensively reviewed elsewhere[19-22].

There is abundant epidemiological evidence that serum SHBG levels are reduced in several metabolic disorders, including obesity, type 2 diabetes and polycystic ovary syndrome (PCOS)[21,23,24]. Furthermore, serum SHBG levels are inversely associated with the metabolic syndrome[25,26] and its individual components, except for blood pressure[27]. For a long time, it was widely accepted that the common denominator of these entities, hyperinsulinemia, accounted for the reduced serum SHBG levels[28-32]. However, growing evidence pleads against a direct role of insulin in the regulation of

SHBG[19,33-35]. For instance, streptozotocin-induced insulin deficiency in transgenic mice carrying human *SHBG* led to a decrease rather than an increase in SHBG levels[33].

Selva et al. provided an alternative explanation for the altered SHBG levels in the hitherto mentioned metabolic disorders[33]. They showed that hepatic de novo lipogenesis – one of the principal pathways involved in the accumulation of intrahepatic lipids[36] – contributed to the regulation of serum SHBG synthesis[33]. In vitro and animal experiments revealed that the conversion of monosaccharides (i.e., fructose and glucose) to palmitate inhibited the expression of HNF-4α and, subsequently, SHBG (Figure 6.1, panel A)[33]. Of interest, other laboratory studies have shown that the reverse may also be possible: SHBG downregulates de novo lipogenesis and thereby reduces intrahepatic lipids[37]. Although these mechanisms have not been confirmed in humans, it is of interest that intrahepatic lipid content is increased in obesity, type 2 diabetes, and PCOS and also inversely associated with serum SHBG levels[21,23,24,38-40]. Of note, other factors, such as inflammatory cytokines and adiponectin, are also altered in obesity, type 2 diabetes, and PCOS and likely contribute to the reduced SHBG levels as well[16-18,41-44].

Ten years ago, Mendelian randomization studies suggested that the function of SHBG is broader than only a carrier protein by showing that genetically predicted SHBG levels are inversely associated with type 2 diabetes[45,46]. Although these Mendelian randomization studies (Box 6.1) inferred a causal role for SHBG in the pathogenesis of type 2 diabetes, as a hepatokine, it cannot be excluded that the effects of SHBG on type 2 diabetes risk are mediated by free testosterone levels, which are inevitably affected by any change in SHBG.

Since then, various experimental studies have further elucidated the mechanisms by which SHBG could affect type 2 diabetes[37,47,48]. Humanized transgenic *SHBG* mice fed a high-fat diet were characterized by an increased white adipose tissue lipolysis and improved glucose homeostasis compared with wild-type mice. Furthermore, the effects on lipolysis were also observed in human mature adipocytes treated with SHBG, even when performed under hormone-deprived conditions[48]. These in vitro experiments allow distinction between direct, primary effects of SHBG and secondary, downstream effects mediated by free testosterone (or oestrogen).

Box 6.1 The concept of Mendelian randomization.

Mendelian randomization is an epidemiological approach that uses genetic variants as instruments to infer causality in observational study designs[89]. At conception, each individual randomly receives an exposure-predisposing or exposure-protective allele that is subject to the Mendelian law of inheritance. Mendelian randomization uses this random distribution of genetic variants in large groups of individuals to study the effect of a certain exposure on an outcome of interest[89]. As such, Mendelian randomization draws many parallels to a randomized controlled trial, which is regarded as the gold standard to study causality.

For instance, common variants in the *SHBG* gene, which are known to affect serum SHBG levels, have been inversely associated with type 2 diabetes risk[46]. Subsequent Mendelian randomization analyses demonstrated that, per unit (of the natural log standard deviation) increase in (genetically predicted) SHBG levels there was an ~72% decreased risk of type 2 diabetes[45].

A well-conducted, unbiased Mendelian randomization analysis is subject to three main assumptions[62]. First, the genetic variant(s) must (robustly) associate with the exposure trait of interest. Second, the genetic variant(s) may not associate with any confounder of the association between the exposure and outcome (known as horizontal pleiotropy). Violation of this assumption can, in part, be tested with an Mendelian randomization-Egger regression. Horizontal pleiotropy should be distinguished from vertical pleiotropy, which refers to the downstream effects of an exposure, and is not a violation of the assumption. Finally, the genetic variant(s) may not be associated with the outcome, other than through the exposure trait.

When these assumptions are met, Mendelian randomization can be a valuable asset in deducing unbiased causal effect estimates in an observational study design. Although the effects of genetic variants on a given phenotype are generally small[90], they do represent a lifelong exposure to the variable of interest. Furthermore, by using genetic variants as instrumental variables, the association of interest is less subject to reverse causation and the influence of confounding factors[89]. Consequently, Mendelian randomization can overcome some of the inherent biases present in traditional observational studies.

New insights into the regulation of SHBG in humans: SHBG as a biomarker

Ruth et al. recently presented the outcomes of a large-scale GWAS for serum SHBG, total testosterone, and free testosterone in men and women using data from the UK Biobank (n=425,097), a population-based cohort study[6]. They identified 658 genetic variants that contributed to the variation in serum SHBG levels (adjusted for body mass index). Careful study of these genes provided evidence of the role of intrahepatic lipid accumulation, and more specifically de novo lipogenesis, in the regulation of SHBG synthesis in humans (Figure 6.1, panel B).

Besides variants in the *SHBG* (rs1799941 and rs6258) and *HNF4α* (rs6073431) genes, several genetic variants that are known to be involved in de novo lipogenesis were associated with lower SHBG levels[6]. First, the GWAS found that a common variant in *GCKR* (rs1260326) was strongly associated with serum SHBG ($p=1.3*10^{-298}$ in men;

p=1.7*10^{-188} in women)[6] and replicated previous findings in smaller cohorts[49,50]. *GCKR* encodes glucokinase regulatory protein (GKRP), a liver-specific protein that binds and thereby inactivates glucokinase, the first, rate-limiting enzyme in glycolysis. Experimental studies have shown that the rs1260326 minor allele encodes a GKRP protein that binds glucokinase less effectively[51,52], resulting in an increased hepatic glucose disposal, glycolysis, and subsequent de novo lipogenesis and intrahepatic lipid accumulation[53-55]. Second, the GWAS also identified a variant in the glucokinase gene (*GCK*; rs1799831) that was associated with serum SHBG levels[6]. Third, serum SHBG levels were affected by a variant in *MLXIPL* (rs17145750)[6], which encodes carbohydrate response element binding protein (ChREBP). ChREBP is one of the major regulators of de novo lipogenesis by serving as a transcription factor of lipogenic enzymes, as well as GKRP and glucokinase[56]. Of note, several genetic variants upstream of sterol regulatory element-binding protein 1c (SREBP1c), another major regulator of de novo lipogenesis, such as insulin receptor substrate 1 (*IRS1*), sterol regulatory element-binding protein cleavage-activating protein (*SCAP*), and phosphatase and tensin homolog (*PTEN*), were also associated with serum SHBG[6]. In contrast, despite the important role of fatty acid synthase, the enzyme that catalyzes the formation of palmitate, in the regulation of SHBG synthesis[33], variants in the fatty acid synthase gene (*FASN*) were not associated with SHBG levels[6]. Whether this absent association is due to the stringent significance thresholds that are applied in GWAS deserves further study.

The striking abundance of genes affecting hepatic de novo lipogenesis in this GWAS of SHBG coincided with the absence of genes affecting intrahepatic lipid content by other pathways than de novo lipogenesis. For instance, variants in *TM6SF2*, which affect intrahepatic lipid content through impairment of very-low-density lipoprotein (VLDL) secretion[57], were not identified[6]. Furthermore, a common variant in the patatin-like phospholipase domain-containing protein 3 gene (*PNPLA3;* rs738409), which is positively associated with intrahepatic lipid content[58], was also positively associated with serum SHBG levels[6]. This variant in *PNPLA3* has been associated with disturbed lipid remodeling and, consequently, impaired VLDL secretion[59]. However, the same variant has also been associated with impaired de novo lipogenesis[60], which may explain the apparent paradoxical positive association with SHBG (Figure 6.1, panel B). These findings imply that de novo lipogenesis specifically, rather than intrahepatic lipids per se, drives the decrease in SHBG levels. These findings at the population level corroborate with previous experimental studies (Figure 6.1, panel A versus B)[33] and also confirm that serum SHBG is a biomarker of metabolic processes in the liver, including de novo lipogenesis.

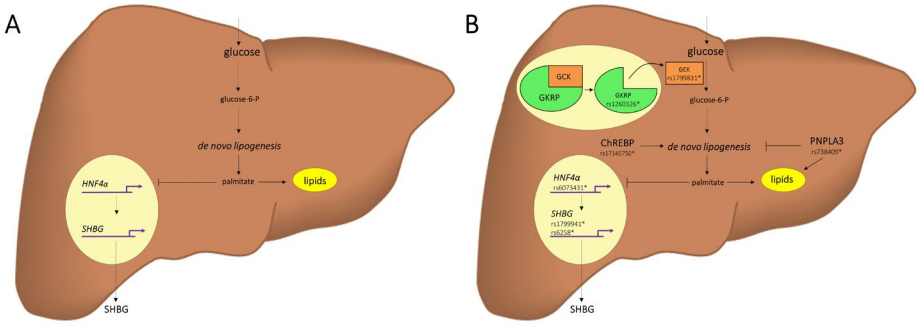

Figure 6.1 Regulation of sex hormone-binding globulin (SHBG) synthesis by carbohydrate-induced de novo lipogenesis, based on previous in vitro and animal experiments and a recent genome-wide association study.
Panel A: Palmitate, the product of de novo lipogenesis from carbohydrates, inhibits the expression of hepatocyte nuclear factor 4 alpha (*HNF4α*) and, consequently, SHBG synthesis[33]. Panel B: a recent genome-wide association study identified several genetic variants (indicated with RefSNP [rs] numbers) that encode proteins involved in de novo lipogenesis and that have been associated with serum SHBG levels[6]. First is a variant in the gene encoding glucokinase regulatory protein (GKRP) (rs1260326), a protein that binds and inactivates glucokinase. Second is a genetic variant in glucokinase (*GCK*) (rs1799831), an enzyme that facilitates the conversion of glucose to glucose-6-phosphate, the first, rate-limiting step in glycolysis. Third is a variant in the gene encoding carbohydrate response element binding protein (ChREBP) (rs17145750), an important transcription factor of lipogenic enzymes. Fourth is a genetic variant in *HNF4α* (rs6073431), an important transcription factor of SHBG, and genetic variants of *SHBG* (rs1799941 and rs6258). These variants were all inversely associated with SHBG levels. Finally, a variant in the patatin-like phospholipase domain-containing protein 3 gene (*PNPLA3*) (rs738409), which inhibits de novo lipogenesis and predisposes to intrahepatic lipid accumulation via an alternate pathway, was positively associated with SHBG levels.

New insights into the causal role of SHBG in type 2 diabetes and PCOS: SHBG as a hepatokine

Ruth et al. also conducted a Mendelian randomization study to assess the relationship of serum SHBG, total testosterone, and free testosterone with several complex disease traits, among others type 2 diabetes and PCOS[6]. In line with previous findings[45,46], the Mendelian randomization analyses showed an inverse association between genetically predicted SHBG levels and risk of type 2 diabetes in both men and women (Figure 6.2, panel A and D, black bars). Furthermore, in accordance with a previous study[61], genetically determined SHBG was also inversely associated with PCOS (Figure 6.2, panel A, pink bars)[6]. Since these analyses used the 658 genetic variants that affect serum SHBG levels as instruments, including the genetic variants that affect de novo lipogenesis, it could be argued that the observed associations are due not to SHBG

itself, but instead to the metabolic changes downstream from processes such as de novo lipogenesis, independent from SHBG (= horizontal pleiotropy, see Box 6.1). However, the authors performed statistical tests, (i.e. Mendelian randomization-Egger regression) to rule out this genetic form of confounding. Moreover, the authors also used only variants in the *SHBG* gene as instruments, which are less prone to horizontal pleiotropy[62,63], and observed similar results.

The subsequent Mendelian randomization analyses for total and free testosterone (Figure 6.2, panel A and D) and the sexual dimorphism that was observed allow differentiation between the primary, direct effects of SHBG on these traits and the secondary effects mediated by free testosterone levels. In this respect, it is also important to consider the (genetic) inter-relationships between SHBG, total testosterone, and free testosterone in men and women, as can be derived from the GWAS results[6].

In women, genetic variants that determine SHBG levels are also inversely associated with free testosterone levels (Figure 6.2, panel B, filled circles), but not with total testosterone levels (Figure 6.2, panel B, open circles)[6]. Furthermore, genetic variants that determine total testosterone are also positively associated with free testosterone levels (Figure 6.2, panel C)[6]. These associations are probably explained by the independent regulation of SHBG (by the liver) and testosterone synthesis (by the adrenals and ovaries), while both affect free testosterone levels (Figure 6.3, panel A). Of note, since a (genetic) change in SHBG alters free testosterone levels, there does not seem to be a feedback loop through the hypothalamic-pituitary-gonadal (HPG) axis that would be expected to maintain free testosterone levels. On the basis of these associations, it is likely that the causal associations of both SHBG and total testosterone with PCOS (Figure 6.2, panel A, pink bars) are explained by their effects on free testosterone (Box 6.2). This is in fact not surprising, given the adoption of (biochemical) hyperandrogenism as a diagnostic criterion of PCOS[64]. In contrast, the absence of a causal association between total testosterone and type 2 diabetes in women (Figure 6.2, panel A, black bars) strongly suggests that free testosterone is not a causal factor in the pathogenesis of type 2 diabetes in women (Box 6.2).

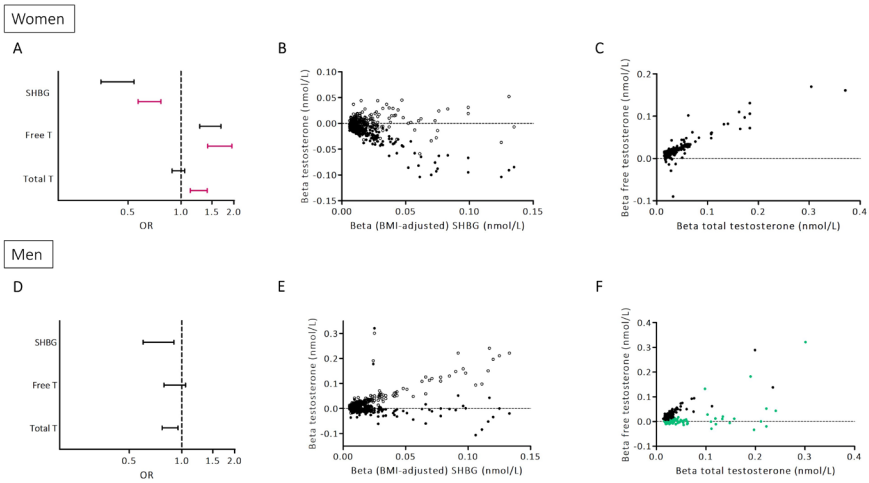

Figure 6.2 Interrelationships between sex hormone-binding globulin (SHBG), total testosterone, and free testosterone, and their associations with type 2 diabetes and polycystic ovary syndrome (PCOS)[6].

Panel A-C; panel A: In women, Mendelian randomization analyses suggested a causal association of SHBG and free testosterone, but not total testosterone, with type 2 diabetes (black). In contrast, all variables, including total testosterone, were causally related to PCOS (pink). Panel B: genetic variants that affect SHBG in women, also affect free testosterone (filled circles), but not total testosterone (open circles). Panel C: genetic variants that affect total testosterone in women also affect free testosterone.

Panel D-F; panel D: in men, Mendelian randomization analyses suggested a causal association of SHBG and total testosterone, but not free testosterone, with type 2 diabetes. Panel E: genetic variants that affect SHBG also affect total testosterone (open circles), but not free testosterone (filled circles). Panel F: genetic variants affecting total testosterone *and* SHBG levels generally have a neutral effect on free testosterone (green), while genetic variants affecting only total testosterone levels have a positive effect on free testosterone (black).

Data are derived from supplementary data of reference[6]. For visualization purposes, two genetic variants have been left out of panel B (rs545206972, *MDPU1*) and panels E and F (rs6258; *SHBG*). The effect sizes of these variants is far greater than that of the other genetic variants.

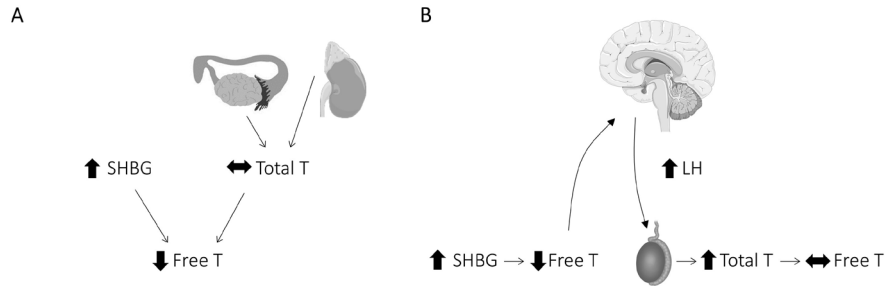

Figure 6.3 The proposed role of the hypothalamic-pituitary-gonadal (HPG) axis in the regulation of free testosterone levels in men and women.
Panel A: In women, a genetic increase in serum sex hormone-binding globulin (SHBG) will result in a decrease in free testosterone. It is anticipated that this does not result in a feedback loop through the HPG-axis, and hence total testosterone levels, which are independently regulated by the ovaries and adrenals, remain unaffected. Panel B: In men, a genetic increase in serum SHBG will decrease free testosterone levels. This initiates a feedback loop via the HPG-axis resulting in increased total testosterone levels and restoration of free testosterone levels.
These biological mechanisms may explain the observations shown in Figure 6.2, panel B, C, E, and F.

In men, the inter-relationships between SHBG, total testosterone, and free testosterone differ from those in women and appear to be more complex. First, in contrast to women, genetic variants that determine SHBG levels are positively associated with total testosterone (Figure 6.2, panel E, open circles), but not with free testosterone (Figure 6.2, panel E, filled circles)[6]. Second, there appears to be a dichotomy in the relationship between genetically determined total testosterone and free testosterone: some genetic variants that increase total testosterone levels also affect free testosterone concentrations, whereas others have a neutral effect (Figure 6.2, panel F)[6]. Of interest, the majority of the latter group also has an effect on SHBG levels (Figure 6.2, panel F, green circles)[6]. These differences may be explained by considering the role of the HPG axis in maintaining free testosterone levels in men. Any (genetic) change in SHBG will result in altered free testosterone levels, which are restored by the HPG axis by changing the production of total testosterone (Figure 6.3, panel B). Consequently, it is anticipated that those genetic variants that affect total *and* free testosterone levels (Figure 6.2, panel F, black circles) are in some way involved in the HPG axis. Indeed, careful study of these genes reveals variants that have been associated with hypogonadotropic hypogonadism[6], either directly (*KISS1*, *FGFR1*[65,66]) or secondarily via an effect on obesity (*FTO, LEPR*[67,68]).

Box 6.2 Is SHBG a true hepatokine? – Application of Occam's razor.

Careful deduction of the sexually dimorphic results of the Mendelian randomization analyses for SHBG, total testosterone, and free testosterone allows differentiation between a primary, direct effect of SHBG or a secondary effect, mediated by free testosterone, on type 2 diabetes and PCOS risk.

Women

Under the following assumptions, it can be derived that free testosterone is the causal factor in the risk of PCOS:

1. Free testosterone is the only bioactive androgen.
2. Free testosterone is determined by both (genetically predicted) total testosterone and SHBG (Figure 6.2, panel B and C).
3. (Genetically predicted) total testosterone and SHBG are not interdependent (Figure 6.2, panel B).

Genetically predicted SHBG *and* total testosterone, both determinants of free testosterone, are associated with PCOS risk (Figure 6.2, panel A, pink bars). In contrast, a different pattern was observed for type 2 diabetes risk (Figure 6.2, panel A, black bars). If free testosterone was the causal factor in the risk of type 2 diabetes, it would have been expected that both determinants of free testosterone (i.e. SHBG *and* total testosterone) were associated with type 2 diabetes risk. Since only genetically predicted SHBG, but *not* total testosterone, was associated with type 2 diabetes risk, these results strongly suggest that SHBG plays a primary, direct role in the pathogenesis of type 2 diabetes.

Of note, the authors also conducted a Mendelian randomization analysis with genetic variants that affect total or free testosterone without any effect on SHBG (a 'testosterone cluster'). This cluster was not associated with type 2 diabetes risk, in support of a primary role of SHBG. Unfortunately, there was no 'SHBG cluster' that included genes with effects on only SHBG, not total or free testosterone, to confirm the conclusions of this deduction.

Men

Under the following assumptions, it can be deduced that the inverse association of genetically predicted SHBG and total testosterone, but not free testosterone, with type 2 diabetes (Figure 6.2, panel D) should be explained by SHBG:

1. Free testosterone is the only bioactive androgen.
2. (Genetically predicted) SHBG determines total testosterone, but not free testosterone (Figure 6.2, panel E).
3. (Genetically predicted) total testosterone affects free testosterone (Figure 6.2, panel F).

However, repeat analysis with the 'testosterone cluster' did show an inverse association with type 2 diabetes risk in men. This discrepancy can be explained by either a type I error (i.e., the observed association between the testosterone cluster and type 2 diabetes is a false positive), a type II error (i.e., the absent association between free testosterone and type 2 diabetes [Figure 6.2, panel D] is a false negative) or refutation of the free hormone hypothesis (assumption 1). These possible explanations deserve further study.

On the basis of these associations, it can be deduced that the causal, inverse association between total testosterone (but not free testosterone) and type 2 diabetes appears to be secondary to the causal, inverse association between SHBG and type 2 diabetes (Figure 6.2, panel D) (Box 6.2). However, Ruth et al. performed additional analyses, which contrast with this conclusion and do not exclude an active role for free testosterone (Box 6.2).

Clinical implications: SHBG-raising therapies

A causal, inverse association between serum SHBG and complex disease traits – either primary (as appears to be the case for type 2 diabetes risk in women) or secondary via free testosterone (i.e., PCOS) – justifies further research and development of SHBG-raising therapies. Such therapies may also be beneficial in the treatment of non-alcoholic fatty liver disease[37]. Currently, several interventions have been shown to increase serum SHBG levels, all with potential benefits and harms.

First, lifestyle interventions have been shown to increase serum SHBG levels[13,69-72]. Studies have identified a beneficial effect of components of the Mediterranean diet (i.e., oleoyl-coenzyme A and resveratrol as found in olive oil and red wine, respectively) on SHBG levels[13,72]. Furthermore, clinically relevant effects on SHBG levels have been observed for lifestyle interventions that achieve weight loss[69-71]. An average weight reduction of 10 kilograms resulted in an ~26% increase in SHBG[71]. It has been estimated that ~20% of the increase in serum SHBG is mediated by a reduction in intrahepatic lipid content[71]. On the basis of hitherto presented GWAS data (Figure 6.1, panel B), it is anticipated that an intervention that specifically targets de novo lipogenesis, such as a low-carbohydrate diet, will be effective in increasing SHBG levels[73,74]. Of note, it is expected that any beneficial metabolic effect of a lifestyle intervention is explained not only by a direct effect of SHBG (as a hepatokine), but also by other pathways downstream from de novo lipogenesis, such as amelioration of hepatic lipid accumulation and insulin resistance (of which SHBG is a biomarker).

Second, it has been well-documented that oral oestrogens give rise to a drastic increase (~320%) in serum SHBG[75]. Of interest, such an effect has not been observed for transdermal oestrogens, which is likely explained by a much lower exposure of the liver[76,77], the primary site of SHBG synthesis[19]. Of additional interest, oral oestrogens, not transdermal oestrogens, have been associated with a 21% lower risk of T2D in postmenopausal women[76,78]. Notably, this risk reduction has not been observed in premenopausal women receiving (combined) oral contraceptives[79-81]. Furthermore, the potential adverse effects of oral oestrogens (e.g., venous thromboembolism and breast and endometrial cancers)[82-84], limit the applicability of oestrogens to prevent type 2 diabetes in women.

Finally, SHBG is also elevated in thyrotoxic states[85], which is explained by a lower rate of de novo lipogenesis[15]. Although treatment of an euthyroid individual with thyroid hormone will have too many adverse effects, thyroid hormone analogues with high affinity for the triiodothyronine receptor beta (THRβ) isoform, which is primarily expressed in the liver, may be beneficial. Initial studies with a first generation THRβ agonist, eprotirome, showed significant effects on SHBG levels (~271% increase), but

the drug was later retracted due to cartilage problems in dogs[86,87]. A next-generation THRβ agonist, resmetirom (MGL-3196), which thus far has shown no extrahepatic adverse effects, likewise has significant effects on SHBG levels (~116% increase)[88]. Both eprotirome and resmetirom also increased serum total testosterone, but not free testosterone levels, in men[86,88], in line with the genetic observations (Figure 6.2, panel E). Furthermore, treatment with eprotirome, which produced a more substantial increase in SHBG levels, also increased serum luteinizing hormone levels, which corroborates the role of the HPG axis (Figure 6.3, panel B). In women, total testosterone levels also increased upon resmetirom treatment[88], which deserves further study. Furthermore, additional studies are warranted to address the effect of resmetirom on glucose metabolism because the original study only reported glycemic outcome data for individuals without diabetes (resmetirom versus placebo change in hemoglobin A1c: -0.10%, 95% confidence interval: -0.23;0.03)[88].

Concluding remarks and future directions

A recent large-scale GWAS and Mendelian randomization study revealed that SHBG has an appreciably greater role in metabolic disorders than it has previously been given credit for, functioning as both a biomarker of metabolic derangements, including de novo lipogenesis, *and* a mediator (either primary or secondary) in the pathogenesis of metabolic disorders such as type 2 diabetes and PCOS.

Thorough study of genetic variants that affect serum SHBG levels suggests that de novo lipogenesis is one of the mechanistic links between the metabolic syndrome and SHBG levels. Furthermore, in women, SHBG not only acts as a carrier protein but also appears to be a true hepatokine involved in the pathogenesis of type 2 diabetes, independent of its effects on free testosterone. Together, this warrants the development of drugs that raise serum SHBG to treat and prevent type 2 diabetes and PCOS. Although trials are still in the early stages, THRβ agonists provide an interesting avenue of research (see Outstanding questions).

Glossary

De novo lipogenesis: the formation of fatty acids from non-lipid precursors, such as glucose.
Free hormone hypothesis: this hypothesis states that the biological effect of a hormone, such as testosterone, is the result of the unbound or free fraction rather than the total concentration of the hormone.
Hepatokine: a liver-derived, signaling protein that affects systemic metabolism.

Mendelian randomization: a statistical method that uses genetic variants as instruments to study the causal association between an exposure and outcome.

Outstanding questions

- What is the molecular mechanism that mediates the protective effect of SHBG on type 2 diabetes?
- Does (free) testosterone have a beneficial effect on type 2 diabetes risk, independent of SHBG, in men?
- How does SHBG affect the individual diagnostic criteria of PCOS?
- Do SHBG-raising therapies, e.g. inhibitors of de novo lipogenesis or thyroid hormone receptor beta agonists, reduce type 2 diabetes and PCOS risk?
- What is the mechanism responsible for the increase in total testosterone levels in resmetirom-treated women?

References

1. Mercier C, Alfsen A, Baulieu E. A testosterone binding globulin. Paper presented at: Proceedings of the second symposium on the steroid hormones. Ghent Belgium. Excerta Medica International Congress Series1966.
2. Dunn JF, Nisula BC, Rodbard D. Transport of steroid hormones: Binding of 21 endogenous steroids to both testosterone-binding globulin and corticosteroid-binding globulin in human plasma. *Journal of Clinical Endocrinology and Metabolism.* 1981;53(1):58-68.
3. Joseph DR. Structure, function, and regulation of androgen-binding protein/sex hormone-binding globulin. *Vitamins and Hormones.* 1994;49:197-280.
4. Mendel CM. The free hormone hypothesis: A physiologically based mathematical model. *Endocrine Reviews.* 1989;10(3):232-274.
5. Goldman AL, Bhasin S, Wu FCW, et al. A reappraisal of testosterone's binding in circulation: Physiological and clinical implications. *Endocrine Reviews.* 2017;38(4):302-324.
6. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nature Medicine.* 2020;26(2):252-258.
7. Grishkovskaya I, Avvakumov GV, Sklenar G, et al. Crystal structure of human sex hormone-binding globulin: Steroid transport by a laminin G-like domain. *EMBO Journal.* 2000;19(4):504-512.
8. Khan MS, Knowles BB, Aden DP, Rosne R. Secretion of testosterone-estradiol-binding globulin by a human hepatoma-derived cell line. *Journal of Clinical Endocrinology and Metabolism.* 1981;53(2): 448-
9. Avvakumov GV, Matveentseva IV, Akhrem LV, Strel'chyonok OA, Akhrem AA. Study of the carbohydrate moiety of human serum sex hormone-binding globulin. *Biochimica et Biophysica Acta (BBA) - General Subjects.* 1983;760(1):104-110.
10. Bocchinfuso WP, Warmels-Rodenhiser S, Hammond GL. Expression and differential glycosylation of human sex hormone-binding globulin by mammalian cell lines. *Molecular Endocrinology.* 1991;5(11):1723-1729.
11. Bérubé D, Séralini GE, Gagné R, Hammond GL. Localization of the human sex hormone-binding globulin gene (SHBG) to the short arm of chromosome 17 (17p12→p13). *Cytogenetic and Genome Research.* 1990;54:65-67.
12. Jänne M, Hammond GL. Hepatocyte nuclear factor-4 controls transcription from a TATA-less human sex hormone-binding globulin gene promoter. *Journal of Biological Chemistry.* 1998;273(51):34105-34114.
13. Saez-Lopez C, Brianso-Llort L, Torres-Torronteras J, et al. Resveratrol increases hepatic SHBG expression through human constitutive androstane receptor: A new contribution to the french paradox. *Scientific Reports.* 2017;7(1):12284.
14. Selva DM, Hammond GL. Peroxisome-proliferator receptor gamma represses hepatic sex hormone-binding globulin expression. *Endocrinology.* 2009;150(5):2183-2189.
15. Selva DM, Hammond GL. Thyroid hormones act indirectly to increase sex hormone-binding globulin production by liver via hepatocyte nuclear factor-4alpha. *Journal of Molecular Endocrinology.* 2009;43(1):19-27.
16. Simó R, Saez-Lopez C, Lecube A, et al. Adiponectin upregulates SHBG production: molecular mechanisms and potential implications. *Endocrinology.* 2014;155(8):2820-2830.
17. Simó R, Barbosa-Desongles A, Lecube A, Hernandez C, Selva DM. Potential role of tumor necrosis factor-α in downregulating sex hormone-binding globulin. *Diabetes.* 2012;61(2):372-382.
18. Simó R, Barbosa-Desongles A, Hernandez C, Selva DM. IL1β down-regulation of sex hormone-binding globulin production by decreasing HNF-4α via MEK-1/2 and JNK MAPK pathways. *Molecular Endocrinology.* 2012;26(11):1917-1927.
19. Simó R, Sáez-López C, Barbosa-Desongles A, Hernández C, Selva DM. Novel insights in SHBG regulation and clinical implications. *Trends in Endocrinology and Metabolism.* 2015;26(7):376-383.
20. Hammond GL. *Biology of Reproduction.* 2011;85(3):431-441.
21. Le TN, Nestler JE, Strauss JF, 3rd, Wickham EP, 3rd. Sex hormone-binding globulin and type 2 diabetes mellitus. *Trends in Endocrinology and Metabolism.* 2012;23(1):32-40.

22. Pugeat M, Nader N, Hogeveen K, et al. Sex hormone-binding globulin gene expression in the liver: drugs and the metabolic syndrome. *Mol Cell Endocrinol.* 2010;316(1):53-59.
23. Jayagopal V, Kilpatrick ES, Jennings PE, Hepburn DA, Atkin SL. The biological variation of testosterone and sex hormone-binding globulin (SHBG) in polycystic ovarian syndrome: Implications for SHBG as a surrogate marker of insulin resistance. *Journal of Clinical Endocrinology and Metabolism.* 2003;88(4):1528-1533.
24. Vermeulen A, Kaufman JM, Giagulli VA. Influence of some biological indexes on sex hormone-binding globulin and androgen levels in aging or obese males. *Journal of Clinical Endocrinology and Metabolism.* 1996;81(5):1821-1826.
25. Weinberg ME, Manson JE, Buring JE, et al. Low sex hormone-binding globulin is associated with the metabolic syndrome in postmenopausal women. *Metabolism: Clinical and Experimental.* 2006;55(11):1473-1480.
26. Laaksonen DE, Niskanen L, Punnonen K, et al. Testosterone and sex hormone–binding globulin predict the metabolic syndrome and diabetes in middle-aged men. *Diabetes Care.* 2004;27(5):1036-1041.
27. Moon H, Choi I, Kim S, et al. Cross-sectional association between testosterone, sex hormone-binding globulin and metabolic syndrome: The Healthy Twin Study. *Clin Endocrinol (Oxf).* 2017;87(5):523-531.
28. Plymate SR, Matej LA, Jones RE, Friedl KE. Inhibition of sex hormone-binding globulin production in the human hepatoma (Hep G2) cell line by insulin and prolactin. *Journal of Clinical Endocrinology and Metabolism.* 1988;67(3):460-464.
29. Nestler JE, Powers LP, Matt DW, et al. A Direct Effect of Hyperinsulinemia on Serum Sex Hormone-Binding Globulin Levels in Obese Women with the Polycystic Ovary Syndrome*. *Journal of Clinical Endocrinology and Metabolism.* 1991;72(1):83-89.
30. Preziosi P, Barrett-Connor E, Papoz L, et al. Interrelation between plasma sex hormone-binding globulin and plasma insulin in healthy adult women: the telecom study. *Journal of Clinical Endocrinology and Metabolism.* 1993;76(2):283-287.
31. Crave JC, Lejeune H, Brebant C, Baret C, Pugeat M. Differential effects of insulin and insulin-like growth factor I on the production of plasma steroid-binding globulins by human hepatoblastoma-derived (Hep G2) cells. *Journal of Clinical Endocrinology and Metabolism.* 1995;80(4):1283-1289.
32. Pasquali R, Casimirri F, De Iasio R, et al. Insulin regulates testosterone and sex hormone-binding globulin concentrations in adult normal weight and obese men. *Journal of Clinical Endocrinology and Metabolism.* 1995;80(2):654-658.
33. Selva DM, Hogeveen KN, Innis SM, Hammond GL. Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. *Journal of Clinical Investigation.* 2007;117(12):3979-3987.
34. Tong G, Hua X, Zhong Y, et al. Intensive insulin therapy increases sex hormone-binding globulin in newly diagnosed type 2 diabetic patients. *European Journal of Endocrinology.* 2014;170(2):237-245.
35. Peter A, Kantartzis K, Machann J, et al. Relationships of circulating sex hormone–binding globulin with metabolic traits in humans. *Diabetes.* 2010;59(12):3167-3173.
36. Donnelly KL, Smith CI, Schwarzenberg SJ, et al. Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *Journal of Clinical Investigation.* 2005;115(5):1343-1351.
37. Saez-Lopez C, Barbosa-Desongles A, Hernandez C, et al. Sex hormone-binding globulin reduction in metabolic disorders may play a role in NAFLD development. *Endocrinology.* 2017;158(3):545-559.
38. Miyake T, Kumagi T, Hirooka M, et al. Body mass index is the most useful predictive factor for the onset of nonalcoholic fatty liver disease: A community-based retrospective longitudinal cohort study. *Journal of Gastroenterology.* 2013;48(3):413-422.
39. Jones H, Sprung VS, Pugh CJ, et al. Polycystic ovary syndrome with hyperandrogenism is characterized by an increased risk of hepatic steatosis compared to nonhyperandrogenic PCOS phenotypes and healthy controls, independent of obesity and insulin resistance. *Journal of Clinical Endocrinology and Metabolism.* 2012;97(10):3709-3716.
40. Anstee QM, McPherson S, Day CP. How big a problem is non-alcoholic fatty liver disease? *BMJ.* 2011;343:d3897.

41. Liu C, Feng X, Li Q, et al. Adiponectin, TNF-α and inflammatory cytokines and risk of type 2 diabetes: A systematic review and meta-analysis. *Cytokine.* 2016;86:100-109.
42. Orio F, Jr., Palomba S, Cascella T, et al. Adiponectin Levels in Women with Polycystic Ovary Syndrome. *Journal of Clinical Endocrinology and Metabolism.* 2003;88(6):2619-2623.
43. Oróstica L, Astorga I, Plaza-Parrochia F, et al. Proinflammatory environment and role of TNF-α in endometrial function of obese women having polycystic ovarian syndrome. *International Journal of Obesity.* 2016;40(11):1715-1722.
44. Vendrell J, Broch M, Vilarrasa N, et al. Resistin, Adiponectin, Ghrelin, Leptin, and Proinflammatory Cytokines: Relationships in Obesity. *Obesity Research.* 2004;12(6):962-971.
45. Ding EL, Song Y, Manson JE, et al. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *New England Journal of Medicine.* 2009;361(12):1152-1163.
46. Perry JRB, Weedon MN, Langenberg C, et al. Genetic evidence that raised sex hormone binding globulin (SHBG) levels reduce the risk of type 2 diabetes. *Human Molecular Genetics.* 2010;19(3): 535-544.
47. Sáez-López C, Salcedo-Allende MT, Hernandez C, et al. Sex hormone-binding globulin expression correlates with acetyl-coenzyme a carboxylase and triglyceride content in human liver. *Journal of Clinical Endocrinology and Metabolism.* 2019;104(5):1500-1507.
48. Saez-Lopez C, Villena JA, Simó R, Selva DM. Sex hormone-binding globulin overexpression protects against high-fat diet-induced obesity in transgenic male mice. *The Journal of Nutritional Biochemistry.* 2020;85:108480.
49. Coviello AD, Haring R, Wellons M, et al. A genome-wide association meta-analysis of circulating sex hormone–binding globulin reveals multiple loci implicated in sex steroid hormone regulation. *Plos Genetics.* 2012;8(7):e1002805.
50. Ruth KS, Campbell PJ, Chew S, et al. Genome-wide association study with 1000 genomes imputation identifies signals for nine sex hormone-related phenotypes. *European Journal of Human Genetics.* 2016;24(2):284-290.
51. Beer NL, Tribble ND, McCulloch LJ, et al. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Human Molecular Genetics.* 2009;18(21):4081-4088.
52. Rees MG, Wincovitch S, Schultz J, et al. Cellular characterisation of the GCKR P446L variant associated with type 2 diabetes risk. *Diabetologia.* 2012;55(1):114-122.
53. Brouwers MCGJ, Jacobs C, Bast A, Stehouwer CDA, Schaper NC. Modulation of glucokinase regulatory protein: A double-edged sword? *Trends in Molecular Medicine.* 2015;21(10):583-594.
54. Santoro N, Caprio S, Pierpont B, et al. Hepatic de novo lipogenesis in obese youth is modulated by a common variant in the GCKR gene. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(8):E1125-E1132.
55. Orho-Melander M, Melander O, Guiducci C, et al. Common missense variant in the glucokinase regulatory protein gene is associated with increased plasma triglyceride and c-reactive protein but lower fasting glucose concentrations. *Diabetes.* 2008;57(11):3112-3121.
56. Kooner JS, Chambers JC, Aguilar-Salinas CA, et al. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nature Genetics.* 2008;40(2):149-151.
57. Kozlitina J, Smagris E, Stender S, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nature Genetics.* 2014;46(4):352-356.
58. Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nature Genetics.* 2008;40(12):1461-1465.
59. Pirazzi C, Adiels M, Burza MA, et al. Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro. *Journal of Hepatology.* 2012;57(6):1276-1282.
60. Mancina RM, Matikainen N, Maglio C, et al. Paradoxical dissociation between hepatic fat content and de novo lipogenesis due to PNPLA3 sequence variant. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(5):E821-E825.
61. Day FR, Hinds DA, Tung JY, et al. Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. *Nature Communications.* 2015;6:8464-8464.

62. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Human Molecular Genetics.* 2018;27(R2):R195-R208.
63. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nature Reviews: Genetics.* 2013;14(7):483-495.
64. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertility and Sterility.* 2004;81(1):19-25.
65. de Roux N, Genin E, Carel JC, et al. Hypogonadotropic hypogonadism due to loss of function of the KiSS1-derived peptide receptor GPR54. *Proceedings of the National Academy of Sciences of the United States of America.* 2003;100(19):10972-10976.
66. Dodé C, Levilliers J, Dupont JM, et al. Loss-of-function mutations in FGFR1 cause autosomal dominant kallmann syndrome. *Nature Genetics.* 2003;33(4):463-465.
67. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science.* 2007;316(5826):889-894.
68. Maffei M, Fei H, Lee GH, et al. Increased expression in adipocytes of ob RNA in mice with lesions of the hypothalamus and with mutations at the db locus. *Proceedings of the National Academy of Sciences of the United States of America.* 1995;92(15):6957-6960.
69. Strain GW, Zumoff B, Miller LK, et al. Effect of massive weight loss on hypothalamic pituitary-gonadal function in obese men. *Journal of Clinical Endocrinology and Metabolism.* 1988;66(5):1019-1023.
70. Niskanen L, Laaksonen DE, Punnonen K, et al. Changes in sex hormone-binding globulin and testosterone during weight loss and weight maintenance in abdominally obese men with the metabolic syndrome. *Diabetes, Obesity & Metabolism.* 2004;6(3):208-215.
71. Telgenkamp I, Kusters Y, Schalkwijk CG, et al. Contribution of liver fat to weight loss-induced changes in serum hepatokines: A randomized controlled trial. *Journal of Clinical Endocrinology and Metabolism.* 2019;104(7):2719-2727.
72. Sáez-López C, Soriguer F, Hernandez C, et al. Oleic acid increases hepatic sex hormone binding globulin production in men. *Molecular Nutrition & Food Research.* 2014;58(4):760-767.
73. Zhang X, Zheng Y, Guo Y, Lai Z. The effect of low carbohydrate diet on polycystic ovary syndrome: A meta-analysis of randomized controlled trials. *International Journal of Endocrinology.* 2019;2019:4386401-4386401.
74. Luukkonen PK, Sädevirta S, Zhou Y, et al. Saturated fat is more metabolically harmful for the human liver than unsaturated fat or simple sugars. *Diabetes Care.* 2018;41(8):1732-1739.
75. Kanaya AM, Herrington D, Vittinghoff E, et al. Glycemic effects of postmenopausal hormone therapy: The Heart and Estrogen/progestin Replacement Study. A randomized, double-blind, placebo-controlled trial. *Annals of Internal Medicine.* 2003;138(1):1-9.
76. Mauvais-Jarvis F, Manson JE, Stevenson JC, Fonseca VA. Menopausal hormone therapy and type 2 diabetes prevention: Evidence, mechanisms, and clinical implications. *Endocrine Reviews.* 2017;38(3):173-188.
77. Ropponen A, Aittomäki K, Vihma V, Tikkanen MJ, Ylikorkala O. Effects of oral and transdermal estradiol administration on levels of sex hormone-binding globulin in postmenopausal women with and without a history of intrahepatic cholestasis of pregnancy. *Journal of Clinical Endocrinology and Metabolism.* 2005;90(6):3431-3434.
78. Margolis KL, Bonds DE, Rodabough RJ, et al. Effect of oestrogen plus progestin on the incidence of diabetes in postmenopausal women: Results from the Women's Health Initiative Hormone Trial. *Diabetologia.* 2004;47(7):1175-1187.
79. Chasan-Taber L, Willett WC, Stampfer MJ, et al. A prospective study of oral contraceptives and NIDDM among U.S. women. *Diabetes Care.* 1997;20(3):330-335.
80. Lopez LM, Grimes DA, Schulz KF. Steroidal contraceptives: effect on carbohydrate metabolism in women without diabetes mellitus. *Cochrane Database of Systematic Reviews.* 2012(4):Cd006133.
81. Kim C, Siscovick DS, Sidney S, et al. Oral contraceptive use and association with glucose, insulin, and diabetes in young adult women: the CARDIA Study. Coronary Artery Risk Development in Young Adults. *Diabetes Care.* 2002;25(6):1027-1032.

82. Sweetland S, Beral V, Balkwill A, et al. Venous thromboembolism risk in relation to use of different types of postmenopausal hormone therapy in a large prospective study. *Journal of Thrombosis and Haemostasis.* 2012;10(11):2277-2286.
83. Beral VMWSC. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet.* 2003;362(9382):419-427.
84. Beral V, Bull D, Reeves G. Endometrial cancer and hormone-replacement therapy in the Million Women Study. *Lancet.* 2005;365(9470):1543-1551.
85. Vierhapper H, Bieglmayer C, Nowotny P, Waldhäusl W. Normal serum concentrations of sex hormone binding-globulin in patients with hyperthyroidism due to subacute thyroiditis. *Thyroid.* 1998;8(12):1107-1111.
86. Ladenson PW, Kristensen JD, Ridgway EC, et al. Use of the thyroid hormone analogue eprotirome in statin-treated dyslipidemia. *New England Journal of Medicine.* 2010;362(10):906-916.
87. Sweetlove M. Phase iii trial of eprotirome. *Pharmaceutical Medicine.* 2012;26(3):185-187.
88. Harrison SA, Bashir MR, Guy CD, et al. Resmetirom (MGL-3196) for the treatment of non-alcoholic steatohepatitis: A multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet.* 2019;394(10213):2012-2024.
89. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics.* 2014;23(R1):R89-R98.
90. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747-753.

PART III

THE ASSOCIATIONS BETWEEN DE NOVO LIPOGENESIS AND CARDIOMETABOLIC DISEASE, AND THE ROLE OF SHBG HEREIN

7

CHAPTER SEVEN

Serum sex hormone-binding globulin mediates the association between intrahepatic lipid content and type 2 diabetes: The Maastricht Study

Pomme I.H.G. Simons, Olivier Valkenburg, Marjo P.H. van de Waarenburg, Marleen M.J. van Greevenbroek, M. Eline Kooi, Jacobus F.A. Jansen, Casper G. Schalkwijk, Coen D.A. Stehouwer, Martijn C.G.J. Brouwers

Submitted

Summary

Objective: Serum sex hormone-binding globulin (SHBG) has been proposed as a hepatokine that contributes to the extrahepatic complications observed in non-alcoholic fatty liver disease (NAFLD). However, it remains uncertain whether serum SHBG contributes to the association between intrahepatic lipids (IHL) and type 2 diabetes. Therefore, we studied whether, and to what extent, serum SHBG mediates the association between IHL content and type 2 diabetes.

Methods: We used cross-sectional data from The Maastricht Study (n=1,554), a population based cohort study with oversampling of individuals with type 2 diabetes. Type 2 diabetes status was assessed by oral glucose tolerance test and IHL content was measured with 3T-Dixon MRI. Mediation analyses were conducted to assess the contribution of serum SHBG to the association between IHL content and type 2 diabetes status.

Results: IHL content was significantly associated with type 2 diabetes in women and men (OR: 1.07, 95% CI: 1.04;1.10 and OR: 1.12, 95% CI: 1.09;1.15, respectively). Serum SHBG significantly mediated the association between IHL content and type 2 diabetes. The contribution of serum SHBG was higher in women 50.9% (95% CI: 26.7;81.3) than in men 17.2% (95% CI: 9.6;27.6). Repeat analyses with proxies of type 2 diabetes and adjustment for covariates did not substantially affect the results.

Discussion: In this large-scale population-based cohort study, serum SHBG mediated the association between IHL content and type 2 diabetes. These findings extend our understanding of the mechanism by which NAFLD contributes to type 2 diabetes and further elaborate on SHBG as a hepatokine.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is a serious health concern that affects approximately 25% of the global population[1]. It encompasses a spectrum of histological abnormalities that result from an excess storage of intrahepatic lipids (IHL)[2]. NAFLD is a precursor for several hepatic complications including liver failure and hepatocellular carcinoma, but is also a risk factor for the development of various extrahepatic complications, such as type 2 diabetes and cardiovascular disease[3-5].

It has been proposed that hepatokines – i.e. liver-derived proteins that have systemic metabolic effects – may, in part, mediate the association between IHL accumulation and extrahepatic complications[6,7]. Serum sex hormone-binding globulin (SHBG) has emerged as a novel hepatokine[8,9]. We have previously extended experimental data[10] by showing that de novo lipogenesis, the principal pathway resulting in the accumulation of IHL[11], is inversely associated with serum SHBG levels in humans[12]. Furthermore, low serum SHBG levels have been causally linked to a higher risk of type 2 diabetes[13,14]. Of note, this effect appears attributable to SHBG itself, independent from the effects of SHBG on free testosterone levels[15].

Thus far, however, it remains to be elucidated to what extent serum SHBG mediates the association between IHL accumulation and type 2 diabetes. Therefore, the aim of the present study was to assess whether, and to what extent, serum SHBG contributes to the association between IHL content and type 2 diabetes, in a population-based cohort.

Methods

Study population

The Maastricht Study is a population-based cohort study with an oversampling of individuals with type 2 diabetes. The study design and rationale have been extensively described previously[16]. In brief, the Maastricht Study focuses on the aetiology, pathophysiology, complications, and comorbidities of type 2 diabetes and is characterized by extensive phenotyping of all participants. All individuals between 40 and 75 years of age living in the southern part of The Netherlands were eligible for participation. Participants were recruited through mass media campaigns and via mailings from the municipal registries and the regional Diabetes Patient Registry.

The present study includes cross-sectional data from 3,340 participants in whom serum SHBG levels were measured and who completed baseline measurements from November

2010 until December 2017. Quantification of IHL content was implemented from December 2013 onwards. Participants were excluded from the current analyses if they were diagnosed with other types of diabetes (n=41), were missing data on intrahepatic lipid content (n=1,161) or covariates (n=584). This resulted in a study population of 1,554 participants (Supplementary Figure S7.1) of whom 369 had type 2 diabetes.

The Maastricht Study has been approved by the institutional medical ethical committee (NL31329.068.10) and the Minister of Health, Welfare and Sports of the Netherlands (Permit 131088-105234-PG). All participants gave written informed consent prior to participation.

Outcome: type 2 diabetes

All participants underwent a standardized 7-point 75g oral glucose tolerance test after an overnight fast, except for individuals using insulin and/or individuals with a fasting capillary glucose ≥11.1 mmol/l. These individuals were automatically classified as having diabetes. Diabetes was defined according to the World Health Organization 2006 diagnostic criteria as a fasting plasma glucose ≥7.0 mmol/l and/or a 2-hour plasma glucose ≥11.1 mmol/l[17]. Finally, participants who used glucose-lowering medication, and were not diagnosed with other types of diabetes, were also defined as having type 2 diabetes[16].

Exposure: intrahepatic lipid content

IHL content was quantified by Dixon MR imaging using a 3.0 Tesla MRI system (MAGNETOM Prismafit, Siemens Healthineers, Erlangen, Germany) with body matrix and supine radiofrequency coils. After a scout scan, transversal two-dimensional T2-weighted True Fast Imaging with Steady-State Free Precession (T2w TRUFI) images were obtained of the liver with the following parameters: voxel size: 1.2 x 1.2 x 5.0 mm^3, repetition time (TR): 422 ms, echo time (TE): 1.65 ms, flip angle: 60°, number of signal averages: 1, parallel imaging (GRAPPA) factor: 2. Next, transversal two-dimensional turbo spin echo Dixon MR images were obtained of the liver during a breathhold using the following parameters: voxel size: 2.0×2.0×6.0 mm^3, number of slices: 4, TR: 500 ms, TE: 31 ms, turbo factor: 5, number of signal averages: 1, parallel imaging (GRAPPA) factor: 3[18]. Three regions-of-interest (ROIs) were drawn in the liver by trained observers on the T2w TRUFI images, while taking care to avoid positioning the ROIs on visible structures, such as vessels and bile ducts, and positioning the ROIs in artifact-free regions. Subsequently, these ROIs were copied to the water and fat Dixon MR images to calculate the intrahepatic lipid percentage expressed as the ratio CH_2/H_20 *100%.

The Dixon MRI method was validated in 36 participants with a broad range of IHL content, and calibrated against 3T proton magnetic resonance spectroscopy (^{1}H-MRS), i.e.

the gold standard for non-invasively quantifying IHL[19]. After calibration, the intra-class correlation coefficient between Dixon MRI and ^{1}H-MRS was 0.989 (95% CI: 0.979; 0.994).

Mediator: serum SHBG levels

Serum SHBG levels were measured using a human SHBG DuoSet solid phase sandwich enzyme-linked immune sorbent assay (ELISA; R&D systems, Minneapolis, MN, USA) according to the manufacturer's instructions. The intra- and interassay coefficients of variation for serum SHBG were 2.8% and 5.1%, respectively. The DuoSet ELISA was validated against chemiluminescent immunometric assay (Immulite XPi instrument, Siemens Healthcare Diagnostics, Mannheim, Germany) in eight samples. The intraclass correlation coefficient was 0.974 (95% CI: 0.862; 0.995).

Measurement of covariates

All participants completed questionnaires regarding age, sex, educational level (low, medium, or high), smoking status (never, former, or current smoker), use of alcohol (grams/day) and menopausal status (postmenopausal status was defined as a most recent menstrual period more than 12 months prior to the time of assessment)[16]. Use of medication was assessed through medication interviews. Anthropometric measurements including weight, height, waist circumference, and office systolic and diastolic blood pressure were measured during physical examination. Body mass index (BMI) was calculated as weight (kg) divided by height (m) squared[16]. Daily total physical activity levels were measured during eight consecutive days using the activPAL3 physical activity monitor (PAL Technologies, Glasgow, UK) and expressed in minutes of stepping activity per day[20]. Fasting levels of glucose, insulin, glycated haemoglobin A1c (HbA1c), and lipid profile (total cholesterol, HDL-cholesterol, LDL-cholesterol and triglycerides) were measured in venous blood samples[16]. Insulin sensitivity was approximated by the Matsuda insulin sensitivity index which was calculated as described previously[21]. Adherence to the Dutch dietary guidelines was assessed based on the Dutch Healthy Diet (DHD) index based on food frequency questionnaires[22]. In the Maastricht Study, the DHD index consists of 14 components. The coffee component was not included as it is based on the type of coffee consumed, which the food frequency questionnaires were unable to distinguish between[23]. Furthermore, as we included alcohol consumption as a separate covariate in the regression models (see below), the DHD index in this study was reported as the DHD-13 (DHD-14 minus the alcohol component).

Statistical analyses

Continuous data are presented as mean ± standard deviation, or as median (interquartile range) in case of non-normal distribution. Categorical data are presented

in percentages. Non-normally distributed variables were ^{10}log-transformed prior to further analyses.

First, univariate regression analyses were conducted to study the associations between IHL content, serum SHBG and type 2 diabetes. All associations were explored for an interaction with sex. Regression coefficients are presented as unstandardized betas.

Second, mediation analyses were conducted to investigate whether the association between IHL content and type 2 diabetes status was mediated by serum SHBG. For serum SHBG to be considered as a potential mediator, the following assumptions had to be met: 1) a statistically significant association between the exposure (IHL content) and mediator (serum SHBG); 2) a statistically significant association between the mediator and outcome (type 2 diabetes status); 3) a statistically significant association between the exposure and outcome; 4) a decrease in the strength of the association between the exposure and outcome upon adjustment for the mediator[24]. Mediation analyses were adjusted for the following confounders: model 1 was adjusted for age; model 2 was additionally adjusted for (proxies of) lifestyle: BMI, alcohol intake, DHD-13, level of education, and total physical activity. In women, we additionally adjusted for menopausal status and use of oestrogen-containing medication (model 3). Furthermore, given the oversampling of type 2 diabetes, analyses were repeated after taking the higher prevalence of type 2 diabetes into consideration using case-control mediation analyses[25]. In addition, the analyses were repeated with adjustment for waist circumference instead of BMI (model 2). Lastly, analyses were repeated with proxies of type 2 diabetes as the outcome variable, i.e. HbA1c and Matsuda index.

For all mediation analyses, the 95% confidence intervals were estimated using nonparametric bootstrapping with the percentile method. The proportion of mediation was estimated as $OR^{Direct}*(OR^{Indirect} - 1)/(OR^{Direct} * OR^{Indirect} - 1) * 100$ in case of a binary outcome[26], or as $\beta^{Indirect}/\beta^{Total} * 100\%$ in case of a continuous outcome[27]. All results were considered statistically significant at p-value <0.05, except for interaction terms where a less stringent p-value threshold was considered statistically significant (p<0.10).

Statistical analyses were performed using IBM Statistical Package of Social Science (SPSS) version 27.0 for Windows (IBM Corp., Chicago, IL, USA) and R statistical software version 4.0.1 (R foundation for Statistical Computing) with the *CMAverse* package[28].

Results

Study population

Table 7.1 shows the characteristics of the overall study population and stratified according to type 2 diabetes status. The overall population had a mean age of 60 ± 8 years, 47.9% was female of whom the majority (79.2%) were postmenopausal. Only a small number of women (4.0%) used oestrogen-containing medication. The overall population had a median IHL content of 3.5% (IQR: 2.1-6.5) and a median serum SHBG level of 35.5 nmol/l (IQR: 25.3-49.8). Participants with type 2 diabetes were more often male, tended to be older, and generally had a poorer metabolic profile (i.e. higher BMI, waist circumference, systolic blood pressure, and serum triglycerides, and lower HDL-cholesterol and Matsuda index). Furthermore, participants with type 2 diabetes were characterized by a higher median IHL content and lower serum SHBG levels.

Table 7.1 Characteristics of the overall study population and stratified according to type 2 diabetes status.

	Overall (n=1,554)	Individuals without type 2 diabetes (n=1,185)	Individuals with type 2 diabetes (n=369)
Age, years	60 ± 8	59 ± 8	62 ± 8
Sex, % women	47.9	53.8	29.0
Postmenopausal, % of women	79.2	78.3	84.1
Use of oestrogen-containing medication, % of women	4.0	3.8	5.6
Education level low/medium/high, %	30.3 / 28.8 / 40.9	27.3 / 28.4 / 44.2	39.8 / 30.1 / 30.1
Smoking, never/former/current, %	38.0 / 50.2 / 11.8	39.0 / 49.3 / 11.7	34.8 / 53.0 / 12.2
Dutch healthy diet index (DHD-13*)	77.2 ± 13.9	78.0 ± 13.9	74.8 ± 13.7
Alcohol, g/day	9.0 (2.0-19.0)	9.8 (2.7-19.5)	5.8 (0.5-16.0)
Physical activity, min/day	120.8 (93.6-148.7)	125.5 (101.1-152.5)	100.9 (74.6-135.8)
BMI, kg/m^2	26.6 ± 4.1	25.8 ± 3.7	29.0 ± 4.3
Waist circumference, cm	94.3 ± 12.5	91.3 ± 11.0	103.8 ± 12.5
Office systolic blood pressure, mmHg	134 ± 17	132 ± 17	140 ± 16
Office diastolic blood pressure, mmHg	76 ± 10	76 ± 10	77 ± 9
Total cholesterol, mmol/l	5.3 ± 1.2	5.6 ± 1.1	4.5 ± 1.0
HDL cholesterol, mmol/l	1.6 ± 0.5	1.7 ± 0.5	1.3 ± 0.4
LDL cholesterol, mmol/l	3.1 ± 1.0	3.3 ± 1.0	2.4 ± 0.9
Triglycerides, mmol/l	1.2 (0.9-1.7)	1.1 (0.8-1.5)	1.5 (1.1-2.1)
Use of lipid-modifying medication, %	32.3	19.7	72.9
HbA1c, %	5.6 (5.4-6.0)	5.4 (5.3-5.7)	6.7 (6.2-7.4)
HbA1c, mmol/mol	38.0 (35.0-42.0)	36.0 (34.0-39.0)	50.0 (44.5-57.0)
Fasting glucose, mmol/l	5.5 (5.0-6.3)	5.3 (4.9-5.7)	7.5 (6.8-8.6)
Fasting insulin, mmol/l	59.1 (41.8-87.8)	55.3 (40.0-77.0)	83.2 (51.3-126.0)
Matsuda-index	3.6 (2.1-5.3)	4.1 (2.6-5.8)	2.0 (1.3-3.0)
Use of glucose-lowering medication, %	17.9	0.0	75.3
Intrahepatic lipid content, %	3.5 (2.1-6.5)	2.9 (1.9-5.1)	5.2 (3.5-10.7)
Serum SHBG, nmol/l	35.5 (25.3-49.8)	38.7 (27.6-54.2)	26.4 (19.6-37.0)

* Dutch Healthy Diet (DHD-14) index minus alcohol component (=DHD-13). Data are presented as mean ± standard deviation, median (interquartile range) unless otherwise indicated.
Abbreviations: BMI body mass index; HDL high-density lipoprotein; LDL low-density lipoprotein; HbA1c haemoglobin A1c; SHBG sex hormone-binding globulin

Univariate regression analyses

Univariate regression analyses were conducted to assess whether the first three assumptions of mediation were met, i.e. 1) the association between IHL content and serum SHBG levels, 2) the association between IHL content and type 2 diabetes status, and 3) the association between serum SHBG levels and type 2 diabetes status. Since there was a statistically significant interaction between the dependent variable and sex on the outcome for all three associations (p=0.001, p=0.040, and p<0.001, respectively), all analyses were subsequently stratified according to sex. The sex-stratified population characteristics are presented in Supplementary Table S7.1.

There was a statistically significant inverse association between IHL content and serum SHBG levels in men and women (β: -0.008, 95% CI: -0.010;-0.006 and β: -0.014, 95% CI: -0.017;-0.011, respectively; Figure 7.1). The strength of association was stronger in women, which remained after exclusion of premenopausal women and women using oestrogen-containing medication (β: -0.015, 95% CI: -0.018;-0.012). Furthermore, there was a statistically significant association between IHL content and type 2 diabetes status in men and women (OR: 1.12, 95% CI: 1.09;1.15 and OR: 1.07, 95% CI: 1.04;1.10, respectively) and a statistically significant inverse association between serum SHBG and type 2 diabetes status in men and women (OR: 0.96, 95% CI: 0.95;0.98 and OR: 0.98, 95% CI: 0.97;0.99, respectively). Adjustment for age, BMI, alcohol intake, DHD-13, level of education, physical activity, and, in women, menopausal status and use of oestrogen-containing medication, did not substantially affect the results (p<0.025 for all analyses; data not shown).

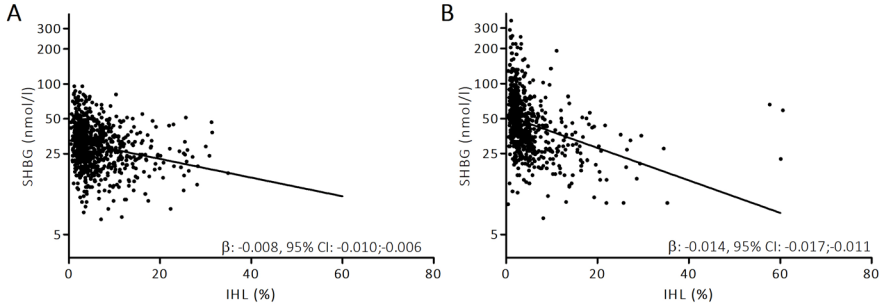

Figure 7.1 Association between intrahepatic lipid (IHL) content and serum sex hormone-binding globulin (SHBG) in men (A) and women (B). The black line represents the line of best fit.

Mediation analyses

Since the first three requirements for mediation were met, we subsequently conducted mediation analyses to assess whether, and to what extent, the relationship between IHL content and type 2 diabetes status was mediated by serum SHBG levels. We found that the association between IHL content and type 2 diabetes was statistically significantly mediated by serum SHBG in men and women (Figure 7.2). In men, serum SHBG was estimated to mediate 17.2% (95% CI: 9.6;27.6) of the association between IHL content and type 2 diabetes, while the proportion of mediation was 50.9% (95% CI: 26.7;81.3) in women. The mediation effect of serum SHBG remained statistically significant after adjustment for age (model 1), BMI, alcohol intake, DHD-13, level of education, and total physical activity (model 2), and, in women, menopausal status and use of oestrogen-containing medication (model 3) (Table 7.2).

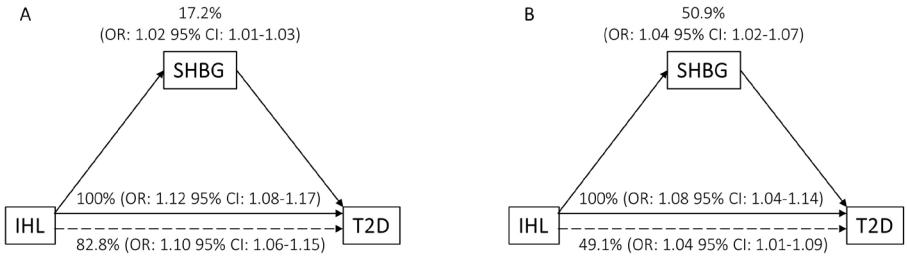

Figure 7.2 Crude association between intrahepatic lipid (IHL) content and type 2 diabetes (T2D) mediated by serum sex hormone-binding globulin (SHBG) in men (A) and women (B). Horizontal, solid arrows represent the total effect, i.e. the association between IHL content and T2D. Horizontal, dashed arrows indicate the direct effect, i.e. the association between IHL content and T2D status not attributable to serum SHBG.

Additional analyses

The mediation analyses were repeated after accounting for the enrichment of type 2 diabetes in the Maastricht Study, which did not substantially affect the results (Supplementary Table S7.2). In addition, the mediation analyses were repeated after adjustment for waist circumference instead of BMI (model 2) which did not materially change the results (Supplementary Table S7.3). Finally, mediation analyses were repeated with proxies for type 2 diabetes (i.e. HbA1c and Matsuda-index) as the dependent variable. All assumptions for mediation were met. In both men and women, serum SHBG remained a statistically significant mediator in the association between IHL content and both HbA1c and Matsuda-index in the crude and fully adjusted models (Supplementary Table S7.4 and S7.5).

Table 7.2 Mediation effect of serum SHBG on the association between intrahepatic lipid content and type 2 diabetes.

	Men (n=810)		Women (n=744)	
	OR (95% CI)	% mediated (95% CI)±	OR (95% CI)	% mediated (95% CI)±
Crude				
Total effect*	1.12 (1.08;1.17)		1.08 (1.04;1.14)	
Direct effect*	1.10 (1.06;1.15)		1.04 (1.01;1.09)	
Indirect effect*	1.02 (1.01;1.03)	17.2 (9.6;27.6)	1.04 (1.02;1.07)	50.9 (26.7;81.3)
Model 1				
Total effect	1.13 (1.09;1.17)		1.08 (1.05;1.13)	
Direct effect	1.10 (1.06;1.14)		1.04 (1.01;1.09)	
Indirect effect	1.03 (1.02;1.04)	24.6 (15.6;36.0)	1.04 (1.02;1.07)	48.5 (24.0;80.4)
Model 2				
Total effect	1.08 (1.05;1.13)		1.04 (1.00;1.09)	
Direct effect	1.07 (1.03;1.11)		1.02 (0.98;1.07)	
Indirect effect	1.01 (1.01;1.02)	17.7 (8.3;32.8)	1.02 (1.00;1.04)	42.6 (2.5;254.1)
Model 3				
Total effect			1.04 (1.00;1.10)	
Direct effect			1.02 (0.98;1.07)	
Indirect effect			1.02 (1.01;1.04)	55.9 (-72.38;337.2)

± % mediated is calculated as $OR^{Direct} * (OR^{Indirect} - 1)/(OR^{Direct} * OR^{Indirect} - 1) * 100$. * Total effect represents association between intrahepatic lipid (IHL) content and type 2 diabetes (T2D); direct effect represents association between IHL and T2D status not attributable to serum SHBG; indirect effect represents association between IHL and T2D attributable to serum SHBG (= mediation).
Model 1: adjusted for age.
Model 2: additionally adjusted for BMI, alcohol intake, Dutch Healthy Diet index (DHD-13), level of education, and total physical activity.
Model 3: additionally adjusted for menopausal status and use of oestrogen-containing medication.

Discussion

In the present study, serum SHBG partially mediated the association between the IHL content and type 2 diabetes status. The contribution of serum SHBG to the association between IHL content and type 2 diabetes was higher in women than in men. Similar results were found when the analyses were repeated with proxies of type 2 diabetes (i.e. HbA1c and Matsuda index) and when adjusted for confounders.

The importance of hepatokines in the pathogenesis of extrahepatic disease, in particular type 2 diabetes, is increasingly recognized[6,7,29,30]. Nevertheless, this is the first study that has assessed the mediation effect of serum SHBG in the association between IHL content and type 2 diabetes. The current findings corroborate the hypothesis that SHBG may have a role not only as carrier protein for testosterone and a biomarker of metabolic disease, but also as a hepatokine affecting type 2 diabetes[14,15,31]. Experimental studies have shown that de novo lipogenesis – which is

one of the primary pathways contributing to the accumulation of IHL[11] – downregulates hepatocyte nuclear factor 4 alpha and subsequently serum SHBG levels[10], a finding which we recently extrapolated to humans[12]. In turn, lower serum SHBG levels have been found to be causally associated with an increased risk of type 2 diabetes[13,14]. The exact biological mechanism by which serum SHBG influences type 2 diabetes is poorly understood, however, it has been proposed that the effect is independent of free testosterone, and, therefore, attributable to SHBG itself[15,32].

It is likely that there are several pathways that mediate the association between IHL content and type 2 diabetes, of which serum SHBG is merely one. Insulin resistance and excess (hepatic) glucose production are other well-known mediators[33]. The accumulation of IHL contributes to an excess of circulating fatty acid metabolites in peripheral tissues which are involved in the pathogenesis of insulin resistance[7,33,34]. In addition, the carbohydrate regulatory element binding protein (ChREBP) – which is one of the principal transcription factors that regulate de novo lipogenesis[35] – activates glucose-6-phosphatase and thereby contributes to an increased hepatic glucose production[36]. It is likely that these pathways are largely responsible for the remaining direct effect of IHL content on type 2 diabetes that was observed in this study.

We observed a relatively high estimated proportion of mediation by serum SHBG in the association between IHL content and type 2 diabetes. This could be an indication of the biological relevance of serum SHBG in the pathogenesis of type 2 diabetes. However, in some analyses there was a considerable uncertainty in the estimated proportion of mediation, in particular when the direct effect (i.e. the effect of IHL content on type 2 diabetes not attributable to serum SHBG) had lost statistical significance. These methodologically implausible intervals – the estimated proportion cannot be negative or exceed 100% – are, therefore, likely due to a statistical limitation rather than an indication of the true uncertainty of the estimate[37]. Alternatively, the high estimated proportion of mediation may partially be the result of potential bidirectionality of the associations. Although it is assumed that there is a causal association between IHL content, serum SHBG and type 2 diabetes – an assumption that is supported by experimental and genetic studies[10,13,14] – it cannot be excluded that, in fact, the associations are bidirectional. Indeed, experimental studies have found that transgenic *SHBG* mice show reduced IHL content and improved glucose homeostasis[38-40], although these findings have not yet been extrapolated to humans. In addition, hyperinsulinemia and hyperglycaemia, which are characteristic for type 2 diabetes[41], stimulate de novo lipogenesis and, consequently, IHL accumulation[42]. As a result of the bidirectionality of these associations, the currently observed estimates should be regarded as the maximum mediation effects.

There was a striking difference in the mediation effect of serum SHBG between men and women, with a higher contribution observed in women. Of interest, previous observational studies have reported similar sexually dimorphic associations between IHL content and serum SHBG[8], in line with the results of the univariate regression analyses in this study[9,13]. Moreover, genetic studies have reported that variants in *GCKR,* which are associated with higher rates of de novo lipogenesis and IHL content[43], have a stronger, inverse association with serum SHBG in women than in men[44]. Nevertheless, the biological mechanisms that account for these sex differences remain poorly understood, and deserve further investigation.

This study has several strengths. By using data from the Maastricht Study we were able to obtain a large cohort of individuals with oversampling of type 2 diabetes, diagnosed by an oral glucose tolerance test. The extensive phenotyping allowed for adjustment for many well-defined confounders, such as use of oestrogen-containing medication and physical activity, assessed by an accelerometer. Furthermore, IHL content was quantified using state-of-the-art methodology (i.e. Dixon MRI). This study also had several limitations. First, as a result of the cross-sectional nature of the data, we cannot draw conclusions on causality. Although experimental and genetic studies support the assumptions of causality in this study, it cannot be excluded that, as hitherto mentioned, the associations are bidirectional. Second, the participants in the current study were primarily Caucasian aged between 40 and 75 years, which resulted in a relatively low number of premenopausal women. Extrapolation to other groups should, therefore, be done with care.

In conclusion, in a large-scale population-based cohort study, we show that serum SHBG mediates the association between IHL content and type 2 diabetes. The mediation effect was larger in women. These findings extend our knowledge on the mechanism that links NAFLD to type 2 diabetes and emphasizes the importance of serum SHBG as a hepatokine.

References

1. Younossi ZM. Non-alcoholic fatty liver disease - A global public health perspective. *Journal of Hepatology.* 2019;70(3):531-544.
2. Ipsen DH, Lykkesfeldt J, Tveden-Nyborg P. Molecular mechanisms of hepatic lipid accumulation in non-alcoholic fatty liver disease. *Cellular and Molecular Life Sciences.* 2018;75(18):3313-3327.
3. Targher G, Corey KE, Byrne CD, Roden M. The complex link between NAFLD and type 2 diabetes mellitus — mechanisms and treatments. *Nature Reviews Gastroenterology & Hepatology.* 2021;18(9):599-612.
4. Mantovani A, Petracca G, Beatrice G, et al. Non-alcoholic fatty liver disease and risk of incident diabetes mellitus: an updated meta-analysis of 501 022 adult individuals. *Gut.* 2021;70(5):962-969.
5. Brouwers MCGJ, Simons N, Stehouwer CDA, Isaacs A. Non-alcoholic fatty liver disease and cardiovascular disease: assessing the evidence for causality. *Diabetologia.* 2020;63(2):253-260.
6. Iroz A, Couty J-P, Postic C. Hepatokines: Unlocking the multi-organ network in metabolic diseases. *Diabetologia.* 2015;58(8):1699-1703.
7. Stefan N, Häring HU. The role of hepatokines in metabolism. *Nature Reviews: Endocrinology.* 2013;9(3):144-152.
8. Jaruvongvanich V, Sanguankeo A, Riangwiwat T, Upala S. Testosterone, sex hormone-binding globulin and nonalcoholic fatty liver disease: A systematic review and meta-analysis. *Annals of Hepatology.* 2017;16(3):382-394.
9. Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: A systematic review and meta-analysis. *JAMA.* 2006;295(11):1288-1299.
10. Selva DM, Hogeveen KN, Innis SM, Hammond GL. Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. *Journal of Clinical Investigation.* 2007;117(12):3979-3987.
11. Donnelly KL, Smith CI, Schwarzenberg SJ, et al. Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *Journal of Clinical Investigation.* 2005;115(5):1343-1351.
12. Simons PIHG, Valkenburg O, Telgenkamp I, et al. Relationship between de novo lipogenesis and serum sex hormone binding globulin in humans. *Clinical Endocrinology.* 2021;95(1):101-106.
13. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nature Medicine.* 2020;26(2):252-258.
14. Ding EL, Song Y, Manson JE, et al. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *New England Journal of Medicine.* 2009;361(12):1152-1163.
15. Simons PIHG, Valkenburg O, Stehouwer CDA, Brouwers MCGJ. Sex hormone-binding globulin: biomarker and hepatokine? *Trends in Endocrinology and Metabolism.* 2021;32(8):544-553.
16. Schram MT, Sep SJ, van der Kallen CJ, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. *European Journal of Epidemiology.* 2014;29(6):439-451.
17. Organization WH. *efinition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycaemia.* Geneva, Switzerland: World Health Organization;2006.
18. WT D. Simple proton spectroscopic imaging. *Radiology.* 1984;153(1):189-194.
19. van Werven JR, Marsman HA, Nederveen AJ, et al. Assessment of hepatic steatosis in patients undergoing liver resection: comparison of US, CT, T1-weighted dual-echo MR imaging, and point-resolved 1H MR spectroscopy. *Radiology.* 2010;256(1):159-168.
20. Vandercappellen EJ, Henry RMA, Savelberg H, et al. Association of the Amount and Pattern of Physical Activity With Arterial Stiffness: The Maastricht Study. *J Am Heart Assoc.* 2020;9(20):e017502.
21. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. *Diabetes Care.* 1999;22(9):1462-1470.
22. van Lee L, Geelen A, van Huysduynen EJCH, et al. The Dutch Healthy Diet index (DHD-index): an instrument to measure adherence to the Dutch Guidelines for a Healthy Diet. *Nutrition Journal.* 2012;11(1):49.

23. Maasen K, Eussen SJPM, Scheijen JLJM, et al. Higher habitual intake of dietary dicarbonyls is associated with higher corresponding plasma dicarbonyl concentrations and skin autofluorescence: the Maastricht Study. *The American Journal of Clinical Nutrition.* 2021.
24. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods.* 2013;18(2):137-150.
25. VanderWeele TJ, Tchetgen Tchetgen EJ. Mediation Analysis With Matched Case-Control Study Designs. *American Journal of Epidemiology.* 2016;183(9):869-870.
26. VanderWeele TJ. Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health.* 2016;37(1):17-32.
27. VanderWeele TJ. Policy-relevant proportions for direct effects. *Epidemiology (Cambridge, Mass).* 2013;24(1):175-176.
28. Shi B, Choirat C, Coull BA, VanderWeele TJ, Valeri L. CMAverse: A Suite of Functions for Reproducible Causal Mediation Analyses. *Epidemiology.* 2021;32(5):e20-e22.
29. Jensen-Cody SO, Potthoff MJ. Hepatokines and metabolism: Deciphering communication from the liver. *Molecular Metabolism.* 2021;44:101138.
30. Misu H. Identification of hepatokines involved in pathology of type 2 diabetes and obesity. *Endocrine Journal.* 2019;66(8):659-662.
31. Meex RCR, Watt MJ. Hepatokines: linking nonalcoholic fatty liver disease and insulin resistance. *Nature Reviews Endocrinology.* 2017;13(9):509-520.
32. Le TN, Nestler JE, Strauss JF, 3rd, Wickham EP, 3rd. Sex hormone-binding globulin and type 2 diabetes mellitus. *Trends in Endocrinology and Metabolism.* 2012;23(1):32-40.
33. Targher G, Byrne CD. Nonalcoholic fatty liver disease: A novel cardiometabolic risk factor for type 2 diabetes and its complications. *Journal of Clinical Endocrinology and Metabolism.* 2013;98(2):483-495.
34. Kitade H, Chen G, Ni Y, Ota T. Nonalcoholic Fatty Liver Disease and Insulin Resistance: New Insights and Potential New Treatments. *Nutrients.* 2017;9(4):387.
35. Strable MS, Ntambi JM. Genetic control of de novo lipogenesis: role in diet-induced obesity. *Critical reviews in biochemistry and molecular biology.* 2010;45(3):199-214.
36. Kim M-S, Krawczyk SA, Doridot L, et al. ChREBP regulates fructose-induced glucose production independently of insulin signaling. *Journal of Clinical Investigation.* 2019;126(11):4372-4386.
37. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annual Review of Psychology.* 2007;58: 593-614.
38. Saez-Lopez C, Barbosa-Desongles A, Hernandez C, et al. Sex hormone-binding globulin reduction in metabolic disorders may play a role in NAFLD development. *Endocrinology.* 2017;158(3):545-559.
39. Sáez-López C, Salcedo-Allende MT, Hernandez C, et al. Sex hormone-binding globulin expression correlates with acetyl-coenzyme a carboxylase and triglyceride content in human liver. *Journal of Clinical Endocrinology and Metabolism.* 2019;104(5):1500-1507.
40. Saez-Lopez C, Villena JA, Simó R, Selva DM. Sex hormone-binding globulin overexpression protects against high-fat diet-induced obesity in transgenic male mice. *The Journal of Nutritional Biochemistry.* 2020;85:108480.
41. DeFronzo RA, Ferrannini E, Groop L, et al. Type 2 diabetes mellitus. *Nature Reviews Disease Primers.* 2015;1(1):15019.
42. Sanders FWB, Griffin JL. De novo lipogenesis in the liver in health and disease: more than just a shunting yard for glucose. *Biological Reviews.* 2016;91(2):452-468.
43. Santoro N, Caprio S, Pierpont B, et al. Hepatic de novo lipogenesis in obese youth is modulated by a common variant in the GCKR gene. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(8):E1125-E1132.
44. Coviello AD, Haring R, Wellons M, et al. A genome-wide association meta-analysis of circulating sex hormone–binding globulin reveals multiple loci implicated in sex steroid hormone regulation. Plos Genetics. 2012;8(7):e1002805.

Supplementary materials

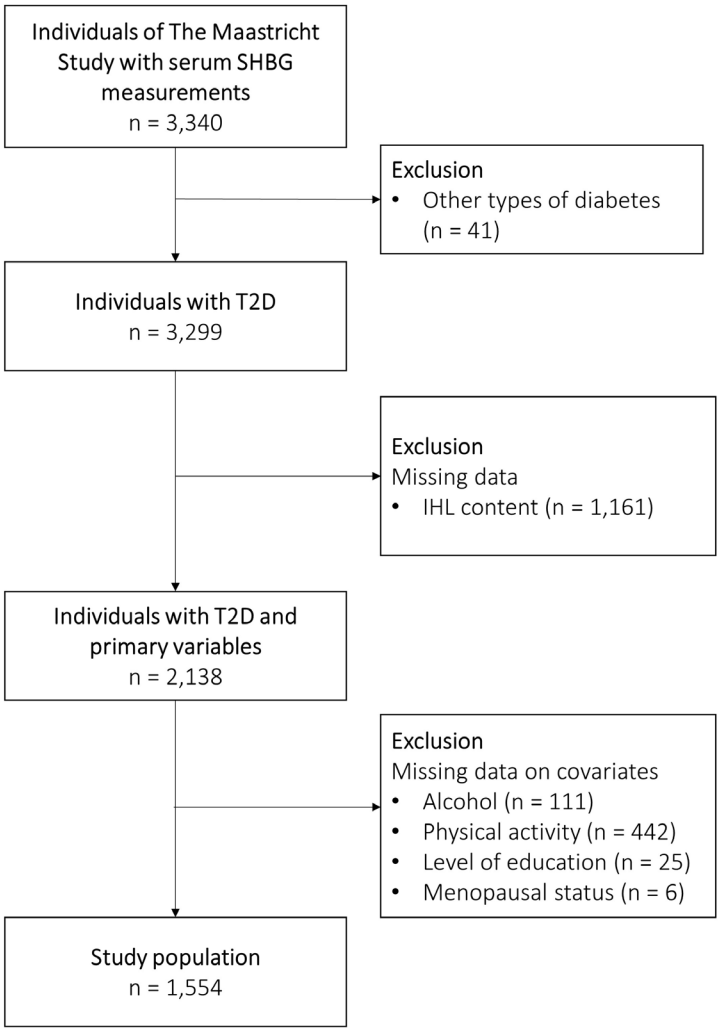

Figure S7.1 Flowchart of study population selection.
Abbreviations: IHL intrahepatic lipid; T2D type 2 diabetes; SHBG sex hormone-binding globulin.

Table S7.1 Characteristics of the study population stratified according to sex.

	Men			Women		
	Overall (n=810)	Individuals without type 2 diabetes (n=548)	Individuals with type 2 diabetes (n=262)	Overall (n=744)	Individuals without type 2 diabetes (n=637)	Individuals with type 2 diabetes (n=107)
Age, years	61 ± 8	60 ± 8	63 ± 8	58 ± 8	58 ± 8	61 ± 8
Postmenopausal, % of women				79.2	78.3	84.1
Use of oestrogen-containing medication, % of women				4.0	3.8	5.6
Education level low/medium/high, %	25.8 / 29.6 / 44.6	21.2 / 29.4 / 49.5	35.5 / 30.2 / 34.4	35.2 / 28.0 / 36.8	32.7 / 27.6 / 39.7	50.5 / 29.9 / 19.6
Smoking, never/former/current, %	36.7 / 52.0 / 11.2	38.1 / 50.7 / 11.1	33.7 / 54.8 / 11.5	39.4 / 48.1 / 12.5	39.7 / 48.0 / 12.2	42.8 / 48.6 / 14.0
Dutch healthy diet index (DHD-13)*	73.9 ± 13.9	74.3 ± 14.0	74.1 ± 13.6	80.9 ± 13.0	81.2 ± 13.0	78.9 ± 12.9
Alcohol, g/day	12.0 (3.9-23.5)	14.3 (5.6-24.7)	7.6 (1.5-20.7)	6.0 (0.8-14.2)	7.1 (1.3-14.7)	5.2 (0.1-8.8)
Physical activity, min/day	119.4 (88.3-149.4)	126.8 (97.1-154.9)	99.0 (72.0-135.3)	122.7 (100.3-148.3)	127.8 (103.4-150.1)	105.0 (80.8-137.0)
BMI, kg/m^2	27.2 ± 3.7	26.3 ± 3.1	28.7 ± 4.1	25.9 ± 4.4	25.3 ± 4.1	28.9 ± 4.9
Waist circumference, cm	99.8 ± 11.0	96.8 ± 9.3	106.1 ± 11.5	88.3 ± 11.4	86.7 ± 10.2	98.2 ± 13.0
Office systolic blood pressure, mmHg	139 ± 16	137 ± 17	142 ± 16	128 ± 16	127 ± 16	135 ± 17
Office diastolic blood pressure, mmHg	79 ± 10	79 ± 10	78 ± 10	73 ± 9	73 ± 9	74 ± 8
Total cholesterol, mmol/l	5.0 ± 1.1	5.4 ± 1.0	4.3 ± 1.0	5.6 ± 1.1	5.8 ± 1.1	4.7 ± 1.1
HDL cholesterol, mmol/l	1.4 ± 0.4	1.5 ± 0.4	1.2 ± 0.3	1.8 ± 0.5	1.8 ± 0.5	1.5 ± 0.4
LDL cholesterol, mmol/l	3.0 ± 1.0	3.3 ± 0.9	2.3 ± 0.9	3.2 ± 1.0	3.4 ± 1.0	2.5 ± 0.9
Triglycerides, mmol/l	1.3 (0.9-1.8)	1.2 (0.9-1.6)	2.2 (1.7-2.8)	1.1 (0.8-1.6)	1.1 (0.8-1.5)	1.5 (1.1-2.2)
Use of lipid-modifying medication, %	40.6	24.3	74.8	23.3	15.7	68.2
HbA1c, %	5.6 (5.4-6.3)	5.4 (5.3-5.7)	6.7 (6.3-7.3)	5.5 (5.3-5.8)	5.4 (5.3-5.7)	6.6 (6.0-7.4)
HbA1c, mmol/mol	38 (35-45)	36 (34-39)	50 (45-56)	37 (34-40)	36 (34-39)	49 (42-57)
Fasting glucose, mmol/l	5.8 (5.2-6.9)	5.4 (5.1-5.9)	7.7 (6.9-8.9)	5.2 (4.9-5.7)	5.1 (4.9-5.5)	7.3 (6.5-8.0)
Fasting insulin, pmol/l	64.3 (45.1-98.3)	60.8 (44.1-86.8)	83.3 (47.4-129.0)	53.3 (38.3-75.9)	51.0 (37.2-69.8)	82.8 (54.4-118.1)
Matsuda index	3.1 (1.9-4.7)	3.5 (2.3-5.1)	2.0 (1.2-3.0)	4.3 (2.6-6.0)	4.5 (3.1-6.2)	2.0 (1.4-2.7)
Use of glucose-lowering medication, %	24.8	0.0	76.7	10.3	0	72.0
Intrahepatic lipid content, %	4.0 (2.5-7.5)	3.5 (2.3-6.0)	6.5 (3.6-10.9)	2.7 (1.7-5.2)	2.5 (1.6-4.5)	5.5 (3.3-10.2)
Serum SHBG, nmol/l	29.6 (22.4-39.7)	31.8 (24.1-41.8)	25.4 (19.1-32.5)	45.0 (31.8-64.9)	47.3 (35.0-66.0)	31.9 (21.5-47.5)

* Dutch Healthy Diet (DHD-14) index minus alcohol component (= DHD-13). Data are presented as mean ± standard deviation or median (interquartile range), unless otherwise indicated.

Abbreviations: BMI body mass index; HDL high density lipoprotein; LDL low density lipoprotein; HbA1c haemoglobin A1c; SHBG sex hormone-binding globulin.

Table S7.2 Mediation effect of serum SHBG on the association between intrahepatic lipid content and type 2 diabetes status after accounting for oversampling of individuals with type 2 diabetes

	Men (n=810)		Women (n=744)	
	OR (95% CI)	% mediated (95% CI)±	OR (95% CI)	% mediated (95% CI)±
Crude				
Total effect*	1.12 (1.08;1.16)		1.07 (1.04;1.14)	
Direct effect*	1.10 (1.07;1.15)		1.04 (1.01;1.09)	
Indirect effect*	1.01 (1.01;1.02)	12.3 (5.5;21.3)	1.03 (1.02;1.06)	47.8 (25.5;78.8)
Model 1				
Total effect	1.12 (1.08;1.16)		1.07 (1.04;1.13)	
Direct effect	1.10 (1.06;1.14)		1.04 (1.01;1.09)	
Indirect effect	1.02 (1.01;1.03)	18.5 (9.7;29.0)	1.03 (1.02;1.06)	45.6 (23.0;79.0)
Model 2				
Total effect	1.08 (1.04;1.13)		1.04 (1.00;1.09)	
Direct effect	1.07 (1.03;1.11)		1.02 (0.98;1.07)	
Indirect effect	1.01 (1.00;1.02)	12.0 (0.5;26.6)	1.01 (1.00;1.09)	39.7 (-15.3;268.5)
Model 3				
Total effect			1.04 (1.00;1.09)	
Direct effect			1.02 (0.98;1.07)	
Indirect effect			1.02 (1.01;1.04)	53.2 (-142.2;332.5)

± % mediated is calculated as $OR^{Direct}(OR^{Indirect} - 1)/(OR^{Direct} * OR^{Indirect} - 1) * 100$. * Total effect represents association between intrahepatic lipid (IHL) content and type 2 diabetes (T2D); direct effect represents association between IHL and T2D status not attributable to serum SHBG; indirect effect represents association between IHL and T2D attributable to serum SHBG (= mediation).

Model 1: adjusted for age

Model 2: additionally adjusted for BMI, alcohol intake, Dutch Healthy Diet index (DHD-13), level of education, and total physical activity

Model 3: additionally adjusted for menopausal status and use of oestrogen-containing medication

Table S7.3 Mediation effect of serum SHBG on the association between intrahepatic lipid content and type 2 diabetes status with adjustment for waist circumference in model 2.

	Men (n=810)		Women (n=744)	
	OR (95% CI)	% mediated (95% CI)[±]	OR (95% CI)	% mediated (95% CI)[±]
Crude				
Total effect[*]	1.12 (1.09;1.17)		1.08 (1.04;1.14)	
Direct effect[*]	1.10 (1.07;1.15)		1.04 (1.01;1.09)	
Indirect effect[*]	1.02 (1.01;1.03)	17.2 (9.6;27.6)	1.04 (1.02;1.07)	50.9 (26.7;81.3)
Model 1				
Total effect	1.13 (1.09;1.18)		1.08 (1.05;1.13)	
Direct effect	1.10 (1.06;1.14)		1.04 (1.01;1.09)	
Indirect effect	1.03 (1.02;1.04)	24.6 (15.6;36.0)	1.04 (1.02;1.07)	48.5 (24.0;80.4)
Model 2				
Total effect	1.08 (1.04;1.12)		1.02 (0.98;1.07)	
Direct effect	1.06 (1.03;1.11)		1.01 (0.97;1.06)	
Indirect effect	1.01 (1.01;1.02)	18.0 (7.8;35.9)	1.01 (1.00;1.03)	49.5 (-45.3;481.0)
Model 3				
Total effect			1.02 (0.98;1.07)	
Direct effect			1.00 (0.96;1.05)	
Indirect effect			1.02 (1.00;1.04)	77.2 (-616.6;721.9)

[±] % mediated is calculated as $OR^{Direct}(OR^{Indirect} - 1)/(OR^{Direct} * OR^{Indirect} - 1) * 100$. [*] Total effect represents association between intrahepatic lipid (IHL) content and type 2 diabetes (T2D); direct effect represents association between IHL and T2D status not attributable to serum SHBG; indirect effect represents association between IHL and T2D attributable to serum SHBG (= mediation).

Model 1: adjusted for age

Model 2: additionally adjusted for waist circumference, alcohol intake, Dutch Healthy Diet index (DHD-13), level of education, and total physical activity

Model 3: additionally adjusted for menopausal status and use of oestrogen-containing medication.

Table S7.4 Mediation effect of serum SHBG on the association between intrahepatic lipids and (log) HbA1c.

	Men (n=810)		Women (n=741)	
	β (95% CI)	% mediated (95% CI)[±]	β (95% CI)	% mediated (95% CI)[±]
Crude				
Total effect[*]	0.0026 (0.0017;0.0036)		0.0020 (0.0012;0.0032)	
Direct effect[*]	0.0020 (0.0012;0.0030)		0.0012 (0.0006;0.0023)	
Indirect effect[*]	0.0005 (0.0003;0.0008)	20.5 (11.4;34.4)	0.0008 (0.0005;0.0012)	39.4 (24.4;56.9)
Model 1				
Total effect	0.0025 (0.0016;0.0035)		0.0018 (0.0011;0.0029)	
Direct effect	0.0017 (0.0009;0.0027)		0.0012 (0.0006;0.0021)	
Indirect effect	0.0008 (0.0005;0.0011)	30.1 (19.0;46.7)	0.0007 (0.0004;0.0010)	36.0 (21.1;54.2)
Model 2				
Total effect	0.0011 (0.0002;0.0021)		0.0011 (0.0006;0.0021)	
Direct effect	0.0008 (-0.0001;0.0017)		0.0008 (0.0003;0.0016)	
Indirect effect	0.0003 (0.0001;0.0005)	28.9 (11.5;107.9)	0.0003 (0.0002;0.0006)	30.2 (14.5;58.4)
Model 3				
Total effect			0.0011 (0.0006;0.0021)	
Direct effect			0.0007 (0.0002;0.0016)	
Indirect effect			0.0004 (0.0002;0.0007)	33.8 (17.2;61.8)

[±] % mediated is calculated as $\beta^{indirect}$ / β^{total} * 100. * Total effect represents association between intrahepatic lipid (IHL) content and haemoglobin A1c (HbA1c); direct effect represents association between IHL and HbA1c not attributable to serum SHBG; indirect effect represents association between IHL and HbA1c attributable to serum SHBG (= mediation).

Model 1: adjusted for age

Model 2: additionally adjusted for BMI, alcohol intake, Dutch Healthy Diet index (DHD-13), level of education, and total physical activity

Model 3: additionally adjusted for menopausal status and use of oestrogen-containing medication.

Table S7.5 Mediation effect of serum SHBG on the association between intrahepatic lipids and (log) Matsuda-index.

	Men (n=722)		Women (n=662)	
	β (95% CI)	% mediated (95% CI)[±]	β (95% CI)	% mediated (95% CI)[±]
Crude				
Total effect[*]	-0.0222 (-0.0276;-0.0179)		-0.0165 (-0.0243;-0.0119)	
Direct effect[*]	-0.0207 (-0.0261;-0.0164)		-0.0136 (-0.0207;-0.0095)	
Indirect effect[*]	-0.0015 (-0.0026;-0.0005)	6.7 (2.4;12.0)	-0.0029 (-0.0053;-0.0009)	17.8 (5.7;29.4)
Model 1				
Total effect	-0.0218 (-0.0268;-0.0176)		-0.0164 (-0.0242;-0.0118)	
Direct effect	-0.0193 (-0.0243;-0.0152)		-0.0136 (-0.0207;-0.0095)	
Indirect effect	-0.0025 (-0.0038;-0.0013)	11.4 (6.2;17.9)	-0.0028 (-0.0051;-0.0009)	17.0 (5.1;28.4)
Model 2				
Total effect	-0.0134 (-0.0182;-0.0094)		-0.0095 (-0.0150;-0.0062)	
Direct effect	-0.0127 (-0.0173;-0.0088)		-0.0087 (-0.0139;-0.0056)	
Indirect effect	-0.0007 (-0.0016;-0.0001)	5.4 (0.7;11.9)	-0.0007 (-0.0023;0.0006)	7.6 (-7.0;21.6)
Model 3				
Total effect			-0.0097 (-0.0150;-0.0065)	
Direct effect			-0.0081 (-0.0127;-0.0051)	
Indirect effect			-0.0016 (-0.0036;-0.0004)	16.3 (4.1;32.4)

[±] % mediated is calculated as $\beta^{indirect}$ / β^{total} * 100. * Total effect represents association between intrahepatic lipid (IHL) content and Matsuda index; direct effect represents association between IHL and Matsuda index not attributable to serum SHBG; indirect effect represents association between IHL and Matsuda index attributable to serum SHBG (= mediation).

Model 1: adjusted for age

Model 2: additionally adjusted for BMI, alcohol intake, Dutch Healthy Diet index (DHD-13), level of education, and total physical activity

Model 3: additionally adjusted for menopausal status and use of oestrogen-containing medication.

8

CHAPTER EIGHT

Association of common gene variants in glucokinase regulatory protein with cardiorenal disease: a systematic review and meta-analysis

Pomme I.H.G. Simons*, Nynke Simons*, Coen D.A. Stehouwer, Casper G. Schalkwijk, Nicolaas C. Schaper, Martijn C.G.J. Brouwers*
* Contributed equally

PLoS One (2018); 13 (10): e0206174

Summary

Background: Small-molecules that disrupt the binding between glucokinase and glucokinase regulatory protein (GKRP) in the liver represent a potential new class of glucose-lowering drugs. It will, however, take years before their effects on clinically relevant cardiovascular endpoints are known. The purpose of this study was to estimate the effects of these drugs on cardiorenal outcomes by studying variants in the GKRP gene (*GCKR*) that mimic glucokinase-GKRP disruptors.

Methods: The MEDLINE and EMBASE databases were searched for studies reporting on the association between *GCKR* variants (rs1260326, rs780094, and rs780093) and coronary artery disease, estimated glomerular filtration rate (eGFR), and chronic kidney disease.

Results: In total 5 coronary artery disease studies (n=274,625 individuals), 7 eGFR studies (n=195,195 individuals), and 4 chronic kidney disease studies (n=31,642 cases and n=408,432 controls) were included. Meta-analysis revealed a significant association between *GCKR* variants and coronary artery disease (OR: 1.02 per risk allele, 95% CI: 1.00;1.04, p=0.01). Sensitivity analyses showed that replacement of one large, influential coronary artery disease study by two other, partly overlapping studies resulted in similar point estimates, albeit less precise (OR: 1.02; 95% CI: 0.98;1.06 and OR: 1.02; 95% CI: 0.99;1.04). *GCKR* was associated with an improved eGFR (+0.49 ml/min, 95% CI: 0.10;0.89, p=0.01) and a trend towards protection from chronic kidney disease (OR: 0.98, 95% CI: 0.95;1.01, p=0.13).

Conclusion: This study suggests that increased glucokinase-GKRP disruption has beneficial effects on eGFR, but these may be offset by a disadvantageous effect on coronary artery disease risk. Further studies are warranted to elucidate the mechanistic link between hepatic glucose metabolism and eGFR.

Introduction

In the current area of precise medicine, there is an ongoing search for new anti-diabetic medication with different modes of action. Drugs that modulate the function of glucokinase have been the scope of diabetes research for more than a decade now[1-4]. Glucokinase plays a pivotal role in regulating pancreatic insulin secretion and hepatic glucose uptake, owing to its unique enzymatic actions[5]. It catalyzes the conversion of glucose to glucose-6-phosphate, the first step in glycolysis. To date, however, clinical trials with glucokinase activators in patients with type 2 diabetes have been disappointing, since the glucose-lowering effects were non-sustained and accompanied by an increased risk of hypoglycemia and hypertriglyceridemia[1]. Hepatoselective glucokinase activators could theoretically bypass some of these side-effects, in particular the risk of hypoglycemia[6].

An alternative way to increase hepatic glucokinase activity is to disrupt the binding between glucokinase and glucokinase regulatory protein (GKRP). GKRP is a liver-specific protein located in the nucleus that binds – and hence inactivates – glucokinase in the fasting state. In the postprandial state, glucokinase dissociates from GKRP and subsequently migrates towards the cytosolic space where it facilitates phosphorylation of glucose[7,8]. Lloyd and colleagues previously demonstrated that small molecules that disrupt the glucokinase-GKRP complex reduce plasma glucose levels without causing hypoglycemia in mice[9]. Although promising, it will probably take years before this new drug can be tested in a clinical setting.

Genetic epidemiology can be helpful to gain more insight into the clinical effects of glucokinase-GKRP disruption in humans. Since individuals are 'randomized' at birth to receive a wildtype allele or a variant that encodes GKRP that binds glucokinase less effectively, the effects of this variant on clinical endpoints can be studied as a surrogate for glucokinase-GKRP disruptors. Such a Mendelian randomization approach has been proven to be effective in predicting the (un)intended effects of new drugs[10].

We previously reviewed current literature on the cardiometabolic effects of variants in the glucokinase regulatory protein gene (*GCKR*)[11]. Individuals carrying the variant that binds glucokinase less effectively are indeed characterized by reduced fasting plasma glucose levels, but this is accompanied by an increased risk of nonalcoholic fatty liver disease (NAFLD), hypertriglyceridemia, and gout[12-14]. Of interest, there are studies suggesting that the same variant protects from chronic kidney disease[15]. Given these opposing effects it is difficult to predict what the net effect will be on coronary artery disease, one of the most clinically relevant outcomes in type 2 diabetes.

The aim of the present study was therefore to elucidate the association between *GCKR* and coronary artery disease and chronic kidney disease by conducting a systematic review and meta-analysis.

Methods

Data sources, searches, and study selection

The MEDLINE and EMBASE databases were searched for: 1) original, genetic association studies addressing the relationship between common variants in *GCKR* (rs1260326, rs780094, or rs780093) and coronary artery disease; and 2) genome-wide association studies (GWAS) on coronary artery disease, as they are likely to include the variants of interest (see Supplementary Table S8.1 for search strategy and Supplementary Figure S8.1 for flow-chart). Coronary artery disease was defined as myocardial infarction (MI), significant stenosis (i.e. ≥50%) in one or more main coronary arteries, or coronary intervention, including coronary artery bypass grafting (CABG) and percutaneous coronary intervention (PCI).

A second search was performed for the association between the common variants in *GCKR* and renal function. Studies reporting serum creatinine levels, estimated glomerular filtration rate (eGFR) (based on serum creatinine or cystatin C), or presence of chronic kidney disease were considered eligible (see Supplementary Table S8.2 for search strategy and Supplementary Figure S8.4 for flowchart).

Cross-sectional articles, written in English, German, or Dutch, were included. No publication date or publication status restrictions were imposed. The electronic searches were conducted by one researcher (P.I.H.G.S.) and completed on March 6, 2018.

Meta-analyses

Two separate systematic reviews and three meta-analyses were conducted to determine the association between 1) common variants in *GCKR* and coronary artery disease; and 2) common variants in *GCKR* and renal function, i.e. eGFR and chronic kidney disease (based on dichotomized eGFR). Selection of variants was primarily based on functionality, i.e. the variant has been shown to be functional and mimics the effects of glucokinase-GKRP disruptors (i.e. rs1260326[16,17]). In addition, variants that are in strong linkage disequilibrium with this functional variant, i.e. rs780094 or rs780093, were included as well ($r^2 \approx 0.92$ for both SNPs in both Europeans and East Asians; source: 1000 Genomes project phase 3). The systematic reviews and meta-

analyses were performed according to the PRISMA statement with the only exception of a (registered) review protocol[18].

Data extraction and quality assessment

Data extraction was done in a two-step, standardized fashion where one researcher (P.I.H.G.S.) extracted the data, which was subsequently checked by two other researchers (N.S. and M.C.G.J.B.). The following variables were extracted from the included studies: odds ratios or unstandardized beta coefficients, with 95% confidence intervals or standard errors. Authors were contacted in case of missing data (in particular for the GWAS). In case of non-response, a reminder was sent three weeks later. When more than one *GCKR* variant was reported, the functional variant (rs1260326) was chosen. The additive model was the preferred model of inheritance, based on previous *GCKR* association studies[19]. Finally, given our interest in the systematic effects of *GCKR* per se, we aimed to obtain the crude outcome variables, i.e. without adjustment for potential mediators (e.g. plasma lipids levels).

To avoid inclusion of study cohorts that were reported more than once, in particular in GWAS consortia, special attention was paid to the origin of the individual study populations. In case of overlap, the study that contained the highest number of participants was included. The quality of the study and the risk of bias were assessed by two independent researchers (P.I.H.G.S. and N.S.) according to the Newcastle-Ottawa Scale (NOS)[20].

Data synthesis and analysis

Back-transformation of the log-transformed difference in eGFR between the two *GCKR* alleles was done as described elsewhere[21]. Odds ratios and beta coefficients were meta-analyzed based on a random-effects model, using the DerSimonian-Laird method to incorporate between-study heterogeneity. Funnel plots were visually examined for asymmetry and analyzed by means of regression (Egger's test).

Since most studies (in particular GWAS) only reported the principal summary measures (i.e. odds ratios) instead of individualized data, it was not feasible to adjust for potential environmental effects, nor was it possible to assess Hardy-Weinberg equilibrium or linkage disequilibrium for each study.

Sensitivity analyses were performed to assess the impact of studies that included subjects with different ancestries, studies with low quality (defined as a NOS score <5 stars), and studies that did not report crude (or age- and/or sex-adjusted) estimates. All analyses were conducted with the 'R' statistical software (R Developmental Core Team) using the *metaphor* package[22].

Results

Systematic review and meta-analysis of the association between common variants in GCKR and coronary artery disease

The electronic search identified 3,051 unique records, which eventually resulted in five studies that were used for the meta-analysis[23-27] (see Supplementary Figure S8.1 for flowchart and reasons for exclusion). All included studies were written in English. Twenty-six studies showed overlap with one of the included studies, i.e. the combined UK Biobank, CARDIoGRAMplusC4D 1000 genomes-based GWAS, and Myocardial Infarction Genetics and CARDIoGRAM Exome dataset[24], and were therefore not included in the meta-analysis (Supplementary Table S8.3). The genetic variants of interest were often not reported in the main article, but could be found in the (online) supplementary materials of the article. For one GWAS, the authors were contacted and the requested data were kindly provided[25].

The characteristics of the included studies are shown in Table 8.1. In total, 274,625 subjects were included. In some, mainly Asian studies, the *GCKR* effect allele – defined as the allele that predisposes to reduced fasting plasma glucose levels (similar to the effect of a glucokinase-GKRP disruptor) – was the predominant allele. The overall quality of the studies was good (Supplementary Table S8.4).

Meta-analysis demonstrated that the *GCKR* effect allele was significantly associated with coronary artery disease (OR: 1.02, 95% CI: 1.00;1.04, p=0.01) (Figure 8.1). Heterogeneity was low (Q=3.30, I^2=0%)[28]. Due to the low number of included studies, a funnel plot (or testing for funnel plot asymmetry) was not included, according to previous recommendations[29,30]. Since the meta-analysis was dominated by one large study – which is composed of 76 sub-studies[31], we conducted several sensitivity analyses to test the robustness of our findings. First, this large study was replaced by another large study that combined the CARDIoGRAMplusC4D 1000 genomes-based GWAS dataset with an additional 56,354 samples (n=260,365 subjects in total, Supplementary Table S8.3)[32]. The subsequent meta-analysis revealed a similar, but less precise point estimate (OR: 1.02, 95% CI: 0.98;1.06, p=0.37, Supplementary Figure S8.2).

The initial large study was also replaced by the CARDIoGRAMplusC4D Metabochip data[33,34], which overlaps for ~55% with the CARDIoGRAMplusC4D 1000 genomes-based GWAS data (Supplementary Table S8.3)[35]. This also allowed a better sensitivity analysis stratified by ancestry, since data for Europeans only have been presented separately[34]. Again, the overall meta-analysis showed a similar, but non-significant point estimate (OR: 1.02, 95% CI: 0.99;1.04, p=0.27, Supplementary Figure S8.3).

The *GCKR* effect allele was significantly associated with coronary artery disease in studies that included subjects of European ancestry only (n=3) (OR: 1.02, 95% CI: 1.00;1.05, p=0.02), but not in studies that included subjects of Asian ancestry only (OR: 1.06, 95% CI: 0.98;1.15, p=0.13; Supplementary Figure S8.3). Of note, these effect sizes were not statistically different (p=0.36). Repeat analysis without the study with low quality[25] (i.e. NOS score <5 stars) did not affect the primary outcome.

Systematic review and meta-analysis of the association between common variants in GCKR and eGFR and chronic kidney disease

Of the 661 eligible records that were selected by our initial search, eight studies fulfilled all in- and exclusion criteria and were used for the meta-analyses (see Supplementary Figure S8.4 for flowchart and reasons for exclusion, and Supplementary Table S8.5 for duplicate studies). All included studies were written in English. The genetic variants of interest were often not reported in the main article, but could be in the (online) supplementary materials of the article. For two GWAS, the authors were contacted and the requested data were kindly provided[36,37]. Six studies reported data on creatinine-based eGFR[36,38-42], one on cystatin C-based eGFR[15], and four on chronic kidney disease[36,37,40,42]. Study characteristics of the eGFR and chronic kidney disease studies are provided in Table 8.2. All studies used only the (creatinine-based) eGFR criterion to define chronic kidney disease. Quality assessment of the included studies yielded an average score of five out of nine stars (Supplementary Table S8.6). Many studies reporting on eGFR scored low on 'comparability', i.e. the analyses were adjusted for covariates more than age and/or sex, whereas we aimed to obtain the crude relationship between *GCKR* and eGFR.

Meta-analysis, including 195,195 individuals, showed that the *GCKR* effect allele was significantly associated with an increased eGFR (0.49 ml/min, 95% CI: 0.10;0.89, p=0.01) (Figure 8.2). Heterogeneity was high (Q=43.12, I^2=88.4%). The only study that reported on cystatin C-based eGFR observed similar effect sizes, which was statistically significant in the discovery cohort (p=0.006), but not in the replication cohort (p=0.16)[15].

Table 8.1 Characteristics of included studies on coronary artery disease (CAD).

Author	Year	Ancestry	Population type	Number of cases	Number of controls	Covariates adjusted for	SNP	EAF	Outcome
Lian[23]	2013	Asian	Hospital	568	494	-	rs780093	0.52	CHD
Nelson[24]	2017	European + non-European	General + hospital	268,744*	-	Array and population structure/ancestry	rs1260326	0.40	CAD
Raffield[25]	2015	European	Type 2 diabetes	212	771	Age, sex	rs1260326	0.39	MI
Takeuchi[26]	2012	Asian	Hospital	1,347	1,337	Not specified	rs780094	0.56	CAD
Zhou[27]	2015	Asian	General + hospital	555	597	-	rs1260326	0.42	CAD

* Number of cases refers to the overall population.
Abbreviations: SNP single nucleotide polymorphism; EAF effect allele frequency; CHD coronary heart disease; MI myocardial infarction.

Table 8.2 Characteristics of included studies on eGFR and chronic kidney disease (CKD).

	Author	Year	Ancestry	Population type	Number of cases*	Number of controls	Adjusted covariates	SNP	EAF	Definition of outcome
eGFR (creatinine-based)	Bonetti[38]	2011	European	T2D	474	-	Age, sex, BMI	rs780094	0.47	MDRD
	Deshmukh[39]	2013	European	T2D	2,970	-	Age, sex, BMI, SBP, HbA1c, T2DM duration	rs1260326		MDRD
	Hishida[40]	2014	Asian	General	3,324	-	Age, sex	rs1260326	0.61	Modified MDRD
	Okada[41]	2012	Asian	General + hospital	42,451	-	Age, sex, alcohol, smoking, BMI	rs1260326	0.52	Modified CKD-EPI
	Pattaro[42]	2016	European	General + T2D	133,413	-	Age, sex	rs1260326	0.42	MDRD
	Yamada[36]	2013	Asian	Hospital	12,563	-	Age, sex	rs1260326	0.57	Modified MDRD
eGFR (cystatin C-based)	Köttgen[15]	2010	European	General + T2D	20,907	-	Age, sex	rs1260326	0.41	76.7 × (serum cystatin c)−1.19
CKD	Hishida[40]	2014	Asian	General	578	2,746	-	rs1260326	0.61	eGFR <60 ml/min/1.73m^2
	Pattaro[42]	2016	European	General + T2D	12,385	104,780	Age, sex	rs1260326	0.42	eGFR <60 ml/min/1.73m^2
	Svein-Bjornsson[37]	2014	European	Hospital	15,594	291,428	Age, sex	rs1260326	0.35	eGFR <60 ml/min/1.73m^2
	Yamada[36]	2013	Asian	Hospital	3,085	9,478	Age, sex	rs1260326	0.57	eGFR <50 ml/min/1.73m^2

* Number of cases for the eGFR trait refers to the overall population.
Abbreviations: SNP single nucleotide polymorphism; EAF effect allele frequency; BMI body mass index; SBP systolic blood pressure; MDRD: modification of diet in renal disease; CKD-EPI chronic kidney disease epidemiology collaboration.

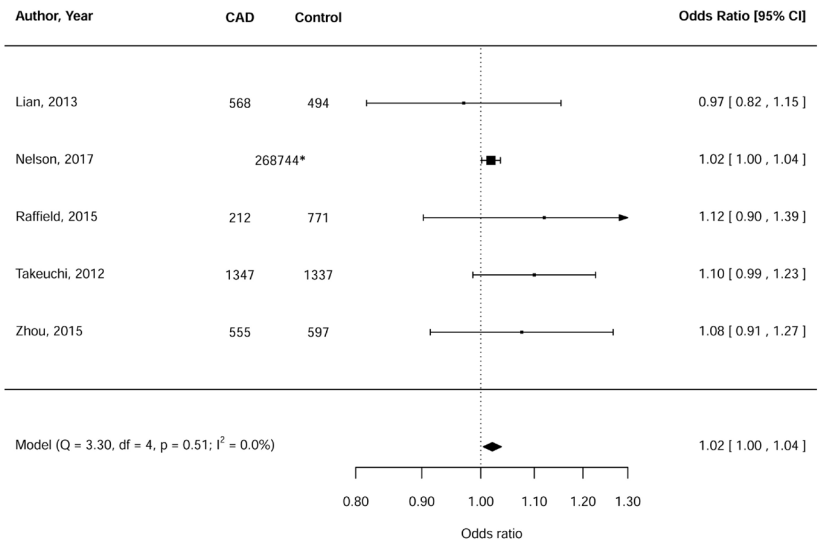

Figure 8.1 Meta-analysis of the relationship between the *GCKR* effect allele and coronary artery disease (CAD). *Number of individuals refers to the overall population.

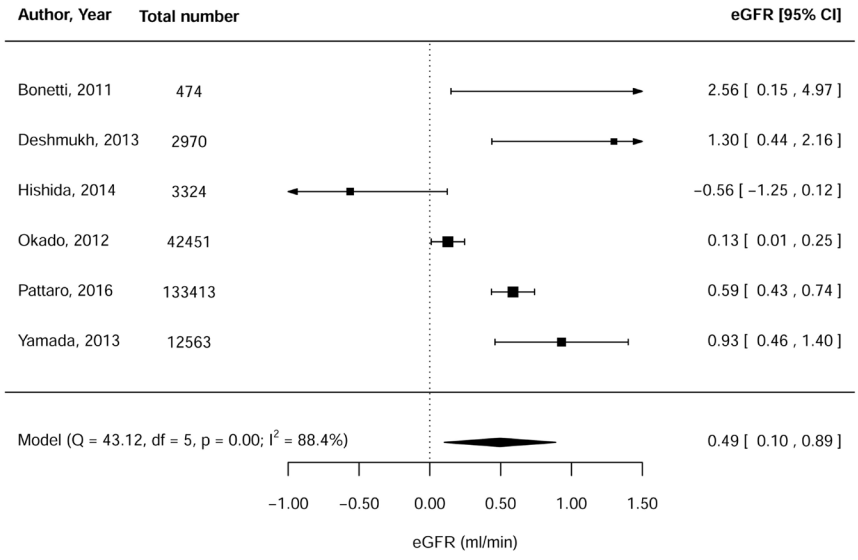

Figure 8.2 Meta-analysis of the relationship between the *GCKR* effect allele and creatinine-based estimated glomerular filtration rate (eGFR).

The meta-analysis for chronic kidney disease, including 31,642 cases and 408,432 controls, showed a protective effect of the *GCKR* effect allele on chronic kidney disease, albeit not statistically significant (OR: 0.98, 95% CI: 0.95;1.01, p=0.13; Q=5.54, I^2=45.9%) (Figure 8.3). The forest plot identified one outlying study that explained the moderate heterogeneity (Figure 8.3). Repeat analysis without this study[40] resulted in a significant, negative relationship (OR: 0.97, 95% CI: 0.95;0.99, p=0.003). The same study also accounted for the non-significant relationship with chronic kidney disease when sensitivity analyses were conducted for Asian studies only (Supplementary Figure S8.5). All chronic kidney disease studies were of sufficient quality (NOS score ≥5 stars) and did not adjust for co-variates other than age and/or sex.

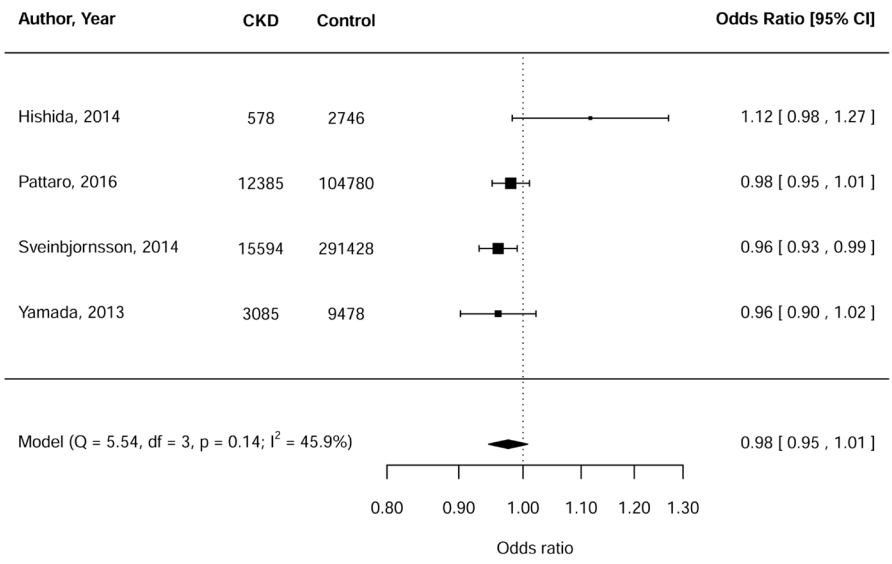

Figure 8.3 Meta-analysis of the relationship between the *GCKR* effect allele and chronic kidney disease (CKD).

Discussion

Glucokinase regulatory protein (GKRP) is a liver-specific protein that plays an important role in the regulation of hepatic glucose uptake and, consequently, de novo lipogenesis, one of the principal pathways in the development of NAFLD[11]. By studying the systemic effects of common variants in *GCKR* it is possible to gain more insight into the interaction between hepatic glucose metabolism and cardiorenal disease. Moreover, it allows an evaluation of small-molecule disruptors of the glucokinase-GKRP

complex as a potential new glucose-lowering treatment. In three meta-analyses using data from at least ~200,000 individuals, we showed that the *GCKR* effect allele – which encodes a GKRP protein that binds glucokinase less effectively – appeared to be associated with coronary artery disease, whereas a protective effect was observed for eGFR.

Previous studies have shown that the *GCKR* effect allele is associated with an atherogenic lipid profile, i.e. higher plasma triglycerides and apolipoprotein B levels, reduced HDL cholesterol levels and the presence of small-dense LDL particles[12,43,44]. In that respect it is of no surprise that we did observe a positive association of *GCKR* on coronary artery disease in our primary analysis. If, however, one would take into account the effect of *GCKR* on only plasma triglycerides, it would be anticipated to already result in an odds ratio of 1.05 to develop CAD[45]. The smaller effect estimate that was found in this study (OR: 1.02, 95% CI: 1.00;1.04) should therefore be accounted for by another, protective factor that blunts the plasma lipid-mediated effects of *GCKR* on coronary artery disease risk. *GCKR* has previously been associated with reduced fasting plasma glucose levels[12]. The hitherto reported protective effect of *GCKR* on eGFR could be another explanatory factor. Previous epidemiological studies have shown that chronic kidney disease is an independent cardiovascular risk factor[46].

The current meta-analyses were confined to creatinine-based renal outcome measures (eGFR and chronic kidney disease), since these were most frequently reported. Köttgen and colleagues showed that the positive relationship between *GCKR* and (creatinine-based) eGFR was also observed for cystatin C-based eGFR[15]. The same authors suggested that another gene, which is in linkage disequilibrium with *GCKR*, is actually responsible for the association with renal function[15]. However, previous experiments in liver-specific glucokinase knockout mice – which are metabolically opposite to increased glucokinase-GKRP disruption – are characterized by *increased* kidney damage[47], which is in line with the current study.

The mechanism by which enhanced glucokinase-GKRP disruption exerts its renoprotective effects remains to be elucidated. The *GCKR* effect allele has been associated with increased NAFLD risk, low HDL cholesterol levels, and higher urate levels[12,13,43,44,48], which in turn have been associated with deterioration of renal function[49-51]. These factors should therefore be outbalanced by factors that protect the kidney, such as lower plasma glucose levels. We cannot exclude that there are also other, yet unknown factors that contribute to the renoprotective effect of the *GCKR* effect allele. Further research is needed to identify these factors as it may have important clinical implications.

The present study may provide a glimpse into the future of what the cardiorenal effects of small-molecule disruptors of the glucokinase-GKRP complex will be as a potential new glucose-lowering drug. Although the protective effect on chronic kidney disease appears to be promising at first sight, it may be outbalanced by an increased risk to develop coronary artery disease. Furthermore, a synergistic effect between *GCKR* and type 2 diabetes on coronary artery disease risk cannot be ruled out. We previously demonstrated that the effects of the *GCKR* effect allele on plasma lipid levels were more pronounced in patients with type 2 diabetes when compared to healthy individuals[52]. A similar interaction between *GCKR* and type 2 diabetes on coronary artery disease risk would seriously decrease the applicability of small molecule disruptors of the glucokinase-GKRP complex as new antidiabetic drug. Unfortunately, there were too few studies that were specifically conducted in type 2 diabetes to formally investigate such an interaction in the current meta-analysis.

This study has several strengths and limitations. First, the meta-analysis of the association of *GCKR* with coronary artery disease depends to a large extent on the combined UK Biobank, CARDIoGRAMplusC4D 1000 genomes-based GWAS, and Myocardial Infarction Genetics and CARDIoGRAM Exome dataset, which is actually a meta-analysis by itself[31]. In subsequent sensitivity analyses we replaced this large dataset by other CARIoGRAMplusC4D-based studies that – despite a substantial overlap with the original study – included a large number of independent samples[32-34]. Although similar effect sizes were observed, statistical significance was not reached. The positive association between the *GCKR* effect allele and coronary artery disease in the primary analysis should therefore be interpreted with some caution.

Second, the definition of chronic kidney disease was only based on eGFR – not the presence of albuminuria – in all of the included studies. Both factors are part of the classification of chronic kidney disease as defined by the Kidney Disease Improving Global Outcomes (KDIGO)[53].The chronic kidney disease Genetics Consortium recently reported that the *GCKR* variant that protects from deterioration of renal function is associated with an *increased* urine albumin-creatinine ratio[51]. These findings emphasize the need for further research on the pathophysiological mechanisms relating GKRP to the kidney. Third, it is not entirely clear whether the effects of genetic variants in *GCKR* and small molecule disruptors of the glucokinase-GKRP complex are truly comparable. This is one of the general limitations of the Mendelian randomization approach in which genetic variants are used as an instrument to study the effects of a specific drug of interest. However, previous experimental studies have shown that both the product of the *GCKR* minor allele and glucokinase-GKRP disruptors cause an increased translocation of glucokinase from the nucleus towards the cytosolic space in the liver[9,17]. This explains the reduced plasma glucose levels that have been associated with both the *GCKR* minor allele and treatment with glucokinase-GKRP disruptors[9,54].

Another aspect that deserves consideration is the moderate-to-high heterogeneity that was observed in some of the meta-analyses. This could be the result of genotyping errors or difference in methodology, such as discrepancies in outcome measures (particularly for coronary artery disease) or study populations (e.g. population-based versus hospital-based). Although ancestry did not account for the moderate-to-high heterogeneity, the number of studies was too small to make strong inferences. Furthermore, differences in diet could contribute to the observed heterogeneity given the previously reported *GCKR*-diet interaction on plasma triglycerides levels[55,56]. It is, however, unlikely that these factors truly account for the opposing effect sizes that were present in the individual studies, e.g. *GCKR* seemed to protect from chronic kidney disease in one Japanese cohort[36,41] whereas a predisposing effect appeared to be present in one other[40]. These opposing effects could simply be the consequence of chance, especially in small-sized studies with few events. Alternatively, *GCKR* could theoretically be in linkage disequilibrium with a gene that exerts an opposing effect on cardiorenal risk in certain but not all populations. These opposing effects could have important therapeutic implications if they would be inherent to GKRP function and therefore deserve further attention.

A final limitation was that we were forced to exclude a considerable amount of studies, and hence a substantial number of subjects, from the meta-analyses because of partial overlap of individual study cohorts. Yet, we were still able to include a high number of individuals, ranging from ~200,000 to 400,000 in the three meta-analyses, which can be attributed to our search strategy that was not confined to studies specifically reporting on *GCKR*. We correctly assumed that GWAS were likely to include our variants of interest without reporting in the manuscript's title or abstract.

Conclusions

The present study extends our knowledge on the systemic effects of enhanced disruption of the glucokinase-GKRP complex by demonstrating that the *GCKR* effect allele is associated with a better eGFR. A disadvantageous effect on coronary artery disease risk can, however, not be ruled out. These findings question the benefits and applicability of small molecule disruptors of the glucokinase-GKRP complex as a potential new class of antidiabetic drugs. Further studies are warranted to identify the factor that mediates the renoprotective effects of enhanced disruption of the glucokinase-GKRP complex.

References

1. Meininger G.E., Scott R., Alba M., et al. Effects of MK-0941, a novel glucokinase activator, on glycemic control in insulin-treated patients with type 2 diabetes. *Diabetes Care.* 2011;34(12):2560-2566.
2. Denney W.S., Denham D.S., Riggs M.R., Amin N.B. Glycemic Effect and Safety of a Systemic, Partial Glucokinase Activator, PF-04937319, in Patients With Type 2 Diabetes Mellitus Inadequately Controlled on Metformin-A Randomized, Crossover, Active-Controlled Study. *Clinical pharmacology in drug development.* 2016;5(6):517-527.
3. Sarabu R., Bizzarro F.T., Corbett W.L., et al. Discovery of piragliatin--first glucokinase activator studied in type 2 diabetic patients. *Journal of medicinal chemistry.* 2012;55(16):7021-7036.
4. Johnson D., Shepherd R.M., Gill D., Gorman T., Smith D.M., Dunne M.J. Glucose-dependent modulation of insulin secretion and intracellular calcium ions by GKA50, a glucokinase activator. *Diabetes.* 2007;56(6):1694-1702.
5. Agius L. Glucokinase and molecular aspects of liver glycogen metabolism. *Biochem J.* 2008;414(1):1-18.
6. Pfefferkorn J.A., Guzman-Perez A., Litchfield J., et al. Discovery of (S)-6-(3-cyclopentyl-2-(4-(trifluoromethyl)-1H-imidazol-1-yl)propanamido)nicotini c acid as a hepatoselective glucokinase activator clinical candidate for treating type 2 diabetes mellitus. *Journal of medicinal chemistry.* 2012;55(3):1318-1333.
7. de la Iglesia N., Veiga-da-Cunha M., Van Schaftingen E., Guinovart J.J., Ferrer J.C. Glucokinase regulatory protein is essential for the proper subcellular localisation of liver glucokinase. *FEBS letters.* 1999;456(2):332-338.
8. van Schaftingen E., Vandercammen A., Detheux M., Davies D.R. The regulatory protein of liver glucokinase. *Adv Enzyme Regul.* 1992;32:133-148.
9. Lloyd D.J., St Jean D.J., Jr., Kurzeja R.J., et al. Antidiabetic effects of glucokinase regulatory protein small-molecule disruptors. *Nature.* 2013;504(7480):437-440.
10. Walker V.M., Davey Smith G., Davies N.M., Martin R.M. Mendelian randomization: a novel approach for the prediction of adverse drug events and drug repurposing opportunities. *Int J Epidemiol.* 2017;46(6):2078-2089.
11. Brouwers M.C., Jacobs C., Bast A., Stehouwer C.D., Schaper N.C. Modulation of Glucokinase Regulatory Protein: A Double-Edged Sword? *Trends Mol Med.* 2015;21(10):583-594.
12. Orho-Melander M., Melander O., Guiducci C., et al. Common missense variant in the glucokinase regulatory protein gene is associated with increased plasma triglyceride and C-reactive protein but lower fasting glucose concentrations. *Diabetes.* 2008;57(11):3112-3121.
13. Speliotes E.K., Yerges-Armstrong L.M., Wu J., et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet.* 2011;7(3):e1001324.
14. Matsuo H., Yamamoto K., Nakaoka H., et al. Genome-wide association study of clinically defined gout identifies multiple risk loci and its association with clinical subtypes. *Annals of the rheumatic diseases.* 2016;75(4):652-659.
15. Kottgen A., Pattaro C., Boger C.A., et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet.* 2010;42(5):376-384.
16. Beer N.L., Tribble N.D., McCulloch L.J., et al. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum Mol Genet.* 2009;18(21):4081-4088.
17. Rees M.G., Wincovitch S., Schultz J., et al. Cellular characterisation of the GCKR P446L variant associated with type 2 diabetes risk. *Diabetologia.* 2012;55(1):114-122.
18. Moher D., Liberati A., Tetzlaff J., Altman D.G. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *The BMJ.* 2009;339:b2535.
19. Sparso T., Andersen G., Nielsen T., et al. The GCKR rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinaemia, and reduced risk of type 2 diabetes. *Diabetologia.* 2008;51(1):70-75.
20. Wells G.A., Shea B., O'Connell D., et al. The Newcastle–Ottawa Scale (NOS) for Assessing the Quality of Non-Randomized Studies in Meta-Analysis. 2010;50(4):1088-1010.

21. Higgins J.P., White I.R., Anzures-Cabrera J. Meta-analysis of skewed data: combining results reported on log-transformed or raw scales. *Stat Med.* 2008;27(29):6072-6092.
22. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software.* 2010;36:1-48.
23. Lian J., Guo J., Chen Z., et al. Positive association between GCKR rs780093 polymorphism and coronary heart disease in the aged Han Chinese. *Dis Markers.* 2013;35(6):863-868.
24. Nelson C.P., Goel A., Butterworth A.S., et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet.* 2017;49(9):1385-1391.
25. Raffield L.M., Cox A.J., Carr J.J., et al. Analysis of a cardiovascular disease genetic risk score in the Diabetes Heart Study. *Acta Diabetol.* 2015;52(4):743-751.
26. Takeuchi F., Isono M., Katsuya T., et al. Association of genetic variants influencing lipid levels with coronary artery disease in Japanese individuals. *PLoS One.* 2012;7(9):e46385.
27. Zhou Y.J., Hong S.C., Yin R.X., Yang Q., Cao X.L., Chen W.X. Polymorphisms in the GCKR are associated with serum lipid traits, the risk of coronary artery disease and ischemic stroke. *Int J Clin Exp Med.* 2015;8(7):10678-10686.
28. Higgins J.P., Thompson S.G., Deeks J.J., Altman D.G. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557-560.
29. Sterne J.A., Sutton A.J., Ioannidis J.P., et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ.* 2011;343:d4002.
30. Higgins J.P.T., Green S., (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from http://handbook.cochrane.org.*
31. Nelson C.P., Goel A., Butterworth A.S., et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet.* 2017;49(9):1385-1391.
32. Zhao W., Rasheed A., Tikkanen E., et al. Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nat Genet.* 2017;49(10):1450-1457.
33. Deloukas P., Kanoni S., Willenborg C., et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nature Genetics.* 2013;45(1):25-33.
34. Schunkert H., Konig I.R., Kathiresan S., et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet.* 2011;43(4):333-338.
35. Nikpay M., Goel A., Won H.H., et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet.* 2015;47(10):1121-1130.
36. Yamada Y., Nishida T., Ichihara S., et al. Identification of chromosome 3q28 and ALPK1 as susceptibility loci for chronic kidney disease in Japanese individuals by a genome-wide association study. *J Med Genet.* 2013;50(6):410-418.
37. Sveinbjornsson G., Mikaelsdottir E., Palsson R., et al. Rare mutations associating with serum creatinine and chronic kidney disease. *Hum Mol Genet.* 2014;23(25):6935-6943.
38. Bonetti S., Trombetta M., Boselli M.L., et al. Variants of GCKR affect both beta-cell and kidney function in patients with newly diagnosed type 2 diabetes: the Verona newly diagnosed type 2 diabetes study 2. *Diabetes Care.* 2011;34(5):1205-1210.
39. Deshmukh H.A., Palmer C.N., Morris A.D., Colhoun H.M. Investigation of known estimated glomerular filtration rate loci in patients with type 2 diabetes. *Diabet Med.* 2013;30(10):1230-1235.
40. Hishida A., Takashima N., Turin T.C., et al. GCK, GCKR polymorphisms and risk of chronic kidney disease in Japanese individuals: data from the J-MICC Study. *J Nephrol.* 2014;27(2):143-149.
41. Okada Y., Sim X., Go M.J., et al. Meta-analysis identifies multiple loci associated with kidney function-related traits in east Asian populations. *Nat Genet.* 2012;44(8):904-909.
42. Pattaro C., Teumer A., Gorski M., et al. Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat Commun.* 2016;7:10023.
43. Chambers J.C., Zhang W., Sehmi J., et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *Nat Genet.* 2011;43(11):1131-1138.
44. Kozian D.H., Barthel A., Cousin E., et al. Glucokinase-activating GCKR polymorphisms increase plasma levels of triglycerides and free fatty acids, but do not elevate cardiovascular risk in the Ludwigshafen

Risk and Cardiovascular Health Study. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme.* 2010;42(7):502-506.
45. Do R., Willer C.J., Schmidt E.M., et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet.* 2013;45(11):1345-1352.
46. Go A.S., Chertow G.M., Fan D., McCulloch C.E., Hsu C.-y. Chronic Kidney Disease and the Risks of Death, Cardiovascular Events, and Hospitalization. *New England Journal of Medicine.* 2004;351(13):1296-1305.
47. Gu Y., Mao Y., Li H., et al. Long-term renal changes in the liver-specific glucokinase knockout mouse: implications for renal disease in MODY2. *Translational research : the journal of laboratory and clinical medicine.* 2011;157(3):111-116.
48. Kolz M., Johnson T., Sanna S., et al. Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. *PLoS Genet.* 2009;5(6):e1000504.
49. Musso G., Gambino R., Tabibian J.H., et al. Association of Non-alcoholic Fatty Liver Disease with Chronic Kidney Disease: A Systematic Review and Meta-analysis. *PLOS Medicine.* 2014;11(7):e1001680.
50. De Cosmo S., Viazzi F., Pacilli A., et al. Serum Uric Acid and Risk of CKD in Type 2 Diabetes. *Clinical journal of the American Society of Nephrology : CJASN.* 2015;10(11):1921-1929.
51. Lanktree M.B., Theriault S., Walsh M., Pare G. HDL Cholesterol, LDL Cholesterol, and Triglycerides as Risk Factors for CKD: A Mendelian Randomization Study. *Am J Kidney Dis.* 2017.
52. Simons N., Dekker J.M., van Greevenbroek M.M., et al. A Common Gene Variant in Glucokinase Regulatory Protein Interacts With Glucose Metabolism on Diabetic Dyslipidemia: the Combined CODAM and Hoorn Studies. *Diabetes Care.* 2016;39(10):1811-1817.
53. Ketteler M., Block G.A., Evenepoel P., et al. Executive summary of the 2017 KDIGO Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD) Guideline Update: what's changed and why it matters. *Kidney international.* 2017;92(1):26-36.
54. Dupuis J., Langenberg C., Prokopenko I., et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet.* 2010;42(2):105-116.
55. Tam C.H., Wang Y., Lee H.M., et al. Early gene-diet interaction between glucokinase regulatory protein (GCKR) polymorphism, vegetable and fish intakes in modulating triglyceride levels in healthy adolescents. *Nutr Metab Cardiovasc Dis.* 2015;25(10):951-958.
56. Rousseaux J., Duhamel A., Dumont J., et al. The n-3 long-chain PUFAs modulate the impact of the GCKR Pro446Leu polymorphism on triglycerides in adolescents. *J Lipid Res.* 2015;56(9):1774-1780.

Supplementary materials

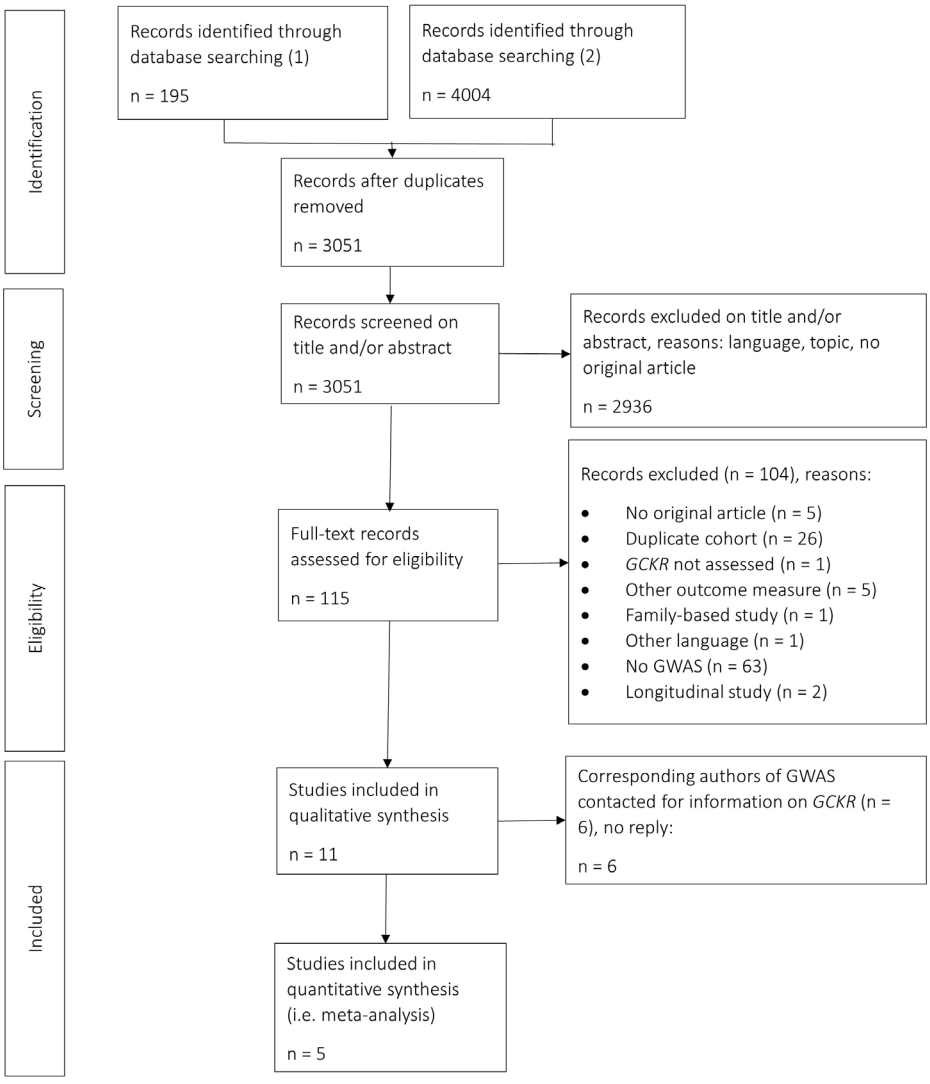

Figure S8.1 Flowchart of the systematic review on coronary artery disease (CAD).

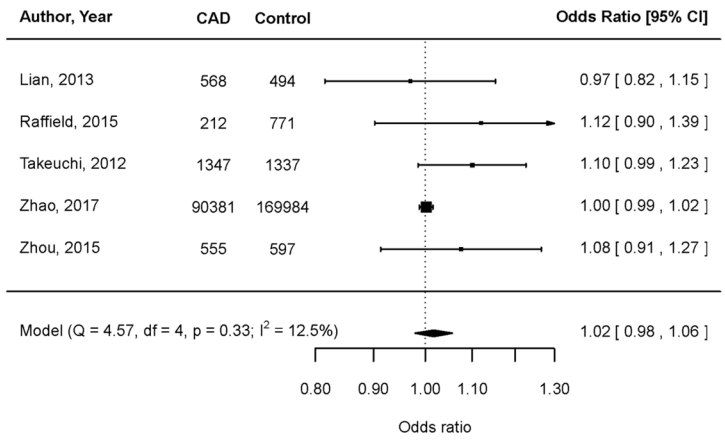

Figure S8.2 Forest plot of the meta-analysis on coronary artery disease (CAD) – sensitivity analysis. For this analysis, the combined UK Biobank, CARDIoGRAMplusC4D 1000 genomes-based GWAS, and Myocardial Infarction Genetics and CARDIoGRAM Exome dataset[1] was replaced by CARDIoGRAMplusC4D 1000 genomes-based GWAS dataset combined with 56,354 samples[2].

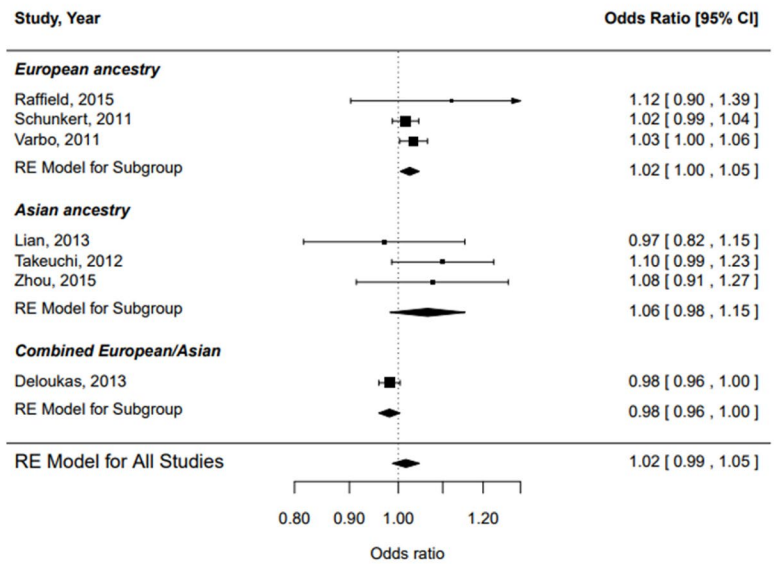

Figure S8.3 Forest plot of the meta-analysis on coronary artery disease (CAD) – stratified by ancestry. For this analysis, the combined UK Biobank, CARDIoGRAMplusC4D 1000 genomes-based GWAS, and Myocardial Infarction Genetics and CARDIoGRAM Exome dataset[1] was replaced by the CARDIoGRAMplusC4D Metabochip dataset[3,4], which allows stratification by ancestry. The CARDIoGRAMplusC4D Metabochip dataset overlaps for ~55% with the CARDIoGRAMplusC4D 1000 genomes-based GWAS. The Copenhagen City Heart Study, the Copenhagen General Population Study, and the Copenhagen Ischemic Heart Disease Study[5] were not part of the CARDIoGRAMplusC4D Metabochip dataset and were therefore included in this meta-analysis.

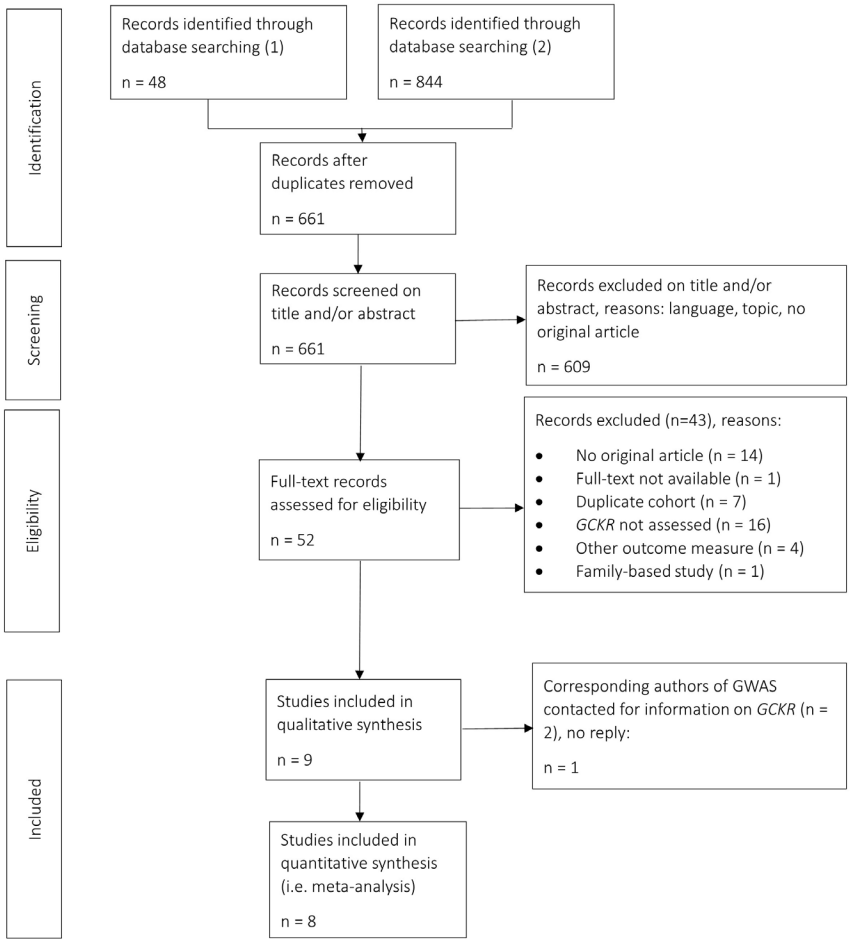

Figure S8.4 Flowchart of the systematic review on eGFR and chronic kidney disease (CKD).

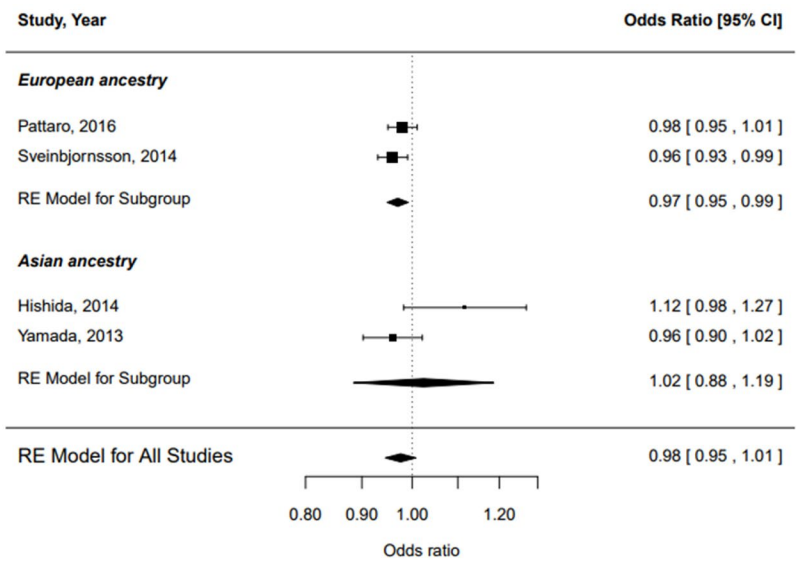

Figure S8.5 Forest plot of the meta-analysis on chronic kidney disease (CKD) – stratified by ancestry.

Table S8.1 Search strategy for coronary artery disease (CAD).

Search strategy: MEDLINE (OVID) 1946 to 2018 week 10, EMBASE (OVID) 1974 to 2018 week 10.

Search strategy 1: CAD and *GCKR*

1. Acute coronary syndrome/
2. Coronary artery disease/
3. Ischemic heart disease/
4. Heart disease/
5. Coronary artery atherosclerosis/
6. Coronary artery occlusion/
7. Coronary artery thrombosis/
8. Cardiovascular disease/
9. Myocardial infarction/
10. Stroke/
11. Cerebrovascular accident/
12. Peripheral vascular disease/
13. Transient ischemic attack/
14. Peripheral arterial disease/
15. Acute coronary syndrome.mp
16. Coronary artery disease.mp
17. Ischemic heart disease.mp
18. Heart disease.mp
19. Coronary artery atherosclerosis.mp
20. Coronary artery occlusion.mp
21. Coronary artery thrombosis.mp
22. CAD.mp
23. Cardiovascular disease.mp
24. CVD.mp
25. Cardiac attack.mp
26. Heart attack.mp
27. Myocardial infarction.mp
28. MI.mp
29. Stroke.mp
30. Cerebrovascular accident.mp
31. Ischemic stroke.mp
32. Peripheral vascular disease.mp
33. Transient ischemic attack.mp
34. TIA.mp
35. Peripheral artery disease.mp
36. CVA.mp
37. GCKR.mp
38. GKRP.mp
39. Glucokinase regulatory protein.mp
40. Rs1260326.mp
41. P446L.mp
42. Rs780094.mp
43. Rs780093.mp
44. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36
45. 37 or 38 or 39 or 40 or 41 or 42 or 43
46. 44 and 45

MEDLINE = 82 retrieved, EMBASE = 113 retrieved

Search strategy 2: CAD and GWAS

1. Acute coronary syndrome/
2. Coronary artery disease/
3. Ischemic heart disease/
4. Heart disease/
5. Coronary artery atherosclerosis/
6. Coronary artery occlusion/
7. Coronary artery thrombosis/
8. Cardiovascular disease/
9. Myocardial infarction/
10. Stroke/
11. Cerebrovascular accident/
12. Peripheral vascular disease/
13. Transient ischemic attack/
14. Peripheral arterial disease/
15. Acute coronary syndrome.mp
16. Coronary artery disease.mp
17. Ischemic heart disease.mp
18. Heart disease.mp
19. Coronary artery atherosclerosis.mp
20. Coronary artery occlusion.mp
21. Coronary artery thrombosis.mp
22. CAD.mp
23. Cardiovascular disease.mp
24. CVD.mp
25. Cardiac attack.mp
26. Heart attack.mp
27. Myocardial infarction.mp
28. MI.mp
29. Stroke.mp
30. Cerebrovascular accident.mp
31. Ischemic stroke.mp
32. Peripheral vascular disease.mp
33. Transient ischemic attack.mp
34. TIA.mp
35. Peripheral artery disease.mp
36. CVA.mp
37. GWAS.mp
38. GWA study.mp
39. Genome-wide association study/
40. Genome wide association study.mp
41. Whole genome association study.mp
42. WGA study.mp
43. WGAS.mp
44. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36
45. 37 or 38 or 39 or 40 or 41 or 42 or 43
46. 43 and 45

MEDLINE = 2026 retrieved, EMBASE = 1978 retrieved

Table S8.2 Search strategy for eGFR and chronic kidney disease (CKD).

Search strategy: MEDLINE (OVID) 1946 to 2018 week 10, EMBASE (OVID) 1974 to 2018 week 10.

Search strategy 1: CKD and *GCKR*

1. Kidney disease/
2. Renal insufficiency, chronic/
3. Kidney failure, chronic/
4. Glomerular filtration rate/
5. Creatine/
6. Cystatin C/
7. Chronic kidney disease/
8. Glomerulopathy/
9. Kidney disease.mp
10. Chronic renal insufficiency.mp
11. Chronic kidney failure.mp
12. Glomerular filtration rate.mp
13. Creatine.mp
14. Cystatin C.mp
15. Chronic kidney disease.mp
16. CKD.mp
17. Renal disease.mp
18. Glomerulopathy.mp
19. GFR.mp
20. Estimated glomerular filtration rate.mp
21. eGFR.mp
22. GCKR.mp
23. GKRP.mp
24. Glucokinase regulatory protein.mp
25. Rs1260326.mp
26. P446L.mp
27. Rs780094.mp
28. Rs780093.mp
29. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21
30. 22 or 23 or 24 or 25 or 26 or 27 or 28
31. 29 and 30

MEDLINE = 20 retrieved, EMBASE = 28 retrieved

Search strategy 2: CKD and GWAS

1. Kidney disease/
2. Renal insufficiency, chronic/
3. Kidney failure, chronic/
4. Glomerular filtration rate/
5. Creatine/
6. Cystatin C/
7. Chronic kidney disease/
8. Glomerulopathy/
9. Kidney disease.mp
10. Chronic renal insufficiency.mp
11. Chronic kidney failure.mp
12. Glomerular filtration rate.mp
13. Creatine.mp

14. Cystatin C.mp
15. Chronic kidney disease.mp
16. CKD.mp
17. Renal disease.mp
18. Glomerulopathy.mp
19. GFR.mp
20. Estimated glomerular filtration rate.mp
21. eGFR.mp
22. GWAS.mp
23. GWA study.mp
24. Genome-wide association study/
25. Genome wide association study.mp
26. Whole genome association study.mp
27. WGA study.mp
28. WGAS.mp
29. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21
30. 22 or 23 or 24 or 25 or 26 or 27 or 28
31. 29 and 30

MEDLINE = 352 retrieved, EMBASE = 492 retrieved

Table S8.3 Overview of the excluded coronary artery disease (CAD) studies with duplicate cohorts.

Excluded study	Study cohorts (with duplicate cohort in bold)	Study in which duplicate cohort is included
Angelakopoulou (2012)[6]	Northwick Park Heart Study II (NPHS II), British Regional Heart Study (BRHS), English Longitudinal Study of Ageing (ELSA), Edinburgh Artery Study (EAS), Whitehall II study, The 1958 Birth Cohort (1958BC), The Medical Research Council National Survey of Health and Development (NSHD), Southampton Atherosclerosis Study (SAS), The Stockholm Heart Epidemiology Program (SHEEP), **The Wellcome Trust Case Control Consortium (WTCCC)**, University College Diabetes and Cardiovascular Study (UDACS), Ealing Diabetes Study of Coagulation (EDS), The MRC British Genetics of Hypertension (BRIGHT)	Nelson (2017)[1]
Bi (2010)[7]	**The atherosclerosis risk in communities (ARIC)***	Nelson (2017)[1]
Burton (2007)[8]	**WTCCC**, 1958 British Birth Cohort, UK blood donor service	Nelson (2017)[1]
Davies (2012)[9]	**The Ottowa heart genomics study (OHGS), Cleveland clinic gene bank (CCGB), WTCCC, INTERHEART, Duke Cathgen Study (DUKE)**	Nelson (2017)[1]
Dehghan (2016)[10]	**Age, gene/environment susceptibility-Reykjavik Study (AGES)*, ARIC***, cardiovascular health study (CHS), **Family heart study (FHS), Rotterdam study**	Nelson (2017)[1]

Table S8.3 (continued)

Excluded study	Study cohorts (with duplicate cohort in bold)	Study in which duplicate cohort is included
Deloukas (2013)[4]	**Artherosclerotic Disease, Vascular Function & Genetic Epidemiology study (ADVANCE), The academic Medical Center Amsterdam Premature Atherosclerosis Study (AMC-PAS),** Angio-Lueb/KORAF3, **Cardiogenics study (CARDIOGENICS)**, The Dietary, lifestyle and genetic determinants of obesity and metabolic syndrome study (DILGOM), **DUKE, Estonian genome center of university of Tartu (EGCUT GWAS, EGCUT metabochip), The European prospective investigation into cancer (EPIC), Functional genomic diagnostic tools for coronary artery disease (FGENTCARD),** Fragmin and Fast Revascularization (FRISCII-GLACIER), Gene x lifestyle interactions and compmlex traits involed in elevated disease risk (GLACIER), **The genetics of diabetes audit and research in Tayside Scotland (GoDARTS), MRC/BHF heart protection study (HPS), The INTERHEART study (ITH), London life sciences population study (LOLIPOP), Ludwigshafen Risk and cardiovascular health study and echonicoccus Multilocularis and internal diseases in Leutkirch study (LURIC-EMIL),** METabolic Syndrome In Men (METSIM), **Monica, Risk, Genetics, Archiving and monograph (MORGAM-FIN, MORGAM-FRA, MORGAM-GER, MORAGM-ITA, MORGAM-UNK), OHGS, Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS),** Pfizer-MGH-Broad (PMD), **PopGEN, European collaborative study of the genetics of precocious coronary artery disease (PROCARDIS), The Pakistan Risk of Myocardial Infarction Study (PROMIS GWAS, PROMIS Metabochip),** SCARF-SHEEP, Swedish Twin Registry (STR), **The Hellenic study of interactions between Snps and eating in atherosclerosis susceptibility (THISEAS), Uppsala longitudinal study of adult men (ULSAM), WTCCC CAD 2, COROGENE,** The Finnish cardiovascular study (FINCAVAS), **Genomics Research in Cardiovascular disease (GenRIC)**	Nelson (2017)[1]
Divers (2017)[11]	**African American Diabetes Heart Study (AA-DHS),** Jackson Heart Study	Raffield (2015)[12]
Erdmann (2010)[13]	**German myocardial infarction family study (GerMIFS III)**	Nelson (2017)[1]
Howson (2017)[14]	**The Copenhagen Ischaemic Heart Disease Study (CIHDS), The Copenhagen General Population Study (CGPS), Copenhagen City Heart Study (CCHS), EPIC-CVD, Bangladesh Risk of Acute Vascular Events (BRAVE), PROMIS, ARIC, Women's Health Initiative (WHI), Myocardial infarction genetics consortium (MIGen),** TAIwan metaboCHIp Consortium (TAICHI)	Nelson (2017)[1]
Kozian (2010)[15]	**LURIC**	Nelson (2017)[1]
Lu (2014)[16]	**Beijing Atherosclerosis Study (BAS), China atherosclerosis study (CAS),** CARDIOGRAM cohorts, China CHS, China Collaborative Study of Cardiovascular Epidemiology	Nelson (2017)[1]
Lettre (2011)[17]	**ARIC***, The coronary artery risk development in young adults (CARDIA), Cleveland family study (CFS), Jackson heart Study (JHS) and Multi-Ethnic Study of Atherosclerosis (MESA)	Nelson (2017)[1]

Table S8.3 (continued)

Excluded study	Study cohorts (with duplicate cohort in bold)	Study in which duplicate cohort is included
Nikpay (2015)[18]	**ADVANCE, AGES*, ARIC*, BAS (Beijing atherosclerosis study), CARDIOGENICS, CAS, CCGB, COROGENE, DUKE, EGCUT, Family Heart Study (FamHS)*, FGENTCARD, FHS, GenRIC, GerMIFS I-IV, GoDARTS, MRC/BHF Heart Protection Study (HPS), HSDSS, BioMe Biobank Program, INTERHEART, LIFE-Heart, LOLIPOP, LURIC, case control study (MAYO-VDB), Medstar cardiac catheterization study (MedStar), MIGen, OHGS, Univeristy of Pennsylvania Medical Centre cardiac catherization study (PennCATH), PIVUS, A subset of FINRISK cohort study (PREDICTCVD), PROCARDIS, PROMIS, Prospective study of pravastatin in the elderly at risk (PROSPER)*, Rotterdam Study*, THISEAS, TWINGENE, ULSAM, Women's genome health study (WGHS), WTCCC**	Nelson (2017)[1]
O'Donnell (2011)[19]	**Cohorts for heart and aging research in genomic epidemiology (CHARGE),** genetic epidemiology network of arteriopathy study (GENOA)	Nelson (2017)[1]
Pfister (2011)[20]	**EPIC Norfolk**	Nelson (2017)[1]
Schunkert (2011)[3]	**ADVANCE**, **Coronary artery disease and omics (CADomics), CHARGE,** Diabetes epidemiology: collaborative analysis of diagnostic criteria in Europe (deCODE CAD), **GerMIFS I, GerMIFS II, GerMIFS III, LURIC/**AtheroRemo 1, **LURIC**/AtheroRemo 2, **MedStar, MIGen, OHGS1, PennCATH, WTCCC**, Acute Myocardial Infarction Gene Study / Dortmund Health Study (AMI/DHS), AMC-PAS, Angio-Lueb/Gokard, CHAOS, **Cleveland Clinic GeneBank/OHGS2, EPIC-CAD,** GENDER, GraceGenetics, **INTERHEART**, Intermountain Heart Collaborative Study (IHCS), Irish Family Study (IFS), **Italian atherosclerosis, thrombosis and vascular biology study (IATVB), LEEDS, Malmo Diet and cancer study-cardiovascular cohort (MDCS),** Mid-America Mid-America Heart Insitute (MAHI), **PopGen**, SAS, Study of Myoarcial Infarction in Leiden (SMILE), SHEEP, The Emory Genebank Study, The Johns Hopkins GeneSTAR Research Program, The New Zealand CAD Study, **THISEAS**, UKMI, **Verona Heart Study (VHS)**	Nelson (2017)[1]
Stark (2009)[21]	**German MI Family Study**	Nelson (2017)[1]
The IBC 50K CAD Consortium (2011)[22]	**ARIC***, BLOODOMICS-Dutch (**AMC-PAS** + AGNES), BLOODOMICS-**German LURIC** + Mannheim), **British Heart Foundation Family Heart Study (BHF-FHS),** The cardiovascular Health Study (CHS), CARDIA, **FHS**, **LOLIPOP**, MONICA-KORA, **PennCATH**, **PROCARDIS**, **PROMIS**	Nelson (2017)[1]
Trégouët (2009)[23]	**WTCCC**	Nelson (2017)[1]
Varbo (2011)[5]	**CCHS, CGPS, CIHDS**	Nelson (2017)[1]
Wang (2011)[24]	**GeneID population China**	Lian (2013)[25]

Table S8.3 (continued)

Excluded study	Study cohorts (with duplicate cohort in bold)	Study in which duplicate cohort is included
Webb (2017)[26]	**ATVB, BHF-FHS, Vanderbilt University Medical Center Biorepository (BioVU), DUKE, EPIC, First-time incidence of myocardial infarction in the AC county 3 (FIA3), GoDARTS, EGCUT, German CAD North, German CAD South, Nord-Trondelag health study (HUNT), BioMe Biobank, MDC, Montreal heart institute study (MHI), OHS, PAS-AMC, PennCath, PROCARDIS, VHS, WHI**	Nelson (2017)[1]
Wild (2012)[27]	The Gutenberg Heart Study (GHS), Atherogene Registry, Gutenberg Heart Express Study (GHSExpress), **CHARGE, GerMIFSI, GerMIFSII, MedStar, PennCATH, The MIGen consortium, WTCCC-CAD,** AngioLueb, Etude Cas-Témoin sur l'Infarctus du Myocarde (ECTIM), **LURIC, MORGAM, Popgen**	Nelson (2017)[1]
Willer (2008)[28]	**WTCCC**	Nelson (2017)[1]
Yaghootkar (2014)[29]	**ADVANCE**, CADomics, CHARGE, DeCode, **GERMifs, LURIC, MedStar, MIGen, OHGS, PennCATH, WTCCC**	Nelson (2017)[1]
Zhao (2017)[2]	**PROMIS**, Risk Assessment of cerebrovascular events study (RACE), **BRAVE**, A prospective cohort to determine evnrionment and genetic determinants of metabolic syndrome related factors (EPIDREAM), **FINRISK, MedStar, MDC, PennCATH, LOLIPOP**, The Singapore Indian Eye Study (SINDI), **The Khatri Sikh Diabetes Study (SDS)**, TAICHI, BioBank Japan (BBJ)	Nelson (2017)[1]

*Cohorts AGES, ARIC, FamHS, FHS, PROSPER, and Rotterdam Study are part of the CHARGE consortium.

Table S8.4 Quality assessment of the coronary artery disease (CAD) studies based on the Newcastle-Ottawa Scale (NOS).

References	Selection				Comparability	Exposure/Outcome			Quality judgment
	1	2	3	4	1	1	2	3	
Lian (2013)[25]	★	★		★	★★	★	★	★	★★★★★★★★★
Nelson (2017)[1]	★	★	★			★	★	★	★★★★★★
Raffield (2015)[12]		★	★		★		★		★★★
Takeuchi (2012)[30]	★	★	★	★		★	★	★	★★★★★★★
Zhou (2015)[31]	★	★	★	★		★	★	★	★★★★★★★

Notes case-control studies (i.e. Lian (2013), Takeuchi (2012), Zhou (2015)): categories of the quality assessment are displayed in bold, with interpretation of each item within the categories for this specific meta-analysis placed between brackets.
Selection: 1. Is the case definition adequate? (if yes, with independent validation (e.g. hospital records), one star; if yes, with record linkage (e.g. ICD-10 code or self-report) or no description, no star); 2. Representativeness of the cases (if consecutive or obviously representative series of cases, one star; if not consecutive or not (clearly) stated, no star); 3. Selection of controls (if community controls, one star; if hospital controls or no description, no star); 4. Definition of controls (if yes, with 'no history of CAD' explicitly stated, one star; if 'no history of CAD' not explicitly stated or no description, no star). **Comparability**: 1. Comparability of cases and controls on the basis of the design or analysis (if study adjusts for no covariates, two stars, if study adjusts for age and/or gender only, one star, if study adjusts for more covariates than age and/or gender, no star). **Exposure**: 1. Ascertainment of exposure (if secure record (e.g. genotyping), one star; if no description, no star); 2. Same method of ascertainment for cases and controls (if yes, one star; if no or no description, no star); 3. Non-response rate (if same rate for both groups, one star; if rate differs for both groups or no designation, no star).

Notes cohort studies (i.e. Nelson (2017), Raffield (2015)): categories of the quality assessment are displayed in bold, with interpretation of each item within the categories for this specific meta-analysis placed between brackets.
Selection: 1. Representativeness of the exposed cohort (if truly or somewhat representative of the average population, one star; if selected group (e.g. patients with type 2 diabetes) or no description, no star). 2. Selection of the non-exposed cohort (if drawn from the same community as the exposed cohort, one star; if drawn from a different source or no description, no star). 3. Ascertainment of exposure (if secure record (e.g. genotyping), one star; if no description, no star). 4. Demonstration that outcome of interest was not present at the start of the study (if explicitly stated, one star; if not explicitly stated, no star). **Comparability**: 1. Comparability of cases and controls on the basis of the design or analysis (if study adjusts for no covariates, two stars, if study adjusts for age and/or gender only, one star, if study adjusts for more covariates than age and/or gender, no star). **Outcome**: 1. Assessment of outcome (if independent blind assessment or record linkage (e.g. hospital records), one star; if with record linkage (e.g. ICD-10 code or self-report) or no description, no star). 2. Was follow-up long enough for outcome to occur (if average age of the sample population minus two standard deviations was equal to or more than 40 years, one star; if average age of the sample population minus two standard deviations was less than 40 years, no star). 3. Adequacy of follow up of cohorts (if explicitly stated why subjects from original cohort were excluded, one star; if not explicitly stated why subjects from original cohort were excluded, no star).

Tabble S8.5 Overview of the excluded eGFR and chronic kidney disease (CKD) studies with duplicate cohorts.

Excluded study	Study cohorts (with duplicate cohort in bold)	Study in which duplicate cohort is included
Gorski (2015)[32]	**Age, Gene/Environment susceptibility-Reykjavik Study (AGES)*, Amish Studies, The atherosclerosis risk in communities study (ARIC)*, Austrian stroke prevention study (ASPS), Cardiovascular health study (CHS),** The Cohort Lausannoise study (CoLaus), **Framingham Heart Study (FHS), Genetic epidemiology network of arteriopathy (GENOA), Health aging and body composition study (HABC), JUPITER, Cooperative research in the region of Augsburg (KORA), multi-ethnic study of atherosclerosis (MESA), Rotterdam Study, Study of health in Pomerania (SHIP), Three Cities (3C)**	Pattaro (2016)[33]
Köttgen (2009)[34]	Cohorts for heart and aging research in genomic epidemiology (CHARGE) which includes **ARIC***, **CHS, FHS, Rotterdam Study**	Pattaro (2016)[33]
Köttgen (2010)[35]**	**AGES*, Amish studies, ARIC*, ASPS, Baltimore longitudinal study of aging (BLSA), CHS, Erasmus rucphen family (ERF), FamHS*, FHS, GENOA,** Gutenburg Heart Study (GHS)**, KORA F3 and F4, Korcula Croatia, Microisolates in South tyrol study (MICROS), The northern Swedish population health study (NSPHS), Orkney complex disease study (ORCADES), Rotterdam Study, SHIP,** VIS **CROATIA, Women's genome health study (WGHS), HABC, Health professionals follow-up study (HPFS), Nurses health study (NHS), POPGen, The Sorbs study (Sorbs),** SPLIT**, Swiss study on air polluation and lung diseases in adults (SAPALDIA), Salzburg Atherosclerosis prevention program in subjects at high individual risk (SAPHIR)**	Pattaro (2016)[33]
Lanktree (2018)[36]	**Atherosclerotic Disease, Vascular function, & Genetic epidemiology study (ADVANCE)**, The academic medical center of Amsterdam Premature Atherosclerosis Cohort (AMC-PAS), **AMISH**, 1958 British Birth Cohort (BC58), D2D 2007 (D2D), deCODE, The diabetes Genetic Study (DIAGEN), The dietary, lifestyle and genetic determinants of obesity and metabolic syndrome study (DILGOM), The finnish diabetes prevention study (DPS), The dose responses to exercise training study (DR'S EXTRA), Edinburgh Artery Study (EAS), **Estonian genome center of university of Tartu (EGCUT)**, Ely, The european prospective investigation into cancer and nutrition (EPIC-CAD), Fenland, The Finnish cardiovascular study (FINCAVAS), Fragmin and fast revascularization during instability in coronary artery disease (FRISCII), FUSION2, Gene x lifestyle interactions and complex traits involved in elevated disease risk (GLACIER), **Genetics of diabetes and audit research Tayside (Go-DARTs),** Nord-Trondelag health study 2 (HUNT), IMPROVE, **KORA F3 AND F4,** Ludwigshafen Risk and Cardiovascular Health Study (LURIC), Malmo diet and cancer study (MDC), Metabolic syndrome in men (METSIM), Northern Finland birth cohort 1986 (NFBC1986), MRC national survey of health and development (NSHD), Prospective investigation of the vasculature in Uppsale seniors (PIVUS), Sardinia study on aging (SardiNIA), SCARFSHEEP, Swedish Twin Register (STR), The Hellenic study of interactions between SNPs and eating in atherosclerosis susceptibility (THISEAS),	Pattaro (2016)[33] Deshmukh (2013)[37]

Table S8.5 (continued)

Excluded study	Study cohorts (with duplicate cohort in bold)	Study in which duplicate cohort is included
Lanktree (2018)[36]	TROMSO, Uppsala longitudinal study of adult men (ULSAM), Whitehall II, Cebu longitudinal health and nutrition survey (CLHNS), Taiwan metabochip consortium (TAICHI), Asian Indian diabetic heart study/Sikh diabetes study (AIDHS/SDS), The Pakistan risk of myocardial infarction study (PROMIS), Family blood pressure project GenNet and HyperGen studies (FBPP), Kingston GXE (GXE), General population cohort study, Uganda (MRC/UVRI GPC), Seychelles tandem study (SEY), Spanishtown (SPT), **AGES**, **ARIC**, TWIN COHORTS, **BLSA**, **CHS**, CoLaus, Invecchiare in Chianti study (InCHIANTI), London life sciences prospective population study (LOLIPOP), National FINRISK Study, PARC, **Rotterdam Study**, Supplementation en vitamins et mineraux antioxydants study (SUMIVAX), **WGHS**, British genetics of hypertension study (BRIGHT), Britisch 1958 birth cohort type 1 diabetes genetics consortium (B58CT1DGC), Diabetes genetics initiative (DGI), FHS, HEATLH2000 GenMets Study, MedStar, PennCATH, **ERF**, Framingham Heart Study (FramHS), **MICROS**, **NSPHS**, **ORCADES**, Vis Study, National FINRISK 1997 study (FINRISK97), Coronary artery disease genomewide replication and meta-analysis study (CARDIoGRAM), COROGENE	Pattaro (2016)[33] Deshmukh (2013)[37]
Li (2017)[38]	**AGES, Amish, ARIC, CHS, CROATIA-Korcula, EGCUT, FamHS, FHS,** Geisinger Genomic Medicine Exome project (GeMEP), Generation Scotland: Scottish family health study (GS:SFHS), **HPFS**, Health and retirement study (HRS), **European network for genetic-epidemiological studies (HYPERGENES), Italian network on genetic isolates (INGI),** Mount Sinai BioMe Biobank program (IPM), **KORA F4**, METSIM, **NHS, Rotterdam Study, SAPALDIA, SHIP, WGHS**, Women's health initiative (WHI), **Young Finns Study (YFS)**, **ARIC**, **CHS**, **GENOA**, Jackson Heart Study (JHS)	Pattaro (2016)[33]
Pattaro (2012)[39]	**AGES*, AMISCH,** SPS**, ARIC*, BLSA, CHS, ERF, FamHS*, FHS, GENOA, HABC, HPFS, KORA F3 and F4, Croatia-Korcula cohort, MICROS, NSPHS, NHS, ORCADES, POPGen, Sorbs, Rotterdam study, SHIP, VIS, WGHS**	Pattaro (2016)[33]
Thio (2017)[40]	**Prevention of renal and vascular endstage disease (PREVEND)**	Pattaro (2016)[33]

*Cohorts AGES, ARIC, FamHS, FHS, PROSPER, and Rotterdam Study are part of the CHARGE consortium
**This study was included for the data-synthesis on cystatin C-based eGFR, since this was not reported by Pattaro (2016)[33].

Table S8.6 Quality assessment of the eGFR and chronic kidney disease (CKD) studies based on the NOS.

References	Selection				Comparability	Exposure/Outcome			Quality judgment
	1	2	3	4	1	1	2	3	
Bonetti (2011)[41]		★	★			★	★		★★★★
Deshmukh (2013)[37]		★	★			★	★		★★★★
Hishida (2014)[42]	★	★	★		★	★		★	★★★★★★
Köttgen (2010)[35]	★	★	★		★	★	★		★★★★★★
Okada (2012)[43]	★	★	★			★	★		★★★★★
Pattaro (2016)[33]	★	★	★		★	★			★★★★★
Sveinbjornsson (2014)[44]	★			★	★	★	★	★	★★★★★★
Yamada (2013)[45]	★	★		★	★	★	★	★	★★★★★★★

Notes case-control studies (i.e. Sveinbjornsson (2014), Yamada (2013)): categories of the quality assessment are displayed in bold, with interpretation of each item within the categories for this specific meta-analysis placed between brackets.

Selection: 1. Is the case definition adequate? (if yes, with independent validation (e.g. hospital records), one star; if yes, with record linkage (e.g. ICD-10 code or self-report) or no description, no star); 2. Representativeness of the cases (if consecutive or obviously representative series of cases, one star; if not consecutive or not (clearly) stated, no star); 3. Selection of controls (if community controls, one star; if hospital controls or no description, no star); 4. Definition of controls (if yes, with 'no history of CKD' explicitly stated, one star; if 'no history of CKD' not explicitly stated or no description, no star). **Comparability**: 1. Comparability of cases and controls on the basis of the design or analysis (if study adjusts for no covariates, two stars, if study adjusts for age and/or gender only, one star, if study adjusts for more covariates than age and/or gender, no star). **Exposure**: 1. Ascertainment of exposure (if secure record (e.g. genotyping), one star; if no description, no star); 2. Same method of ascertainment for cases and controls (if yes, one star; if no or no description, no star); 3. Non-response rate (if same rate for both groups, one star; if rate differs for both groups or no designation, no star).

Notes cohort studies (i.e. Bonetti (2011), Deshmukh (2013), Hishida (2014), Köttgen (2010), Okado (2012), Pattaro (2016)): categories of the quality assessment are displayed in bold, with interpretation of each item within the categories for this specific meta-analysis placed between brackets.

Selection: 1. Representativeness of the exposed cohort (if truly or somewhat representative of the average population, one star; if selected group (e.g. patients with type 2 diabetes) or no description, no star). 2. Selection of the non-exposed cohort (if drawn from the same community as the exposed cohort, one star; if drawn from a different source or no description, no star). 3. Ascertainment of exposure (if secure record (e.g. genotyping), one star; if no description, no star). 4. Demonstration that outcome of interest was not present at the start of the study (if dichotomous variable (e.g. CKD yes/no) and 'no history of CKD' explicitly stated, one star; if dichotomous variable (e.g. CKD yes/no) and 'no history of CKD' not explicitly stated, or continuous variable (e.g. eGFR), no star). **Comparability**: 1. Comparability of cases and controls on the basis of the design or analysis (if study adjusts for no covariates, two stars, if study adjusts for age and/or gender only, one star, if study adjusts for more covariates than age and/or gender, no star). **Outcome**: 1. Assessment of outcome (if independent blind assessment or record linkage (e.g. hospital records), one star; if with record linkage (e.g. ICD-10 code or self-report) or no description, no star). 2. Was follow-up long enough for outcome to occur (if average age of the sample population minus two standard deviations was equal to or more than 40 years or variable was continuous (e.g. eGFR), one star; if average age of the sample population minus two standard deviations was less than 40 years, no star). 3. Adequacy of follow up of cohorts (if explicitly stated why subjects from original cohort were excluded, one star; if not explicitly stated why subjects from original cohort were excluded, no star).

References

1. Nelson C.P., Goel A., Butterworth A.S., et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet.* 2017;49(9):1385-1391.
2. Zhao B., Lu Q., Cheng Y., et al. A Genome-wide association study to identify single-nucleotide polymorphisms for acute kidney injury. *American Journal of Respiratory and Critical Care Medicine.* 2017;195(4):482-490.
3. Schunkert H., Konig I.R., Kathiresan S., et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet.* 2011;43(4):333-338.
4. Deloukas P., Kanoni S., Willenborg C., et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet.* 2013;45(1):25-33.
5. Varbo A., Benn M., Tybjaerg-Hansen A., Grande P., Nordestgaard B.G. TRIB1 and GCKR polymorphisms, lipid levels, and risk of ischemic heart disease in the general population. *Arteriosclerosis, Thrombosis, and Vascular Biology.* 2011;31(2):451-457.
6. Angelakopoulou A., Shah T., Sofat R., et al. Comparative analysis of genome-wide association studies signals for lipids, diabetes, and coronary heart disease: Cardiovascular Biomarker Genetics Collaboration. *European Heart Journal.* 2012;33(3):393-407.
7. Bi M., Kao W.H., Boerwinkle E., et al. Association of rs780094 in GCKR with metabolic traits and incident diabetes and cardiovascular disease: the ARIC Study. *PLoS One.* 2010;5(7):e11690.
8. The Wellcome Trust Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007;447:661.
9. Davies R.W., Wells G.A., Stewart A.F., et al. A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. *Circulation Cardiovascular genetics.* 2012;5(2):217-225.
10. Dehghan A., Bis J.C., White C.C., et al. Genome-Wide Association Study for Incident Myocardial Infarction and Coronary Heart Disease in Prospective Cohort Studies: The CHARGE Consortium. *PLOS ONE.* 2016;11(3):e0144997.
11. Divers J., Palmer N.D., Langefeld C.D., et al. Genome-wide association study of coronary artery calcified atherosclerotic plaque in African Americans with type 2 diabetes. *BMC Genetics.* 2017;18 (1) (no pagination)(105).
12. Raffield L.M., Cox A.J., Carr J.J., et al. Analysis of a cardiovascular disease genetic risk score in the Diabetes Heart Study. *Acta Diabetol.* 2015;52(4):743-751.
13. Erdmann J., Willenborg C., Nahrstaedt J., et al. Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. *European heart journal.* 2011;32(2):158-168.
14. Howson J.M.M., Zhao W., Barnes D.R., et al. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. *Nature Genetics.* 2017;49(7):1113-1119.
15. Kozian D.H., Barthel A., Cousin E., et al. Glucokinase-activating GCKR polymorphisms increase plasma levels of triglycerides and free fatty acids, but do not elevate cardiovascular risk in the Ludwigshafen Risk and Cardiovascular Health Study. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme.* 2010;42(7):502-506.
16. Sevastianova K., Santos A., Kotronen A., et al. Effect of short-term carbohydrate overfeeding and long-term weight loss on liver fat in overweight humans. *Am J Clin Nutr.* 2012;96(4):727-734.
17. Lettre G., Palmer C.D., Young T., et al. Genome-Wide Association Study of Coronary Heart Disease and Its Risk Factors in 8,090 African Americans: The NHLBI CARe Project. *PLOS Genetics.* 2011;7(2):e1001300.
18. Nikpay M., Goel A., Won H.H., et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet.* 2015;47(10):1121-1130.
19. Speliotes E.K., Yerges-Armstrong L.M., Wu J., et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet.* 2011;7(3):e1001324.
20. Pfister R., Barnes D., Luben R.N., Khaw K.T., Wareham N.J., Langenberg C. Individual and cumulative effect of type 2 diabetes genetic susceptibility variants on risk of coronary heart disease. *Diabetologia.* 2011;54(9):2283-2287.

21. Stark K., Reinhard W., Grassl M., et al. Common Polymorphisms Influencing Serum Uric Acid Levels Contribute to Susceptibility to Gout, but Not to Coronary Artery Disease. *PLoS ONE.* 2009;4(11):e7729.
22. The I.B.C.K.C.A.D.C. Large-Scale Gene-Centric Analysis Identifies Novel Variants for Coronary Artery Disease. *PLOS Genetics.* 2011;7(9):e1002260.
23. Tregouet D.A., Konig I.R., Erdmann J., et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nature genetics.* 2009;41(3):283-285.
24. Wang F., Xu C.Q., He Q., et al. Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. *Nature genetics.* 2011;43(4):345-349.
25. Lian J., Guo J., Chen Z., et al. Positive association between GCKR rs780093 polymorphism and coronary heart disease in the aged Han Chinese. *Dis Markers.* 2013;35(6):863-868.
26. Webb T.R., Erdmann J., Stirrups K.E., et al. Systematic Evaluation of Pleiotropy Identifies 6 Further Loci Associated With Coronary Artery Disease. *Journal of the American College of Cardiology.* 2017;69(7):823-836.
27. Wild P.S., Zeller T., Schillert A., et al. A genome-wide association study identifies LIPA as a susceptibility gene for coronary artery disease. *Circulation Cardiovascular genetics.* 2011;4(4):403-412.
28. Willer C.J., Sanna S., Jackson A.U., et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet.* 2008;40(2):161-169.
29. Yaghootkar H., Scott R.A., White C.C., et al. Genetic evidence for a normal-weight "metabolically obese" phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2 diabetes. *Diabetes.* 2014;63(12):4369-4377.
30. Takeuchi F., Isono M., Katsuya T., et al. Association of genetic variants influencing lipid levels with coronary artery disease in Japanese individuals. *PLoS One.* 2012;7(9):e46385.
31. Zhou Y.J., Hong S.C., Yin R.X., Yang Q., Cao X.L., Chen W.X. Polymorphisms in the GCKR are associated with serum lipid traits, the risk of coronary artery disease and ischemic stroke. *Int J Clin Exp Med.* 2015;8(7):10678-10686.
32. Gorski M., Tin A., Garnaas M., et al. Genome-wide association study of kidney function decline in individuals of European descent. *Kidney international.* 2015;87(5):1017-1029.
33. Pattaro C., Teumer A., Gorski M., et al. Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat Commun.* 2016;7:10023.
34. Köttgen A., Glazer N.L., Dehghan A., et al. Multiple Novel Loci are Associated with Indices of Renal Function and Chronic Kidney Disease. *Nature genetics.* 2009;41(6):712-717.
35. Kottgen A., Pattaro C., Boger C.A., et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet.* 2010;42(5):376-384.
36. Lanktree M.B., Theriault S., Walsh M., Pare G. HDL Cholesterol, LDL Cholesterol, and Triglycerides as Risk Factors for CKD: A Mendelian Randomization Study. *American journal of kidney diseases : the official journal of the National Kidney Foundation.* 2018;71(2):166-172.
37. Deshmukh H.A., Palmer C.N., Morris A.D., Colhoun H.M. Investigation of known estimated glomerular filtration rate loci in patients with type 2 diabetes. *Diabet Med.* 2013;30(10):1230-1235.
38. Li M., Li Y., Weeks O., et al. SOS2 and ACP1 Loci Identified through Large-Scale Exome Chip Analysis Regulate Kidney Development and Function. *Journal of the American Society of Nephrology : JASN.* 2017;28(3):981-994.
39. Pattaro C., Köttgen A., Teumer A., et al. Genome-Wide Association and Functional Follow-Up Reveals New Loci for Kidney Function. *PLOS Genetics.* 2012;8(3):e1002584.
40. Thio C.H.L., van der Most P.J., Nolte I.M., et al. Evaluation of a genetic risk score based on creatinine-estimated glomerular filtration rate and its association with kidney outcomes. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association.* 2017.
41. Bonetti S., Trombetta M., Boselli M.L., et al. Variants of GCKR affect both beta-cell and kidney function in patients with newly diagnosed type 2 diabetes: the Verona newly diagnosed type 2 diabetes study 2. *Diabetes Care.* 2011;34(5):1205-1210.
42. Hishida A., Takashima N., Turin T.C., et al. GCK, GCKR polymorphisms and risk of chronic kidney disease in Japanese individuals: data from the J-MICC Study. *J Nephrol.* 2014;27(2):143-149.

43. Okada Y., Sim X., Go M.J., et al. Meta-analysis identifies multiple loci associated with kidney function-related traits in east Asian populations. *Nat Genet.* 2012;44(8):904-909.
44. Sveinbjornsson G., Mikaelsdottir E., Palsson R., et al. Rare mutations associating with serum creatinine and chronic kidney disease. *Hum Mol Genet.* 2014;23(25):6935-6943.
45. Yamada Y., Nishida T., Ichihara S., et al. Identification of chromosome 3q28 and ALPK1 as susceptibility loci for chronic kidney disease in Japanese individuals by a genome-wide association study. *J Med Genet.* 2013;50(6):410-418.

9

CHAPTER NINE

Association between de novo lipogenesis susceptibility genes and coronary artery disease

Pomme I.H.G. Simons, Olivier Valkenburg,
Coen D.A. Stehouwer, Martijn C.G.J. Brouwers

Submitted

Summary

Aim: Coronary artery disease is the principal cause of death in individuals with non-alcoholic fatty liver disease (NAFLD). The aim of this study was to assess the causal effect of de novo lipogenesis, one of the major pathways leading to NAFLD, on coronary artery disease risk.

Methods: De novo lipogenesis susceptibility genes were used as instruments and selected using three approaches: 1) genes that are associated with both high serum triglycerides and low sex hormone-binding globulin, both downstream consequences of de novo lipogenesis (unbiased approach), 2) genes that have a known role in de novo lipogenesis (biased approach), and 3) genes that have been associated with serum fatty acids, used as a proxy of de novo lipogenesis. Gene-coronary artery disease effect estimates were retrieved from the meta-analysis of CARDIoGRAM and the UK Biobank (~76,014 cases and ~264,785 controls). Effect estimates were clustered using a fixed-effects meta-analysis.

Results: Twenty-two de novo lipogenesis susceptibility genes were identified by the unbiased approach, nine genes by the biased approach and seven genes were associated with plasma fatty acids. Clustering of genes selected in the unbiased and biased approach showed a statistically significant association with coronary artery disease (OR: 1.016, 95% CI: 1.012;1.020 and OR: 1.013, 95% CI: 1.007;1.020, respectively), while clustering of fatty acid genes did not (OR: 1.004, 95% CI: 0.996;1.011). Subsequent exclusion of potential influential outliers did reveal a statistically significant association (OR: 1.009, 95% CI: 1.000;1.018).

Conclusion: De novo lipogenesis susceptibility genes are associated with an increased risk of coronary artery disease. These findings suggest that de novo lipogenesis is causally involved in the pathogenesis of coronary artery disease and favor further development of strategies that target NAFLD through de novo lipogenesis.

Introduction

Non-alcoholic fatty liver disease (NAFLD) has become a substantial health burden that is associated with hepatic complications including end-stage liver failure and hepatocellular carcinoma[1]. Notably, NAFLD is also strongly associated with extrahepatic complications such as type 2 diabetes and coronary artery disease, the latter of which has become the leading cause of death in patients with NAFLD[2].

Over the past years, there has been on ongoing discussion on the causal role of NAFLD in the development of coronary artery disease. The central, systemic role of the liver in metabolic processes and the need for long-term follow-up complicates conventional epidemiological and intervention studies that target NAFLD[3].

Genetic epidemiology can serve as an alternative method to infer causality. As genetic variants that predispose to or protect from an exposure of interest (such as NAFLD) are randomly distributed at conception, they can be used as an instrument to study the causal effect of the exposure on the outcome (such as coronary artery disease)[4]. We previously applied this approach and showed that genetic variants that result in NAFLD through impaired secretion of very-low-density lipoproteins (VLDL) protect from coronary artery disease[5]. However, it remains uncertain how other more principal pathways that result in NAFLD – in particular de novo lipogenesis[6] – contribute to the risk of coronary artery disease.

In the present study we, therefore, aimed to use genetic epidemiology to gain more insight into the causal effect of de novo lipogenesis on coronary artery disease.

Methods

Selection of de novo lipogenesis susceptibility genes

Unbiased approach

Since de novo lipogenesis inhibits the synthesis of SHBG and stimulates VLDL production[7-10], we assumed that genetic variants that predispose to both low serum SHBG *and* high triglyceride levels are likely de novo lipogenesis susceptibility genes. The susceptibility genes were retrieved from genome-wide association (GWA) studies in the Global Lipids Genetics Consortium (serum triglycerides) and UK Biobank (BMI-adjusted SHBG)[11,12]. Genetic variants were included if the associations with SHBG and triglycerides reached genome-wide significance ($p<5*10^{-8}$), and the effect allele was positively associated with serum triglyceride levels and inversely associated with serum

SHBG levels. Genetic variants were excluded if they were in linkage disequilibrium (r^2>0.1, the variant with the largest absolute effect estimate was retained).

Biased approach

In the biased approach, we screened all genome-wide significant SHBG susceptibility genes in the UK Biobank for a potential involvement in de novo lipogenesis (based on genecards.org). Genes, and their corresponding genetic variants, that were of interest were further explored in existing literature to verify their role in de novo lipogenesis. The effect allele of the genetic variant was chosen as the allele that decreases serum SHBG levels. Genetic variants were excluded if they were in linkage disequilibrium, as described previously.

Fatty acid approach

In the fatty acid approach, we selected genetic variants that have previously been identified from a GWA study for palmitic acid (16:0), stearic acid (18:0), palmitoleic acid (16:ln-7), or oleic acid (18:ln-9), used as biomarkers of de novo lipogenesis[13]. The effect allele was chosen as the allele that increases plasma fatty acids. If a genetic variant had effects on multiple plasma fatty acids, the effect allele was chosen as the allele that increases the concentration of the fatty acid that is most proximal in the pathway of de novo lipogenesis (see Supplementary Figure S9.1). Genetic variants were excluded if they were in linkage disequilibrium, as described previously.

Associations with coronary artery disease

Summary-level data for the association of the selected genetic variants with coronary artery disease was retrieved from the publicly available data of the CARDIoGRAMplusC4D 1000 Genomes-based GWAS, Myocardial Infarction Genetics and CARDIoGRAM Exome chip, and UK Biobank SOFT CAD study[14]. This dataset includes ~76,014 cases and ~264,785 controls. Coronary artery disease was defined as a history of fatal or nonfatal myocardial infarction, percutaneous transluminal coronary angioplasty (PTCA) or coronary artery bypass grafting (CABG), chronic ischemic heart disease (IHD) or angina, based on self-reported data or hospital records.

Statistical analyses

For each approach, a fixed-effect meta-analysis was conducted to combine the gene-coronary artery disease effect estimates. The overall effect estimate should, therefore, be interpreted as the average coronary artery disease risk conferred by one de novo lipogenesis risk allele[5]. Higgin's I^2 and Cochran's Q statistic were calculated to identify heterogeneity of the effect estimates. Potential influential outliers were identified statistically using the leave-one-out method[15]. Results were considered statistically

significant at p<0.05. All analyses were conducted with the R statistical software (R Developmental Core Team) using the *metaphor* package[16].

Results

Unbiased approach

Thirty-one genes that reached genome-wide significance ($p<5*10^{8}$) for the association with both serum SHBG *and* triglycerides were identified. Nine genes were subsequently excluded because of linkage disequilibrium (*SNX17*), misalignment of the predefined direction of the association with serum triglycerides and SHBG (*AKR1C4*, *APOC1*, *APOC1P1*, and *MET*), or absence in the outcome dataset (*GATAD2A*, *HSD17B13*, *MACF1*, and *NRBF2*). Therefore, twenty-two genes that predisposed to low serum SHBG and high triglyceride levels were included in the final analysis (Supplementary Table S9.1). Clustering of these genetic variants resulted in a statistically significant association with coronary artery disease (OR: 1.016, 95% CI: 1.012;1.020, I^2: 72.7%, Q: 76.9) (Figure 9.1). As the I^2 statistic indicated significant heterogeneity, the analysis was repeated after exclusion of the most influential outliers (*JMJD1C*, *MYRF*, and *TRIB1*). Exclusion of these genes reduced the heterogeneity, and did not affect the strength of the association (OR: 1.020, 95% CI: 1.015;1.024, I^2: 41.8%, Q: 30.9).

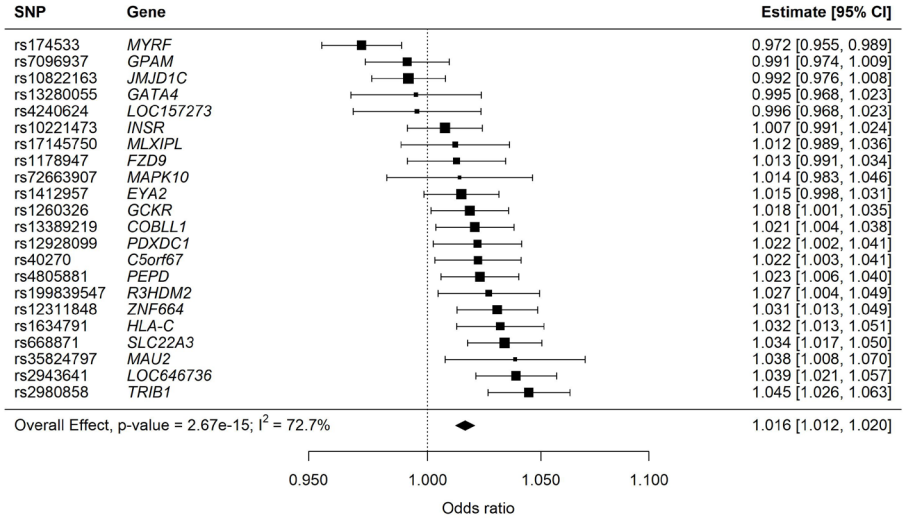

Figure 9.1 Association between de novo lipogenesis susceptibility genes, identified by an unbiased approach, and coronary artery disease. The overall effect estimate represents the average risk of coronary artery disease conferred by one de novo lipogenesis risk allele.

Biased approach

Ten SHBG susceptibility genes were identified that are known to be involved in de novo lipogenesis. One gene (*IRS1*) was excluded as it was unavailable in the outcome dataset. The remaining nine genes (*GCK*, *GCKR*, *GPAM*, *INSR*, *MLXIPL*, *PNPLA3*, *PTEN*, *SCAP*, and *TRIB1*) were included in the analysis (Supplementary Table S9.2). Their putative role in de novo lipogenesis is shown in Supplementary Figure S9.1. Clustering of these genetic variants resulted in a statistically significant association with coronary artery disease (OR: 1.013, 95% CI: 1.007;1.020, I^2: 74.3%, Q: 31.1) (Figure 9.2).

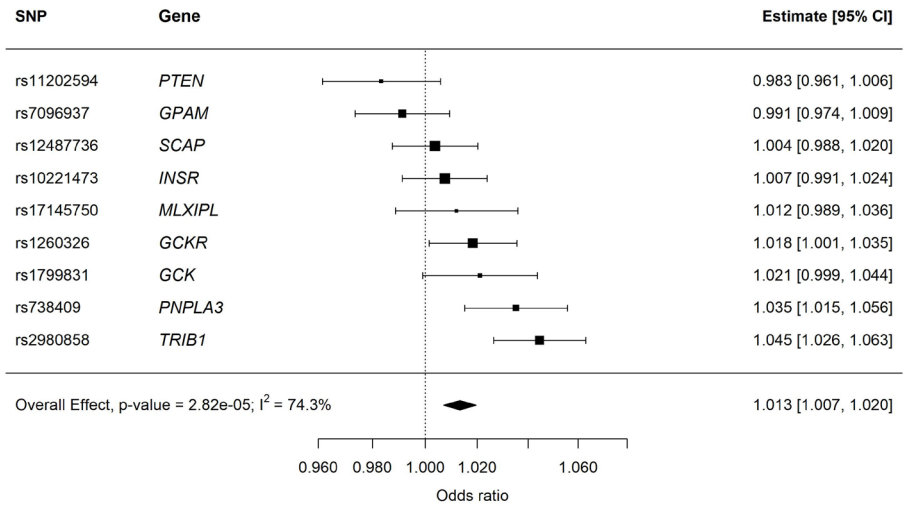

Figure 9.2 Association between de novo lipogenesis susceptibility genes, identified by a biased approach, and coronary artery disease. Overall effect estimate represents the average risk of coronary artery disease conferred by one de novo lipogenesis risk allele.

Since the *PNPLA3* major allele, which according to the selection criteria was associated with higher rates of de novo lipogenesis and lower serum SHBG levels (Supplementary Table S9.2), is also associated with a higher VLDL secretion rate and a lower intrahepatic lipid content[17,18], the analysis was repeated after exclusion of this variant. This did not affect the strength of the association (OR: 1.011, 95% CI: 1.004;1.018, I^2: 73.2%, Q: 26.1). Furthermore, we repeated the analysis after exclusion of influential outliers (*GPAM*, *PNPLA3*, *PTEN*, and *TRIB1*). The strength of the association remained materially unchanged while the heterogeneity was reduced (OR: 1.012, 95% CI: 1.003;1.020, I^2: 0.0%, Q: 2.4).

Fatty acid approach

Of the eight fatty acid susceptibility genes that were previously identified in a GWA study[13], one gene (*ALG14*) was excluded from the current analysis because of linkage disequilibrium. The remaining genes are presented in Supplementary Table S9.3. Clustering of these genetic variants did not result in a statistically significant association with coronary artery disease (OR: 1.004, 95% CI: 0.996;1.011, I^2: 74.3%, Q: 24.0) (Figure 9.3). After exclusion of the influential outliers (*GCKR* and *PKD2L1*), and a consequent reduction in heterogeneity, the strength of the association increased and reached statistical significance (OR: 1.009, 95% CI: 1.000;1.018, I^2: 0.0%, Q: 3.4).

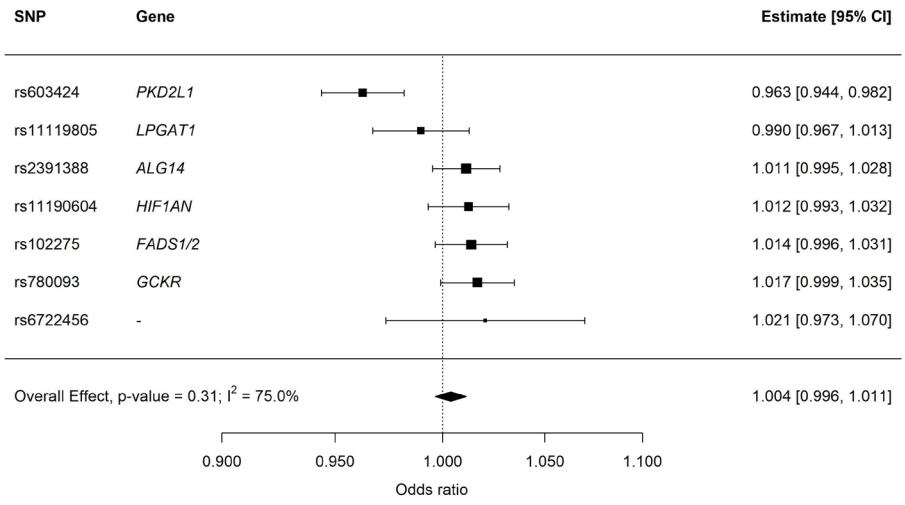

Figure 9.3 Association between fatty acid susceptibility genes and coronary artery disease. Overall effect estimate represents the average risk of coronary artery disease conferred by one fatty acid risk allele.

9

Discussion

The aim of the present study was to assess the association between de novo lipogenesis susceptibility genes and coronary artery disease. De novo lipogenesis susceptibility genes were identified using an unbiased and biased selection approach, as well as by using fatty acid susceptibility genes as a proxy for de novo lipogenesis. Clustering of these genes revealed a statistically significant association between de novo lipogenesis susceptibility genes, but not fatty acid genes, with coronary artery disease.

Both experimental and observational studies have shown that an increase in de novo lipogenesis results in an increase in VLDL secretion as well as a reduction in serum SHBG levels[7,8,10]. We, therefore, assumed that the overlap in the triglyceride and SHBG susceptibility genes likely represent genes that also predispose to de novo lipogenesis. As triglycerides are a well-known risk factor for cardiovascular disease, it is perhaps unsurprising that the genes identified in this approach predispose to coronary artery disease[19-21]. Serum triglycerides are, therefore, a likely mediator in this association. Of interest, we previously showed that the direction of the association between NAFLD susceptibility genes and coronary artery disease depends on their effect on serum lipids[5], which further corroborates the mediation effect of serum lipids on coronary artery disease risk. As a limitation of this unbiased approach, we cannot exclude that other processes may also predispose to both serum SHBG and triglyceride levels, in particular upstream factors of de novo lipogenesis such as obesity[22]. To address this, we used BMI-adjusted SHBG susceptibility genes, which meant that well-known obesity genes, such as *FTO*[23], were not identified. Nevertheless, despite our efforts to reduce the effect of obesity, we cannot exclude residual confounding.

To overcome some of the limitations of the unbiased approach, we also selected genes based on their involvement in de novo lipogenesis. We identified nine genes that are known to regulate the process of de novo lipogenesis, including *MLXIPL* and *GCKR*[24,25]. Nonetheless, careful evaluation of the genes reveals the absence of genes that encode several other well-known lipogenic enzymes, such as acetyl-coenzyme A carboxylase (*ACC*) and fatty acid synthase (*FASN*)[26,27]. As these genes did not predispose to serum SHBG level – a selection criteria we enforced to ensure that the genetic variants are associated with downstream consequences of de novo lipogenesis and, hence, are likely to be functional variants (or in linkage disequilibrium with a variant that is) – they could not be included. Furthermore, there was a notable discrepancy between the de novo lipogenesis susceptibility genes identified by the biased and unbiased approach. This was somewhat surprising, as we had anticipated that all genes involved in de novo lipogenesis (biased approach) would also be both SHBG and triglyceride susceptibility genes (unbiased approach). This discrepancy could be the result of the very stringent significant p-value threshold applied for the selection of both SHBG and triglyceride susceptibility genes ($p<5*10^{-8}$).

In the present study, we did not observe an association between fatty acid susceptibility genes and coronary artery disease, although exclusion of potential outliers did reveal a statistically significant association. Previous observational studies that used fatty acids as a proxy for de novo lipogenesis found inconclusive associations with cardiovascular disease[28-30]. The use of fatty acids as a proxy of de novo lipogenesis has, more recently, been scrutinized by stable isotope studies. De novo lipogenesis associated only weakly, though significantly, with palmitic acid (16:0) and stearic acid

(18:0), the direct products of de novo lipogenesis, but not with its derivatives, such as palmitoleic acid (16:ln-7) or oleic acid (18:ln-9)[31]. The validity of these fatty acids as a serum biomarker of de novo lipogenesis can, therefore, be questioned.

Despite it being marked as a statistical outlier in the fatty acid approach, *GCKR* may be one of the most valid de novo lipogenesis susceptibility genes included in the present study. First, it is the only gene that was identified as a de novo lipogenesis susceptibility gene in all three approaches. Second, there is ample biological plausibility that variants in *GCKR* affect de novo lipogenesis. The minor allele in *GCKR* encodes a variant of glucokinase regulatory protein (GKRP), a liver-specific protein, which binds glucokinase less effectively[32,33]. Thereby, it increases the hepatic influx of glucose resulting in higher availability of substrate for de novo lipogenesis (Supplementary Figure S9.1)[34]. Indeed, stable isotope studies have shown that individuals carrying the minor allele of *GCKR* have higher rates of de novo lipogenesis[25]. The statistically significant association of the *GCKR* minor allele with coronary artery disease in this study and our previous meta-analysis[35], therefore, further establishes a causal role for de novo lipogenesis in the pathogenesis of coronary artery disease.

The findings in this study may provide a glimpse into the long-term consequences of therapies that affect de novo lipogenesis. On the one hand, therapies that reduce de novo lipogenesis, such as ACC inhibitors which are currently undergoing phase II trials as a potential treatment for NAFLD[36], may in the long-term also have beneficial cardiovascular effects. This beneficial side-effect is desirable as cardiovascular disease is the principal cause of death in individuals with NAFLD[2]. As previously indicated, in this study we were unable to assess the effects of genetic variants in ACC specifically, although we did study upstream variants, including *MLXIPL*. On the other hand, the present findings also indicate that therapies that stimulate de novo lipogenesis should be avoided. Currently, compounds that augment hepatic glucose uptake, such as liver-specific glucokinase activators and disruptors of the GKRP-glucokinase complex, are under investigation as a new class of glucose-lowering medication[37,38]. These drugs have biological analogies with variants in *GCK* and *GCKR* and may, therefore, stimulate de novo lipogenesis and, hence, cause NAFLD and coronary artery disease[39].

This study has several strengths. Given the absence of GWA studies for de novo lipogenesis, we used three independent methods to identify de novo lipogenesis susceptibility genes, which allowed us to test the robustness of our findings. Furthermore, by retrieving gene-coronary artery disease effect estimates from the CARDIoGRAM and UK Biobank dataset, which includes more than 340,000 individuals, we had sufficient statistical power to assess the relationship between de novo lipogenesis susceptibility genes and coronary artery disease. Finally, as indicated, the current results can shed light on the possible long-term consequences of drug

therapies that affect de novo lipogenesis, a finding which would otherwise require years of follow-up in conventional research.

In addition to the hitherto described considerations, this study has several additional limitations. First, a primary assumption of all Mendelian randomization studies is that the instrumental genes do not affect the outcome, other than through the exposure, i.e. there should be no horizontal pleiotropy[40]. *PNPLA3* illustrates this risk of horizontal pleiotropy. The major allele of *PNPLA3*, which was included in the biased approach based on its association with lower serum SHBG levels and higher rates of de novo lipogenesis[18], is also associated with higher VLDL secretion[17,18]. The latter, which is thought to be the primary effect of *PNPLA3*, is also known to be a risk factor for coronary artery disease[5]. Exclusion of *PNPLA3* in the current analyses did not, however, affect the strength of the associations. Likewise, pleiotropic effects may also explain why two variants (i.e. *MYRF* and *PKD2L1*) were found to be statistically significantly protective for coronary artery disease, which was in direct contrast with the average effect of de novo lipogenesis genes found in this study. Second, as gene-exposure data, i.e. gene-de novo lipogenesis data, was unavailable, we were unable to conduct full Mendelian randomization analyses. Consequently, it is not possible to quantify the effect size that de novo lipogenesis may have on coronary artery disease. The results of the current study should, therefore, be interpreted as the average risk of coronary artery disease conferred by one de novo lipogenesis genetic variant, which explains the observed small effect sizes. If gene-de novo lipogenesis data becomes available in the future – which is not likely given the laborious nature of quantifying de novo lipogenesis – the current study should be repeated to draw conclusions on the extent to which de novo lipogenesis contributes to the risk of coronary artery disease.

In summary, de novo lipogenesis susceptibility genes, but not fatty acid susceptibility genes, are associated with coronary artery disease. These findings enhance our understanding of the contribution of different pathways of intrahepatic lipid accumulation in the risk of cardiovascular disease, and suggest that augmented de novo lipogenesis may have negative consequences on the risk of coronary artery disease. The current findings justify further studies of the long-term consequences of therapies targeting de novo lipogenesis as a means to not only treat NAFLD, but also to reduce the risk of extrahepatic complications of NAFLD, such as coronary artery disease.

References

1. Mantovani A, Scorletti E, Mosca A, et al. Complications, morbidity and mortality of nonalcoholic fatty liver disease. *Metabolism.* 2020;111:154170.
2. Ong JP, Pitts A, Younossi ZM. Increased overall mortality and liver-related mortality in non-alcoholic fatty liver disease. *Journal of Hepatology.* 2008;49(4):608-612.
3. Brouwers MCGJ, Simons N, Stehouwer CDA, Isaacs A. Non-alcoholic fatty liver disease and cardiovascular disease: assessing the evidence for causality. *Diabetologia.* 2020;63(2):253-260.
4. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics.* 2014;23(R1):R89-R98.
5. Brouwers MCGJ, Simons N, Stehouwer CDA, et al. Relationship between nonalcoholic fatty liver disease susceptibility genes and coronary artery disease. *Hepatology Communications.* 2019;3(4):587-596.
6. Donnelly KL, Smith CI, Schwarzenberg SJ, et al. Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *Journal of Clinical Investigation.* 2005;115(5):1343-1351.
7. Diraison F, Moulin P, Beylot M. Contribution of hepatic de novo lipogenesis and reesterification of plasma non esterified fatty acids to plasma triglyceride synthesis during non-alcoholic fatty liver disease. *Diabetes & Metabolism Journal.* 2003;29(5):478-485.
8. Selva DM, Hogeveen KN, Innis SM, Hammond GL. Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. *Journal of Clinical Investigation.* 2007;117(12):3979-3987.
9. Simons PIHG, Valkenburg O, Stehouwer CDA, Brouwers MCGJ. Sex hormone-binding globulin: biomarker and hepatokine? *Trends in Endocrinology and Metabolism.* 2021;32(8):544-553.
10. Simons PIHG, Valkenburg O, Telgenkamp I, et al. Relationship between de novo lipogenesis and serum sex hormone binding globulin in humans. *Clinical Endocrinology.* 2021;95(1):101-106.
11. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nature Genetics.* 2013;45(11):1274-1283.
12. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nature Medicine.* 2020;26(2):252-258.
13. Wu JH, Lemaitre RN, Manichaikul A, et al. Genome-wide association study identifies novel loci associated with concentrations of four plasma phospholipid fatty acids in the de novo lipogenesis pathway: results from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. *Circulation: Cardiovascular Genetics.* 2013;6(2):171-183.
14. Nelson CP, Goel A, Butterworth AS, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nature Genetics.* 2017;49(9):1385-1391.
15. Viechtbauer W, Cheung MW-L. Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods.* 2010;1(2):112-125.
16. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software.* 2010;36(3):1-48.
17. Pirazzi C, Adiels M, Burza MA, et al. Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro. *Journal of Hepatology.* 2012;57(6):1276-1282.
18. Mancina RM, Matikainen N, Maglio C, et al. Paradoxical dissociation between hepatic fat content and de novo lipogenesis due to PNPLA3 sequence variant. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(5):E821-E825.
19. Marston NA, Giugliano RP, Im K, et al. Association between triglyceride lowering and reduction of cardiovascular risk across multiple lipid-lowering therapeutic classes. *Circulation.* 2019;140(16): 1308-1317.
20. Allara E, Morani G, Carter P, et al. Genetic determinants of lipids and cardiovascular disease outcomes. *Circulation: Genomic and Precision Medicine.* 2019;12(12):e002711.
21. Do R, Willer CJ, Schmidt EM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nature Genetics.* 2013;45(11):1345-1352.

22. Smith GI, Shankaran M, Yoshino M, et al. Insulin resistance drives hepatic de novo lipogenesis in nonalcoholic fatty liver disease. *Journal of Clinical Investigation.* 2020;130(3):1453-1460.
23. Dina C, Meyre D, Gallina S, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nature Genetics.* 2007;39(6):724-726.
24. Eissing L, Scherer T, Tödter K, et al. De novo lipogenesis in human fat and liver is linked to ChREBP-β and metabolic health. *Nature Communications.* 2013;4:1528-1528.
25. Santoro N, Caprio S, Pierpont B, et al. Hepatic de novo lipogenesis in obese youth is modulated by a common variant in the GCKR gene. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(8):E1125-E1132.
26. Strable MS, Ntambi JM. Genetic control of de novo lipogenesis: role in diet-induced obesity. *Critical reviews in biochemistry and molecular biology.* 2010;45(3):199-214.
27. Kim K-H. Regulation of mammalian acetyl-coenzyme a carboxylase. *Annual Review of Nutrition.* 1997;17(1):77-99.
28. Simon JA, Hodgkins ML, Browner WS, et al. Serum Fatty Acids and the Risk of Coronary Heart Disease. *American Journal of Epidemiology.* 1995;142(5):469-476.
29. Wang L, Folsom AR, Eckfeldt JH, the ASI. Plasma fatty acid composition and incidence of coronary heart disease in middle aged adults: The Atherosclerosis Risk in Communities (ARIC) Study. *Nutrition, Metabolism and Cardiovascular Diseases.* 2003;13(5):256-266.
30. Wu JHY, Lemaitre RN, Imamura F, et al. Fatty acids in the de novo lipogenesis pathway and risk of coronary heart disease: the Cardiovascular Health Study. *American Journal of Clinical Nutrition.* 2011;94(2):431-438.
31. Rosqvist F, McNeil CA, Pramfalk C, et al. Fasting hepatic de novo lipogenesis is not reliably assessed using circulating fatty acid markers. *American Journal of Clinical Nutrition.* 2019;109(2):260-268.
32. Beer NL, Tribble ND, McCulloch LJ, et al. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Human Molecular Genetics.* 2009;18(21):4081-4088.
33. Rees MG, Wincovitch S, Schultz J, et al. Cellular characterisation of the GCKR P446L variant associated with type 2 diabetes risk. *Diabetologia.* 2012;55(1):114-122.
34. Brouwers MCGJ, Jacobs C, Bast A, Stehouwer CDA, Schaper NC. Modulation of glucokinase regulatory protein: A double-edged sword? *Trends in Molecular Medicine.* 2015;21(10):583-594.
35. Simons PIHG, Simons N, Stehouwer CDA, et al. Association of common gene variants in glucokinase regulatory protein with cardiorenal disease: A systematic review and meta-analysis. *PloS One.* 2018;13(10):e0206174.
36. Calle RA, Amin NB, Carvajal-Gonzalez S, et al. ACC inhibitor alone or co-administered with a DGAT2 inhibitor in patients with non-alcoholic fatty liver disease: two parallel, placebo-controlled, randomized phase 2a trials. *Nature Medicine.* 2021.
37. Lloyd DJ, St Jean DJ, Kurzeja RJM, et al. Antidiabetic effects of glucokinase regulatory protein small-molecule disruptors. *Nature.* 2013;504(7480):437-440.
38. Zhu D, Gan S, Liu Y, et al. Dorzagliatin monotherapy in Chinese patients with type 2 diabetes: a dose-ranging, randomised, double-blind, placebo-controlled, phase 2 study. *Lancet Diabetes Endocrinol.* 2018;6(8):627-636.
39. Hodson L, Brouwers MCGJ. Non-alcoholic fatty liver disease concerns with glucokinase activators. *Lancet Diabetes Endocrinol.* 2018;6(9):684-685.
40. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Human Molecular Genetics.* 2018;27(R2):R195-R208.

Table S9.1 De novo lipogenesis susceptibility genes identified by their association with low serum SHBG and high serum triglycerides (unbiased approach).

SNP	Gene	Chr:Pos[†]	Effect allele	Other allele	EAF	SHBG		Triglycerides	
						β	SE	β	SE
rs40270	*C5orf67*	5:55804552	C	A	0.728	-0.016	0.001	0.038	0.003
rs13389219	*COBLL1*	2:165528876	C	T	0.606	-0.016	0.001	0.027	0.003
rs1412957	*EYA2*	20:45557065	G	A	0.446	-0.007	0.001	0.012	0.005
rs1178947	*FZD9*	7:72850178	T	C	0.806	-0.008	0.001	0.081	0.006
rs13280055	*GATA4*	8:11522353	A	G	0.118	-0.008	0.002	0.038	0.006
rs1260326	*GCKR*	2:27730940	T	C	0.398	-0.041	0.001	0.115	0.003
rs7096937	*GPAM*	10:113950418	C	T	0.719	-0.009	0.001	0.012	0.005
rs1634791	*HLA-C*	6:31276777	G	A	0.585	-0.012	0.001	0.010	0.003
rs10221473	*INSR*	19:7236626	A	G	0.007	-0.007	0.001	0.004	0.002
rs10822163	*JMJD1C*	10:65124098	C	G	0.523	-0.053	0.001	0.031	0.005
rs4240624	*LOC157273*	8:9184231	G	A	0.113	-0.022	0.001	0.028	0.006
rs2943641	*LOC646736*	2:227093745	C	T	0.647	-0.014	0.001	0.026	0.004
rs72663907	*MAPK10*	4:87185271	G	T	0.104	-0.009	0.001	0.015	0.004
rs35824797	*MAU2*	19:19456264	T	C	0.076	-0.012	0.002	0.022	0.009
rs17145750	*MLXIPL*	7:73026378	C	T	0.847	-0.013	0.002	0.102	0.005
rs174533	*MYRF*	11:61549025	A	G	0.328	-0.010	0.001	0.054	0.003
rs12928099	*PDXDC1*	16:15150505	C	A	0.731	-0.010	0.001	0.029	0.003
rs4805881	*PEPD*	19:33896432	A	C	0.364	-0.009	0.001	0.018	0.003
rs199839547	*R3HDM2*	12:57701103	GTT	G	0.805	-0.017	0.001	0.027	0.003
rs668871	*SLC22A3*	6:160769811	C	T	0.535	-0.018	0.001	0.013	0.002
rs2980858	*TRIB1*	8:126501177	T	C	0.305	-0.010	0.001	0.085	0.005
rs12311848	*ZNF664*	12:124486851	A	G	0.683	-0.010	0.001	0.024	0.003

[†] Based on the human assembly GRCh37 reference genome assembly. Data for this supplementary table are retrieved from[1-5].
Abbreviations: SNP single nucleotide polymorphism; Chr chromosome; Pos position; EAF effect allele frequency; SE standard error: SHBG sex hormone binding globulin.

Table S9.2 De novo lipogenesis susceptibility genes identified by their association with low serum SHBG and documented involvement in de novo lipogenesis (biased approach).

SNP	Gene	Chr:Pos†	Effect allele	Other allele	EAF	SHBG		Involvement in DNL (reference)
						β	SE	
rs1799831	*GCK*	7:44199142	T	C	0.161	-0.009	0.001	[6]
rs1260326	*GCKR*	2:27730940	T	C	0.398	-0.041	0.001	[7]
rs7096937	*GPAM*	10:113950418	C	T	0.719	-0.009	0.001	[8]
rs10221473	*INSR*	19:7236626	A	G	0.454	-0.007	0.001	[9]
rs17145750	*MLXIPL*	7:73026378	C	T	0.847	-0.013	0.002	[10]
rs738409	*PNPLA3*	22:44324727	C	G	0.778	-0.023	0.001	[11,12]
rs11202594	*PTEN*	10:89641222	A	G	0.150	-0.006	0.001	[13]
rs12487736	*SCAP*	3:47459679	T	C	0.565	-0.009	0.001	[14]
rs2980858	*TRIB1*	8:126501177	T	C	0.305	-0.010	0.001	[15]

† Based on the human assembly GRCh37 reference genome assembly, Data for this supplementary table are retrieved from[1,2]
Abbreviations: SNP single nucleotide polymorphism; Chr chromosome; Pos position; EAF effect allele frequency; SE standard error; SHBG sex hormone binding globulin; DNL de novo lipogenesis.

Table S9.3. Genetic variants that associate with serum fatty acids.

SNP	Gene	Chr:Pos†	Effect allele	Other allele	EAF	Fatty acid
rs2391388	*ALG14*	1:95485825	C	A	0.459	16:0; 18:0
rs102275	*FADS1/2*	11:61557803	T	C	0.651	18:0; 18:ln-9; 16:ln-7
rs780093	*GCKR*	2:27742603	T	C	0.387	16:ln-7
rs11190604	*HIF1AN*	10:102302457	G	A	0.213	16:ln-7
rs11119805	*LPGAT1*	1:211918244	T	A	0.866	18:0
rs603424	*PKD2L1*	10:102075479	G	A	0.790	16:ln-7
rs6722456	-	2:134529091	G	A	0.971	16:ln-7

† Based on the human assembly GRCh37 reference genome assembly. Data for this supplementary table are retrieved from[1,16].
Abbreviations: SNP single nucleotide polymorphism; Chr chromosome; Pos position; EAF effect allele frequency.

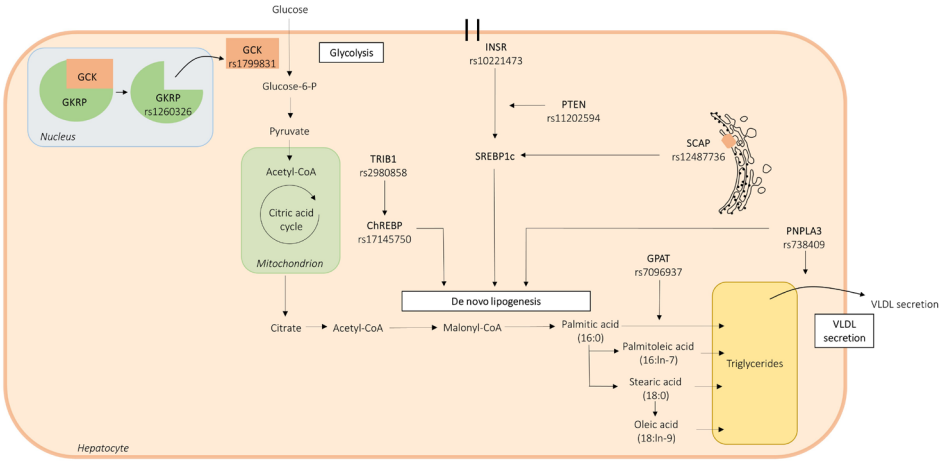

Figure S9.1. Involvement of genetic variants (identified by the biased approach) in the regulation of de novo lipogenesis. The RefSNP (rs) numbers refer to variants in the de novo lipogenesis susceptibility genes that were identified by a biased approach (Supplementary Table S9.2). The effect alleles of the genetic variants (see Supplementary Table S9.2) have been associated with low serum SHBG levels, a downstream product of enhanced de novo lipogenesis. The putative role of these variants in stimulating de novo lipogenesis is as detailed below:

De novo lipogenesis encompasses the synthesis of palmitic acid and other more complex fatty acids, from non-lipid precursors. Acetyl-coenzyme A (acetyl-CoA), which can be formed as a product of the glycolytic pathway, serves as a substrate for de novo lipogenesis. During de novo lipogenesis, acetyl-coA is elongated to form malonyl-coenzyme A (malonyl-coA) and palmitic acid. Palmitic acid can subsequently be processed to form other fatty acids (palmitoleic acid, stearic acid or oleic acid) and triglycerides, which can be stored in the liver or secreted in the form of very-low-density lipoprotein (VLDL) particles.

The formation of fatty acids in the liver is tightly regulated by several enzymes and transcription factors. The first, rate-limiting step in glycolysis is facilitated by glucokinase (GCK), which converts glucose into glucose-6-phosphate. The activity of GCK is regulated by glucokinase regulatory protein (GRKP), which binds, and thereby inactivates, GCK. The expression of lipogenic enzymes that facilitate de novo lipogenesis, such as ACC and FASN, is regulated by the transcription factors carbohydrate-response element-binding protein (ChREBP) and sterol regulatory element-binding protein 1c (SREBP-1c). In turn, the expression of ChREBP and SREBP-1c is regulated by several upstream factors. For instance, the insulin receptor (INSR), SREBP cleavage-activating protein (SCAP), and phosphate and tensin homolog (PTEN) activate SREBP1c expression. Tribbles homolog 1 (TRIB1) stimulates the expression of ChREBP. Finally, the conversion of palmitic acid to triglycerides is facilitated by glycerol-3-phosphate acyltransferase (GPAT). The major allele of the *PNPLA3* gene which encodes a variant in patatin-like phospholipase domain-containing protein 3 (PNPLA3) has been shown to be associated with de novo lipogenesis, although the exact mechanism remains unknown. In addition, the *PNPLA3* major allele has also been associated with VLDL remodelling and serum triglyceride levels.

References

1. Nelson CP, Goel A, Butterworth AS, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nature Genetics.* 2017;49(9):1385-1391.
2. Ruth KS, Day FR, Tyrrell J, et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nature Medicine.* 2020;26(2):252-258.
3. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nature Genetics.* 2013;45(11):1274-1283.
4. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562(7726):203-209.
5. Nealelab UK Biobank GWAS. http://www.nealelab.is/uk-biobank/. Published 2018. Accessed.
6. Peter A, Stefan N, Cegan A, et al. Hepatic glucokinase expression is associated with lipogenesis and fatty liver in humans. *Journal of Clinical Endocrinology and Metabolism.* 2011;96(7):E1126-E1130.
7. Santoro N, Caprio S, Pierpont B, et al. Hepatic de novo lipogenesis in obese youth is modulated by a common variant in the GCKR gene. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(8): E1125-E1132.
8. Wendel AA, Cooper DE, Ilkayeva OR, Muoio DM, Coleman RA. Glycerol-3-phosphate acyltransferase (GPAT)-1, but not GPAT4, incorporates newly synthesized fatty acids into triacylglycerol and diminishes fatty acid oxidation. *Journal of Biological Chemistry.* 2013;288(38):27299-27306.
9. Cook JR, Langlet F, Kido Y, Accili D. Pathogenesis of selective insulin resistance in isolated hepatocytes. *Journal of Biological Chemistry.* 2015;290(22):13972-13980.
10. Eissing L, Scherer T, Tödter K, et al. De novo lipogenesis in human fat and liver is linked to ChREBP-β and metabolic health. *Nature Communications.* 2013;4:1528-1528.
11. Mancina RM, Matikainen N, Maglio C, et al. Paradoxical dissociation between hepatic fat content and de novo lipogenesis due to PNPLA3 sequence variant. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(5):E821-E825.
12. Pirazzi C, Adiels M, Burza MA, et al. Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro. *Journal of Hepatology.* 2012;57(6): 1276-1282.
13. Qiu W, Federico L, Naples M, et al. Phosphatase and tensin homolog (PTEN) regulates hepatic lipogenesis, microsomal triglyceride transfer protein, and the secretion of apolipoprotein B-containing lipoproteins. *Hepatology.* 2008;48(6):1799-1809.
14. Moon Y-A, Liang G, Xie X, et al. The Scap/SREBP pathway is essential for developing diabetic fatty liver and carbohydrate-induced hypertriglyceridemia in animals. *Cell Metabolism.* 2012;15(2):240-246.
15. Ishizuka Y, Nakayama K, Ogawa A, et al. TRIB1 downregulates hepatic lipogenesis and glycogenesis via multiple molecular interactions. *Journal of Molecular Endocrinology.* 2014;52(2):145-158.
16. Wu JH, Lemaitre RN, Manichaikul A, et al. Genome-wide association study identifies novel loci associated with concentrations of four plasma phospholipid fatty acids in the de novo lipogenesis pathway: results from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. *Circulation: Cardiovascular Genetics.* 2013;6(2):171-183.

10

CHAPTER TEN

Summary and general discussion

Polycystic ovary syndrome is estimated to be present in one in every ten women of childbearing age, and, as such, it is the most common endocrine disorder amongst premenopausal women[1]. Women with PCOS have an increased risk of developing cardiometabolic disease, including coronary artery disease and type 2 diabetes[2-6]. However, as obesity is also highly prevalent in these women[7,8], it has remained uncertain whether the increased risk of cardiometabolic disease is the result of PCOS per se. Alternatively, obesity, and its downstream metabolic complications such as de novo lipogenesis, may be the common aetiological factor that increases the risk of PCOS and other cardiometabolic disease.

This thesis, therefore, aimed to investigate the association between PCOS and cardiometabolic disease, and to assess the role of de novo lipogenesis herein. To unravel this triangular relationship this thesis had three sub-aims:

1. To assess the association between (risk factors of) PCOS and cardiometabolic disease.
2. To assess whether de novo lipogenesis decreases serum sex hormone-binding globulin (SHBG).
3. To assess the associations between de novo lipogenesis and cardiometabolic disorders and the role of serum SHBG herein.

In this chapter the main findings of this thesis are summarized and discussed in the context of current scientific literature, and in the context of methodological considerations. Furthermore in light of the findings of this thesis, possible directions for future research are explored.

Summary and general discussion

Part 1: PCOS and the association with (risk factors of) cardiometabolic disease

Although PCOS associates with an increased risk of cardiometabolic disease, this may not be the result of a causal association. Indeed, while type 2 diabetes is more common in women with PCOS[9], a previous Mendelian randomization study has reported that this is not the result of a *causal* effect of PCOS on type 2 diabetes[10]. In that study it remained uncertain whether similar conclusions can be drawn for the effect of PCOS on coronary artery disease, as the association with coronary artery disease was studied in a dataset of women *and* men combined[10]. In chapter two, we repeated the Mendelian randomization analysis in a sex-specific dataset, and observed no statistically significant association between genetically predicted risk of PCOS and risk of coronary artery disease. Although methodological limitations should be taken

into account (see '*Methodological considerations*'), these findings suggest that PCOS per se does not seem to have a causal effect on the risk of coronary artery disease.

While PCOS may not be causal in influencing the risk of coronary artery disease, the coexistence of the two conditions implies there should be a common underlying factor. Obesity is highly prevalent in both PCOS and coronary artery disease[8,11], and the strengths of observational associations between PCOS and coronary artery disease are attenuated upon adjustment for body mass index (BMI)[3]. Therefore, we conducted sex-specific Mendelian randomization analyses to explore the associations between genetically predicted BMI with risk of PCOS and coronary artery disease (**chapter two**). The results suggest a causal role for obesity, and its downstream metabolic complications, as the common ground for both PCOS and coronary artery disease. This is in line with the results of previous overall Mendelian randomization studies[12-14], and with intervention studies that report that weight loss results in improvements of cardiovascular risk factors and PCOS phenotype[15,16].

The mechanisms that link obesity and PCOS have been the subject of much debate. It is postulated that hyperandrogenism (i.e. high levels of free testosterone) may play a role. Nevertheless, it has remained uncertain whether free testosterone levels are a cause or consequence of (other) features of PCOS[17,18]. Free testosterone levels are the result of the changes in its two determinants, i.e. total testosterone (originating from the ovaries and adrenal glands[19]) and SHBG (originating from the liver[20]). These determinants are independently regulated in women (**chapter six**). Therefore, it can be reasoned that the differential associations between features of PCOS with SHBG and total testosterone may hint at the direction of the effects (Figure 10.1). If free testosterone is causal in the development of features of PCOS, then both SHBG and total testosterone would be associated with the feature of interest (Figure 10.1, panel A). If, however, free testosterone is the consequence of features in PCOS, then only one of the determinants of free testosterone would be associated (Figure 10.1, panel B).

In a well-phenotyped group of PCOS patients (**chapter three**), metabolic features of PCOS (i.e. BMI, serum lipids and insulin resistance) associated primarily with serum SHBG, but not with total testosterone. On the other hand, reproductive features of PCOS (i.e. antral follicle count and anti-Müllerian hormone [AMH]) associated with total testosterone, but not with serum SHBG. These findings suggest that changes in free testosterone are more likely to be the consequence of metabolic and reproductive disturbances in PCOS rather than their cause. In line with this, in **chapter six** we observed that free testosterone does not drive the risk of type 2 diabetes in women[21].

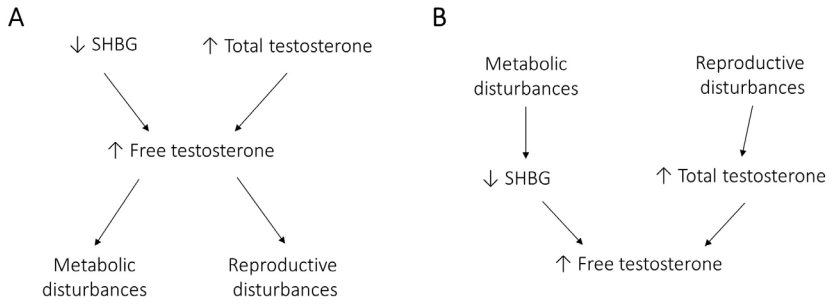

Figure 10.1 Potential associations between features of PCOS and androgen markers. If free testosterone has a causal effect on metabolic and reproductive disturbances in PCOS, then both determinants of free testosterone (i.e. SHBG and total testosterone) should associate with these features (A). If free testosterone levels are a consequence of metabolic or reproductive disturbances in PCOS, then these features would associate with only one of the determinants of free testosterone (B).

In conclusion, in the first part of this thesis we aimed to investigate the causal association between PCOS and (risk factors of) cardiometabolic disease. Using Mendelian randomization analyses we conclude that PCOS per se does not seem to be causal in the risk of coronary artery disease. Rather, obesity appears to be the common denominator of both PCOS and coronary artery disease. Furthermore, cardiometabolic features of PCOS, such as dyslipidaemia and insulin resistance, associate primarily with a reduced serum SHBG level, suggesting that changes in serum SHBG may be a consequence of metabolic dysfunction.

Part 2: The effect of de novo lipogenesis on sex hormone-binding globulin

Serum SHBG is synthesized in the liver, a process which is tightly regulated by several transcription factors, including hepatocyte nuclear factor 4α (HNF-4α)[22]. Several hormonal, nutritional and metabolic factors are known to influence serum SHBG levels[23-28]. Animal studies have shown that monosaccharide-induced de novo lipogenesis is one of the processes decreasing the levels of HNF-4α and, consequently, SHBG[29]. In part two of this thesis, we have used different methodological approaches to extrapolate these data to humans. First, we observed that in women with varying degrees of obesity and hepatic steatosis, de novo lipogenesis, measured with stable isotopes, was statistically significantly associated with serum SHBG levels (**chapter four**). Of interest, this association was not observed in men, although this may be the result of insufficient statistical power rather than biological differences. Second, by studying patients with glycogen storage disease type 1a (GSD1a), who are characterized by higher rates of de novo lipogenesis[30], we unravelled that de novo

lipogenesis seems to causally affect serum SHBG levels (**chapter five**). Finally, from a large-scale GWA study of serum SHBG[21], we identified that genes involved in the regulation of de novo lipogenesis were associated with serum SHBG levels (**chapter six**). Of note, the effect of de novo lipogenesis on serum SHBG seems to be independent of insulin.

The exact mechanisms that mediate the effect of de novo lipogenesis on serum SHBG remain uncertain. Genetic variants that affect very-low-density lipoprotein (VLDL) secretion from the liver, such as variants in *TM6SF2*[31], were not found to associate with serum SHBG levels (**chapter six**). This suggests that liver fat per se is not involved in regulating SHBG. In experimental studies, incubation of HepG2 cells with palmitate – the end-product of de novo lipogenesis[32] – reduced HNF-4α and SHBG levels, thereby indicating a potential role for palmitate[29]. We observed that in patients with GSD1a, the fraction of intrahepatic saturated fatty acids – which is a measure of palmitate – also associated with serum SHBG levels (**chapter five**). Whether the effect of palmitate on HNF-4α is direct, or mediated by other factors deserves further study. Moreover, SHBG may also be regulated by other pathways of intrahepatic lipid (IHL) accumulation which increase palmitate levels, such as the influx of free fatty acids from the adipose tissue or diet into the liver. In future studies, the contribution of this pathway to serum SHBG levels could be studied with stable isotopes, or, alternatively, by studying patients with lipodystrophy, who are characterized by an increased flux of free fatty acids[33].

As discussed hitherto, by studying different genetic models we found an effect of de novo lipogenesis on serum SHBG, however, this does not exclude the possibility that the relationship is bidirectional. Indeed, experimental studies have found that transgenic *SHBG* mice fed a high-fructose diet, are protected from IHL accumulation and have a reduced expression of lipogenic enzymes, including acetyl-coenzyme A carboxylase (ACC) and fatty acid synthase (FAS)[34]. Nevertheless, these experimental data have not been extrapolated to humans. For this, intervention trials that target serum SHBG (e.g. through weight loss, oestrogen-containing oral contraceptives or thyroid hormones analogues[35-38]) would be needed. However, these interventions are plagued by a high risk of pleiotropic effects, which would challenge any conclusions drawn from such a study. Alternatively, the rates of de novo lipogenesis could be studied in individuals with a variant in *SHBG*, although identifying sufficient number of individuals with this variant may limit the feasibility of such a study[39].

We observed a sexually dimorphic association between de novo lipogenesis and serum SHBG. In **chapter four** and **chapter seven**, de novo lipogenesis and IHL content associated more strongly with serum SHBG levels in women than in men. In line with our observational findings, variants in *GCKR*, that are associated with higher rates of de

novo lipogenesis, also associate more strongly with serum SHBG levels in women than in men[40]. The biological mechanisms that account for these sex differences are not well understood, and deserve further investigation.

In conclusion, in the second part of this thesis we observe that de novo lipogenesis seems to regulate serum SHBG levels in humans. Future studies should assess whether other pathways of intrahepatic accumulation of palmitate (or other fatty acids) also contribute to the regulation of serum SHBG levels. Moreover, it cannot be exclude that the association between de novo lipogenesis and serum SHBG is bidirectional, i.e. that serum SHBG may also affect de novo lipogenesis.

Part 3: The association between de novo lipogenesis and cardiometabolic disease, and the role of SHBG herein

In **chapter two** we observed that obesity predisposes to PCOS *and* coronary artery disease. Individuals with obesity are more likely to have higher rates of de novo lipogenesis[41,42]. Therefore, in part three of this thesis, we hypothesized that de novo lipogenesis may be the common factor that predisposes to both PCOS and other cardiometabolic disease, and furthermore, that these relationships may be mediated by SHBG.

Part 3.1 The association between de novo lipogenesis and polycystic ovary syndrome

The importance of non-alcoholic fatty liver disease (NAFLD) in the pathophysiology of PCOS is increasingly being recognised[43]. In particular, it has been proposed that NAFLD associates with the hyperandrogenic phenotype of PCOS[44]. The findings presented in this thesis elaborate thereon, by implicating de novo lipogenesis – one of the primary pathways leading to NAFLD – in the pathophysiology of (hyperandrogenic) PCOS.

First, we observed that de novo lipogenesis decreases serum SHBG levels (**chapters four, five and six**), which, in turn, affects free testosterone levels. Higher levels of free testosterone are a diagnostic criterion of PCOS, and, unsurprisingly, SHBG was, therefore, found to be causal in the risk of PCOS (**chapter six**). Free testosterone (and therefore SHBG) may contribute to the reproductive pathophysiology of PCOS by disturbing the ovarian antral follicle development[45]. The presence of numerous small antral follicles then increases the synthesis and secretion of AMH by the granulosa cells[46]. This may, in turn, disrupt the aromatase-induced conversion of androgens to oestrogens and dysregulate the pulsatile secretion of gonadotrophin-releasing hormone (GnRH)[46-49]. Both aspects further enhance the ovarian androgen concentration. Elevated levels of AMH also reduce the sensitivity of antral follicles to follicle-stimulating hormone (FSH), which further exacerbates abnormal follicular development[50]. Despite the aforementioned possible effect of free testosterone on

reproductive features, we did not observe an association between free testosterone with AMH, antral follicle count or LH (chapter three). Therefore, whether, and the extent to which, free testosterone contributes to the ovarian features of PCOS deserves further investigation.

Second, it has been speculated that de novo lipogenesis associates with insulin resistance, which may also contribute to the dysregulation of ovarian follicle development. In this thesis, we were unable to study the effect of de novo lipogenesis on insulin resistance specifically, although we did observe an association between IHL content and type 2 diabetes (chapter seven). The subsequent effect of insulin resistance on PCOS remains much debated, but it has been suggested that within theca cells, insulin may act in synergy with LH to stimulate the release of testosterone[48]. Testosterone may then affect ovarian function, as discussed previously. Moreover, it is thought that insulin promotes the arrest of pre-antral follicular development, thereby contributing to the polycystic ovarian morphology[51,52].

Part 3.2 The association between de novo lipogenesis and type 2 diabetes

In chapter seven we observed a statistically significant association between IHL content, assessed with magnetic resonance imaging, and type 2 diabetes. This corroborates previous observational studies, although these primarily used abdominal ultrasound to quantify IHL content[53-57]. Unfortunately, as a result of the practical restrictions of measuring de novo lipogenesis in large-scale studies, we were unable to extrapolate our findings to the effect of de novo lipogenesis on type 2 diabetes. Although de novo lipogenesis is one of the primary pathways leading to the accumulation of IHL, as a result of the partitioning of IHL (i.e. the differential distribution of IHL to the storage pool, oxidation pathway or secretion as VLDL particles[58]) there is not always a linear relationship between de novo lipogenesis and IHL content[59], as is also illustrated in chapter four.

In chapter seven we observed that the association between IHL content and type 2 diabetes is mediated by serum SHBG, thereby implicating SHBG as a hepatokine (i.e. a liver-derived protein with systemic metabolic effects)[60]. This is in line with experimental studies that have previously observed an effect of SHBG on glucose homeostasis. Human transgenic *SHBG* mice fed a high-fat diet, were protected against diet-induced dysregulation of insulin levels and demonstrated improved glucose homeostasis and insulin sensitivity[61]. In ex vivo studies the effects of SHBG on metabolic parameters were also seen under hormone-deprived conditions, indicating that the metabolic effects of SHBG may be direct, rather than mediated by free testosterone[61]. In chapter six, we extrapolated these findings to humans, as we observed that genetically predicted SHBG had a causal effect on type 2 diabetes, independent of its effects on free testosterone.

Nevertheless, the effect of IHL content on type 2 diabetes is likely to be the result of several mediating pathways, of which serum SHBG is merely one. Other well-known mediators are the effect of IHL content on fatty acid metabolites and thereby insulin resistance, as well as the effect of IHL content on excess (hepatic) glucose production[62]. In support of mediating pathways independent of serum SHBG, genetic studies identified that variants in *TM6SF2*, which are known to affect IHL content but not serum SHBG, were also associated with type 2 diabetes[21,31,63].

Part 3.3 The association between de novo lipogenesis and chronic kidney disease

In **chapter eight**, we used genetic epidemiology to study the association between de novo lipogenesis and chronic kidney disease. Unexpectedly, the minor allele of *GCKR*, which predisposes to higher rates of de novo lipogenesis[64], was associated with a higher estimated glomerular filtration rate (eGFR) and lower risk of chronic kidney disease, albeit the latter was not statistically significant. These findings were in contrast with several observational studies that report an association between NAFLD – a downstream consequence of de novo lipogenesis – and an increased risk of chronic kidney disease[65-71]. The contradicting findings may be the result of genetic variants that are in high linkage disequilibrium with *GCKR* and that positively affect kidney function. Alternatively, variants in *GCKR* may have pleiotropic effects which positively affect kidney function. For instance, *GCKR* is known to decrease fasting plasma glucose, which may result in an overall renal protective effect[72]. Nevertheless, these findings warrant further research, as experimental mice studies report that liver-specific knock-out of glucokinase (which is associated with lower de novo lipogenesis rates and higher fasting glucose[73,74], and is hence opposite to the effect of *GCKR*) is associated with increased kidney damage, and thereby corroborate our findings[75].

Part 3.4 The association between de novo lipogenesis and coronary artery disease

In **chapter eight** and **chapter nine,** we used genetic epidemiology – with a single-variant and multiple-variant approach – to study the association between de novo lipogenesis susceptibility genes and coronary artery disease. The minor allele of *GCKR* significantly associated with the risk of coronary artery disease. The multiple-variant approach likewise identified a significant association between de novo lipogenesis susceptibility genes, but not fatty acid genes, and coronary artery disease. Although NAFLD has been extensively associated with cardiovascular disease in observational data[76,77], genetic epidemiology suggests that not all pathways of IHL accumulation contribute equally. Previous studies have shown that genetic variants that contribute to NAFLD development through impaired VLDL secretion are protective of coronary artery disease[78], while the present studies suggest that de novo lipogenesis susceptibility genes increase the risk of coronary artery disease. Of note, as gene-de novo lipogenesis

data was unavailable, we were unable to conduct full Mendelian randomization analyses. This limits the ability to draw strong conclusions on the causal role of de novo lipogenesis in coronary artery disease. Randomized controlled trials are needed to validate our findings and assess how much risk reduction can be achieved by reducing de novo lipogenesis.

It is likely that the effect of de novo lipogenesis susceptibility genes on coronary artery disease is mediated by serum triglycerides and SHBG levels, which are both downstream consequences of de novo lipogenesis[64,79,80]. The causal effect of serum triglycerides on coronary artery disease has been well-established in both randomized controlled trials as well as Mendelian randomization studies[81-83]. On the contrary, the effect of serum SHBG on coronary artery disease has not been as extensively studied, although initial findings do support a potential role for SHBG[84-87]. Given that recent studies observe an effect of SHBG on cardiometabolic factors under hormone-deprived conditions[61,88], it would be of interest to study whether the effect of SHBG on coronary artery disease is direct, or mediated by free testosterone.

In conclusion, in the third part of this thesis we found that a dysregulation of de novo lipogenesis may (causally) contribute to the increased risk of several cardiometabolic complications, including PCOS, type 2 diabetes and coronary artery disease. Serum SHBG acts as an important mediator of these associations. Serum SHBG should, therefore, be regarded as a hepatokine with metabolic effects. Unexpectedly, the minor allele of *GCKR* associated with a reduced risk of chronic kidney disease, which is a finding that deserves further investigation.

Methodological considerations

The studies conducted as part of this thesis further our understanding of the relationships between PCOS and cardiometabolic disease and the role of de novo lipogenesis herein. However, the variety of research methodologies used in this thesis each have notable limitations that should be considered when interpreting their results.

Methodological considerations in observational studies

Although a significant strength of observational studies is the ability to explore associations in large groups of individuals, several forms of bias can challenge the results of observational studies.
First, error in the measurement of variables can result in information bias[89]. Therefore, where possible, we used gold-standard measurement techniques to measure the most

important variables in this thesis. Stable isotopes were used to quantify de novo lipogenesis and proton magnetic resonance spectroscopy was used to non-invasively measure IHL content. Serum biomarkers, such as serum SHBG, were measured using standardized laboratory protocols, and physical examinations (i.e. ovarian volume or antral follicle count) were conducted by an experienced gynaecologist. In spite of our efforts to reduce this bias, some forms of measurement error are still likely to persist or could not be avoided all together. This is particularly likely for subjective measures assessed in PCOS patients, such as the Ferriman-Gallwey score or recall of the length of the menstrual cycle.

Second, selection of women with PCOS from a tertiary medical centre outpatient clinic for menstrual cycle disorders (**chapter three**) may have led to observations which are not reflective of the true population, i.e. referral bias. Cases selected from this setting, rather than the general population, may suffer from a more severe phenotype. Extrapolation of the current results to all women with PCOS should therefore be done with care. Likewise, it is reasonable that in **chapter seven** the recruitment strategy of the Maastricht Study as well as the in- and exclusion criteria applied may have led to a selective inclusion of healthier patients, i.e. selection bias[89]. This could have resulted in an underestimation of some associations. Selection bias can also be introduced when cases are matched to controls[90], as was done in **chapter five**.

Third, the association between an independent and dependent variable can be influenced by extraneous factors, i.e. confounding. Where possible, associations were adjusted for the most important confounders to assess the robustness of our findings. However, as a result of a small sample size, adjustment for confounders was unachievable in **chapter five**. Therefore, to account for potential confounding, we matched cases and controls on the most important factors. We cannot, however, exclude residual confounding in this, and other studies. In addition, we aimed to be considerate of variables that are, in fact, mediators or upstream factors (i.e. instrumental variables), as adjustment for these variables may lead to overadjustment bias[91,92]. This is illustrated in **chapter four**, where statistical significance was lost in the association between de novo lipogenesis and SHBG after adjustment for serum insulin. This was not unexpected as insulin is an upstream factor of de novo lipogenesis, and, therefore, this may be indicative of overadjustment bias.

Finally, an inherent limitation of observational studies is the inability to deduce the direction of an effect, i.e. temporal ambiguity[89]. In **chapter four** we found a significant association between de novo lipogenesis and serum SHBG, however as the study was of an observational nature we could not deduce whether de novo lipogenesis preceded serum SHBG, or vice versa. Using genetic epidemiology we found that de novo

lipogenesis affects serum SHBG, but still cannot exclude the possibility that, in fact, the association is bidirectional.

Methodological considerations in genetic epidemiology

Common genetic variants

Genetic epidemiology can overcome some of the traditional limitations seen in observational studies. Mendelian randomization is a form of genetic epidemiology that can infer causal associations in observational data[93]. It is based on the concept that each individual randomly receives an exposure-predisposing or an exposure-protective allele at conception. Similar to a randomized-controlled trial, this random distribution of variants amongst individuals allows us to gain insight into the effect of lifelong exposure to the variable of interest on the outcome. Any well-conducted Mendelian randomization study is subject to three primary assumptions[94]. First, the genetic variant(s) must (robustly) associate with the exposure trait of interest. Second, the genetic variant(s) may not associate with any confounder of the association between the exposure and outcome (i.e. horizontal pleiotropy). Finally, the genetic variant(s) may not associate with the outcome of interest, other than through the exposure trait.

Mendelian randomization was originally conducted with a single genetic variant that directly affected the functioning of a protein. Over time it has evolved to be used with clusters of genetic variants that associate with more complex, heterogeneous traits, such as PCOS or de novo lipogenesis[95]. The use of Mendelian randomization with complex traits has several limitations. First, the gene-exposure associations tend to be weaker and are less likely to be specific to the exposure[96]. For instance, the PCOS susceptibility genes used in **chapter two** associated more strongly with the reproductive features rather than the metabolic features of PCOS, and consequently may represent only a subset of the PCOS phenotype. Second, the use of less specific variants increases the risk of horizontal pleiotropy, i.e. the genetic variants may affect the outcome through a trait other than the studied exposure trait[97].

However, when the limitations are adequately addressed, Mendelian randomization can be a valuable asset in assessing causality within observational study designs. As the inheritance of exposure-predisposing or exposure-protective alleles is independent and cannot be influenced by the outcome, the results of Mendelian randomization analyses are less prone to unmeasured confounding or reverse causation[98]. In addition, the two-sample Mendelian randomization approach, in which the exposure and outcome variable are derived from different datasets, allows long-term causal consequences of a disease to be studied[99]. As such, the effect of PCOS (which is diagnosed pre-menopause) on coronary artery disease (which tends to occur post-menopause) can be

assessed. Moreover, the two-sample approach has, by design, a weak instrument bias towards the null rather than in the direction of the confounded observation, as a result of which there is a lower likelihood of a type I error[100].

Rare mutations

Where Mendelian randomization uses common genetic variants to study causal associations between an exposure and outcome, the study of rare mutations can also provide insight into causal consequences of a certain trait. Patients with GSD1a have a genetic mutation in *G6PC* that results in increased substrate for de novo lipogenesis[30]. Therefore, studying these patients can shed light on the downstream consequences of higher rates of de novo lipogenesis (**chapter five**). Although the extreme effects seen with rare mutations, such as GSD1a, ensure that the differences between the diseased and the controls are more evident, it also increases the risk of pleiotropic effects[101]. Indeed, rare monogenetic mutations are often diagnosed based on the clustering of symptoms, marking the high risk of pleiotropy[101]. Furthermore, given the fasting intolerance in GSD1a patients, all measures were conducted after nocturnal feeding, which may also have induced pleiotropic effects. Lastly, as the mutations are rare, it can be challenging to find a sample size sufficiently large to achieve adequate power for detection of significant differences. Indeed, in **chapter five** we found several associations that neared, but did not reach, statistical significance.

Statistical power: population and effect size

The effect of a genetic variant is inversely related to the sample size required to achieve sufficient statistical power. In case of variants with small effects (i.e. common genetic variants) large groups are required. In case of variants with large effects (i.e. rare mutations) small groups are required[102]. This is illustrated in **chapter eight**, where we combined data from several large-scale cohorts in order to obtain a sample size sufficiently large enough to detect a significant association between the common variant in *GCKR* and coronary artery disease. In **chapter five**, we studied patients with a GSD1a and, as a result of the extreme phenotype of this rare mutation, we were able to detect significant differences despite the small sample size.

Population and effect sizes are not only of importance in genetic studies. Also in observational studies a sample size large enough to obtain sufficient statistical power is needed to detect significant associations. In **chapter four** there was a statistically significant association between de novo lipogenesis and serum SHBG in women, but not in men. Although this could be the result of biological differences, it could also be the result of insufficient statistical power to detect the association in men. Therefore, the study should be repeated in larger populations to further investigate these differences.

Conclusions & future directions

In conclusion, in this thesis we have studied the association between PCOS and cardiometabolic disease and assessed the role of de novo lipogenesis and SHBG herein. We conclude that PCOS per se is not causal in the risk of coronary artery disease. Rather, obesity seems the common denominator of both PCOS and coronary artery disease (Figure 10.2). More specifically, de novo lipogenesis – a downstream consequence of obesity[41] – is likely to be at the root of coexisting cardiometabolic disorders such as PCOS and cardiometabolic disease (Figure 10.2). Finally, in this thesis we have shown that the association between de novo lipogenesis with PCOS and other cardiometabolic disease is, in part, mediated by SHBG.

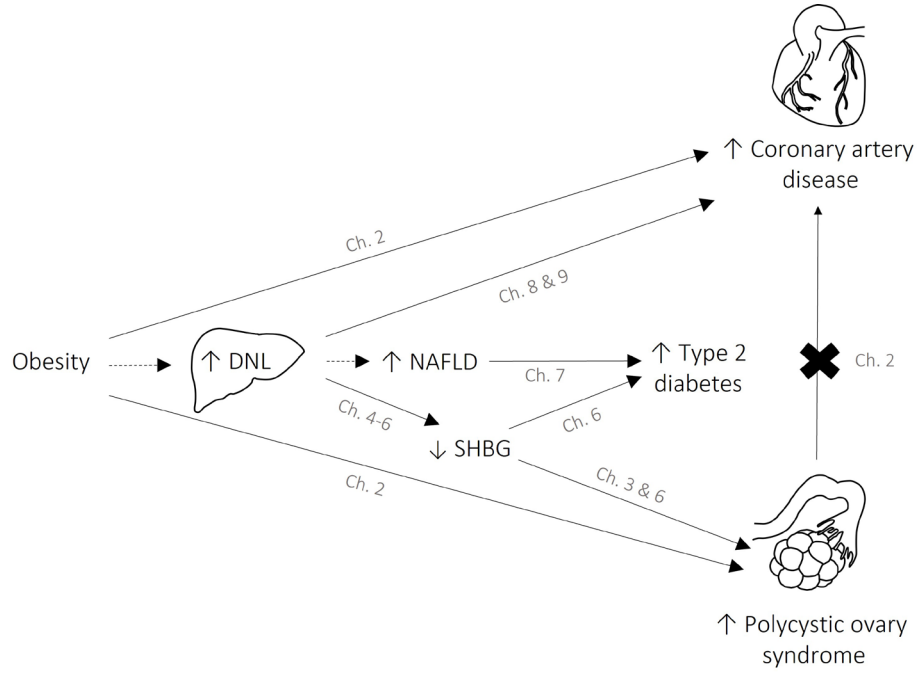

Figure 10.2 Overview of the findings in this thesis. Black lines represent (causal) associations studied in this thesis while dotted lines represent associations that have been assumed from previous literature.
Abbreviations: DNL de novo lipogenesis; NAFLD non-alcoholic fatty liver disease; SHBG sex hormone-binding globulin.

The impact of the findings in this thesis are elaborately discussed in the next chapter (*'Impact paragraph'*). In brief, the findings in this thesis justify an individualized approach in the treatment of women with PCOS. In these patients, reduced SHBG

levels may be a biomarker of metabolic dysfunction and could, therefore, indicate an increased risk of developing cardiometabolic disorders. The treatment plan of these patients should, therefore focus on the screening of cardiometabolic risk factors. Second, the findings in this thesis indicate that a dysregulation of de novo lipogenesis has several cardiometabolic consequences, including an increased risk of coronary artery disease, type 2 diabetes and PCOS. Future research endeavours should explore treatment options that target de novo lipogenesis, in order to reduce the risk of these complications. Finally, this thesis highlights the systemic metabolic effects of the SHBG, and its contribution to cardiometabolic disease. The exact mechanism through which SHBG contributes to these disorders, and the extent to which it plays a role in increasing the risk, deserves further investigation.

References

1. Wolf W.M., Wattick R.A., Kinkade O.N., Olfert M.D. Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity. *International Journal of Environmental Research and Public Health.* 2018;15(11):2589.
2. Berni T.R., Morgan C.L., Rees D.A. Women with polycystic ovary syndrome have an increased risk of major cardiovascular events: a population study. *Journal of Clinical Endocrinology and Metabolism.* 2021;106(9):e3369-3380.
3. de Groot P.C., Dekkers O.M., Romijn J.A., Dieben S.W., Helmerhorst F.M. PCOS, coronary heart disease, stroke and the influence of obesity: A systematic review and meta-analysis. *Hum Reprod Update.* 2011;17(4):495-500.
4. Glintborg D., Rubin K.H., Nybo M., Abrahamsen B., Andersen M. Cardiovascular disease in a nationwide population of Danish women with polycystic ovary syndrome. *Cardiovasc Diabetol.* 2018;17(1):37.
5. Lo J.C., Feigenbaum S.L., Yang J., Pressman A.R., Selby J.V., Go A.S. Epidemiology and adverse cardiovascular risk profile of diagnosed polycystic ovary syndrome. *Journal of Clinical Endocrinology and Metabolism.* 2006;91(4):1357-1363.
6. Zhou Y., Wang X., Jiang Y., et al. Association between polycystic ovary syndrome and the risk of stroke and all-cause mortality: insights from a meta-analysis. *Gynecol Endocrinol.* 2017;33(12):904-910.
7. Lim S.S., Norman R.J., Davies M.J., Moran L.J. The effect of obesity on polycystic ovary syndrome: a systematic review and meta-analysis. *Obes Rev.* 2013;14(2):95-109.
8. Bogers R.P., Bemelmans W.J.E., Hoogenveen R.T., et al. Association of Overweight With Increased Risk of Coronary Heart Disease Partly Independent of Blood Pressure and Cholesterol Levels: A Meta-analysis of 21 Cohort Studies Including More Than 300 000 Persons. *Arch Intern Med.* 2007;167(16):1720-1728.
9. Kakoly N.S., Khomami M.B., Joham A.E., et al. Ethnicity, obesity and the prevalence of impaired glucose tolerance and type 2 diabetes in PCOS: A systematic review and meta-regression. *Hum Reprod Update.* 2018;24(4):455-467.
10. Zhu T., Cui J., Goodarzi M.O. Polycystic ovary syndrome and risk of type 2 diabetes, coronary heart disease, and stroke. *Diabetes.* 2020.
11. Lim S.S., Davies M.J., Norman R.J., Moran L.J. Overweight, obesity and central obesity in women with polycystic ovary syndrome: A systematic review and meta-analysis. *Hum Reprod Update.* 2012;18(6):618-637.
12. Brower M.A., Hai Y., Jones M.R., et al. Bidirectional Mendelian randomization to explore the causal relationships between body mass index and polycystic ovary syndrome. *Hum Reprod.* 2018;34(1): 127-136.
13. Zhao Y., Xu Y., Wang X., et al. Body mass index and polycystic ovary syndrome: A 2-sample bidirectional Mendelian randomization study. *Journal of Clinical Endocrinology and Metabolism.* 2020;105(6).
14. Larsson S.C., Bäck M., Rees J.M.B., Mason A.M., Burgess S. Body mass index and body composition in relation to 14 cardiovascular conditions in UK Biobank: a Mendelian randomization study. *Eur Heart J.* 2020;41(2):221-226.
15. Haase C.L., Lopes S., Olsen A.H., Satylganova A., Schnecke V., McEwan P. Weight loss and risk reduction of obesity-related outcomes in 0.5 million people: evidence from a UK primary care database. *International Journal of Obesity.* 2021;45(6):1249-1258.
16. Marzouk T.M., Sayed Ahmed W.A. Effect of Dietary Weight Loss on Menstrual Regularity in Obese Young Adult Women with Polycystic Ovary Syndrome. *J Pediatr Adolesc Gynecol.* 2015;28(6):457-461.
17. Azziz R., Sanchez L.A., Knochenhauer E.S., et al. Androgen excess in women: Experience with over 1000 consecutive patients. *Journal of Clinical Endocrinology and Metabolism.* 2004;89(2):453-462.
18. Livadas S., Pappas C., Karachalios A., et al. Prevalence and impact of hyperandrogenemia in 1,218 women with polycystic ovary syndrome. *Endocrine.* 2014;47(2):631-638.
19. Burger H.G. Androgen production in women. *Fertil Steril.* 2002;77:3-5.
20. Hammond G.L. Plasma steroid-binding proteins: primary gatekeepers of steroid hormone action. *J Endocrinol.* 2016;230(1):R13-25.

21. Ruth K.S., Day F.R., Tyrrell J., et al. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat Med.* 2020;26(2):252-258.
22. Jänne M., Hammond G.L. Hepatocyte nuclear factor-4 controls transcription from a TATA-less human sex hormone-binding globulin gene promoter. *J Biol Chem.* 1998;273(51):34105-34114.
23. Saez-Lopez C., Brianso-Llort L., Torres-Torronteras J., Simó R., Hammond G.L., Selva D.M. Resveratrol increases hepatic SHBG expression through human constitutive androstane receptor: A new contribution to the french paradox. *Sci Rep.* 2017;7(1):12284.
24. Selva D.M., Hammond G.L. Thyroid hormones act indirectly to increase sex hormone-binding globulin production by liver via hepatocyte nuclear factor-4alpha. *J Mol Endocrinol.* 2009;43(1):19-27.
25. Selva D.M., Hammond G.L. Peroxisome-proliferator receptor gamma represses hepatic sex hormone-binding globulin expression. *Endocrinology.* 2009;150(5):2183-2189.
26. Simó R., Saez-Lopez C., Lecube A., Hernandez C., Fort J.M., Selva D.M. Adiponectin upregulates SHBG production: molecular mechanisms and potential implications. *Endocrinology.* 2014;155(8):2820-2830.
27. Simó R., Barbosa-Desongles A., Lecube A., Hernandez C., Selva D.M. Potential role of tumor necrosis factor-α in downregulating sex hormone-binding globulin. *Diabetes.* 2012;61(2):372-382.
28. Simó R., Sáez-López C., Barbosa-Desongles A., Hernández C., Selva D.M. Novel insights in SHBG regulation and clinical implications. *Trends in Endocrinology and Metabolism.* 2015;26(7):376-383.
29. Selva D.M., Hogeveen K.N., Innis S.M., Hammond G.L. Monosaccharide-induced lipogenesis regulates the human hepatic sex hormone-binding globulin gene. *J Clin Invest.* 2007;117(12):3979-3987.
30. Bandsma R.H., Prinsen B.H., van Der Velden Mde S., et al. Increased de novo lipogenesis and delayed conversion of large VLDL into intermediate density lipoprotein particles contribute to hyperlipidemia in glycogen storage disease type 1a. *Pediatr Res.* 2008;63(6):702-707.
31. Kozlitina J., Smagris E., Stender S., et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet.* 2014;46(4):352-356.
32. Carta G., Murru E., Banni S., Manca C. Palmitic Acid: Physiological Role, Metabolism and Nutritional Implications. *Front Physiol.* 2017;8:902.
33. Joseph J., Shamburek R.D., Cochran E.K., Gorden P., Brown R.J. Lipid regulation in lipodystrophy versus the obesity-associated metabolic syndrome: The dissociation of HDL-C and triglycerides. *Journal of Clinical Endocrinology and Metabolism.* 2014;99(9):E1676-E1680.
34. Saez-Lopez C., Barbosa-Desongles A., Hernandez C., et al. Sex hormone-binding globulin reduction in metabolic disorders may play a role in NAFLD development. *Endocrinology.* 2017;158(3):545-559.
35. Niskanen L., Laaksonen D.E., Punnonen K., Mustajoki P., Kaukua J., Rissanen A. Changes in sex hormone-binding globulin and testosterone during weight loss and weight maintenance in abdominally obese men with the metabolic syndrome. *Diabetes Obes Metab.* 2004;6(3):208-215.
36. Telgenkamp I., Kusters Y.H.A.M., Schalkwijk C.G., et al. Contribution of Liver Fat to Weight Loss–Induced Changes in Serum Hepatokines: A Randomized Controlled Trial. *The Journal of Clinical Endocrinology & Metabolism.* 2019;104(7):2719-2727.
37. Kanaya A.M., Herrington D., Vittinghoff E., et al. Glycemic effects of postmenopausal hormone therapy: The Heart and Estrogen/progestin Replacement Study. A randomized, double-blind, placebo-controlled trial. *Ann Intern Med.* 2003;138(1):1-9.
38. Harrison S.A., Bashir M.R., Guy C.D., et al. Resmetirom (MGL-3196) for the treatment of non-alcoholic steatohepatitis: A multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet.* 2019;394(10213):2012-2024.
39. Ohlsson C., Wallaschofski H., Lunetta K.L., et al. Genetic determinants of serum testosterone concentrations in men. *PLoS Genet.* 2011;7(10):e1002313-e1002313.
40. Coviello A.D., Haring R., Wellons M., et al. A genome-wide association meta-analysis of circulating sex hormone–binding globulin reveals multiple loci implicated in sex steroid hormone regulation. *PLoS Genet.* 2012;8(7):e1002805.
41. Smith G.I., Shankaran M., Yoshino M., et al. Insulin resistance drives hepatic de novo lipogenesis in nonalcoholic fatty liver disease. *J Clin Invest.* 2020;130(3):1453-1460.
42. Lambert J.E., Ramos-Roman M.A., Browning J.D., Parks E.J. Increased de novo lipogenesis is a distinct characteristic of individuals with nonalcoholic fatty liver disease. *Gastroenterology.* 2014;146(3): 726-735.

43. Allen L.A., Shrikrishnapalasuriyar N., Aled Rees D. Long-term health outcomes in young women with polycystic ovary syndrome: a narrative review. *Clin Endocrinol (Oxf).* 2021.
44. Vassilatou E., Lafoyianni S., Vryonidou A., et al. Increased androgen bioavailability is associated with non-alcoholic fatty liver disease in women with polycystic ovary syndrome. *Hum Reprod.* 2010;25(1):212-220.
45. Lebbe M., Woodruff T.K. Involvement of androgens in ovarian health and disease. *Molecular Human Reproduction.* 2013;19(12):828-837.
46. Broekmans F.J., Visser J.A., Laven J.S., Broer S.L., Themmen A.P., Fauser B.C. Anti-Müllerian hormone and ovarian dysfunction. *Trends in Endocrinology and Metabolism.* 2008;19(9):340-347.
47. Chang H.M., Klausen C., Leung P.C. Antimüllerian hormone inhibits follicle-stimulating hormone-induced adenylyl cyclase activation, aromatase expression, and estradiol production in human granulosa-lutein cells. *Fertil Steril.* 2013;100(2):585-592.e581.
48. Azziz R., Carmina E., Chen Z., et al. Polycystic ovary syndrome. *Nature Reviews Disease Primers.* 2016;2(1):16057.
49. Walters K.A., Gilchrist R.B., Ledger W.L., Teede H.J., Handelsman D.J., Campbell R.E. New Perspectives on the Pathogenesis of PCOS: Neuroendocrine Origins. *Trends in Endocrinology and Metabolism.* 2018;29(12):841-852.
50. Rosenfield R.L., Ehrmann D.A. The pathogenesis of polycystic ovary syndrome (PCOS): The hypothesis of PCOS as functional ovarian hyperandrogenism revisited. *Endocr Rev.* 2016;37(5):467-520.
51. Baptiste C.G., Battista M.-C., Trottier A., Baillargeon J.-P. Insulin and hyperandrogenism in women with polycystic ovary syndrome. *Journal of Steroid Biochemistry and Molecular Biology*. 2010;122(1-3): 42-52.
52. Goodarzi M.O., Dumesic D.A., Chazenbalk G., Azziz R. Polycystic ovary syndrome: etiology, pathogenesis and diagnosis. *Nat Rev Endocrinol.* 2011;7(4):219-231.
53. Mantovani A., Petracca G., Beatrice G., Tilg H., Byrne C.D., Targher G. Non-alcoholic fatty liver disease and risk of incident diabetes mellitus: an updated meta-analysis of 501 022 adult individuals. *Gut.* 2021;70(5):962-969.
54. Naeem M., Bülow R., Schipf S., et al. Association of hepatic steatosis derived from ultrasound and quantitative MRI with prediabetes in the general population. *Sci Rep.* 2021;11(1):13276.
55. Chang Y., Jung H.S., Yun K.E., Cho J., Cho Y.K., Ryu S. Cohort study of non-alcoholic fatty liver disease, NAFLD fibrosis score, and the risk of incident diabetes in a Korean population. *Am J Gastroenterol.* 2013;108(12):1861-1868.
56. Tokita Y., Maejima Y., Shimomura K., et al. Non-alcoholic Fatty Liver Disease Is a Risk Factor for Type 2 Diabetes in Middle-aged Japanese Men and Women. *Intern Med.* 2017;56(7):763-771.
57. Shibata M., Kihara Y., Taguchi M., Tashiro M., Otsuki M. Nonalcoholic fatty liver disease is a risk factor for type 2 diabetes in middle-aged Japanese men. *Diabetes Care.* 2007;30(11):2940-2944.
58. Hodson L., Frayn K.N. Hepatic fatty acid partitioning. *Curr Opin Lipidol.* 2011;22(3).
59. Pramfalk C., Pavlides M., Banerjee R., et al. Fasting plasma insulin concentrations are associated with changes in hepatic fatty acid synthesis and partitioning prior to changes in liver fat content in healthy adults. *Diabetes.* 2016;65(7):1858-1867.
60. Stefan N., Häring H.U. The role of hepatokines in metabolism. *Nat Rev Endocrinol.* 2013;9(3):144-152.
61. Saez-Lopez C., Villena J.A., Simó R., Selva D.M. Sex hormone-binding globulin overexpression protects against high-fat diet-induced obesity in transgenic male mice. *J Nutr Biochem.* 2020;85:108480.
62. Targher G., Byrne C.D. Nonalcoholic fatty liver disease: A novel cardiometabolic risk factor for type 2 diabetes and its complications. *Journal of Clinical Endocrinology and Metabolism.* 2013;98(2):483-495.
63. Fan Y., Wolford B.N., Lu H., et al. Type 2 diabetes sex-specific effects associated with E167K coding variant in TM6SF2. *iScience.* 2021;24(11):103196.
64. Santoro N., Caprio S., Pierpont B., Van Name M., Savoye M., Parks E.J. Hepatic de novo lipogenesis in obese youth is modulated by a common variant in the GCKR gene. *Journal of Clinical Endocrinology and Metabolism.* 2015;100(8):E1125-E1132.
65. Sirota J.C., McFann K., Targher G., Chonchol M., Jalal D.I. Association between nonalcoholic liver disease and chronic kidney disease: an ultrasound analysis from NHANES 1988–1994. *Am J Nephrol.* 2012;36(5):466-471.

66. Pan L.L., Zhang H.J., Huang Z.F., et al. Intrahepatic triglyceride content is independently associated with chronic kidney disease in obese adults: A cross-sectional study. *Metabolism.* 2015;64(9):1077-1085.
67. Targher G., Bertolini L., Rodella S., et al. Non-alcoholic fatty liver disease is independently associated with an increased prevalence of chronic kidney disease and proliferative/laser-treated retinopathy in type 2 diabetic patients. *Diabetologia.* 2008;51(3):444-450.
68. Huh J.H., Kim J.Y., Choi E., Kim J.S., Chang Y., Sung K.-C. The fatty liver index as a predictor of incident chronic kidney disease in a 10-year prospective cohort study. *PLoS One.* 2017;12(7):e0180951.
69. Sinn D.H., Kang D., Jang H.R., et al. Development of chronic kidney disease in patients with non-alcoholic fatty liver disease: A cohort study. *J Hepatol.* 2017;67(6):1274-1280.
70. Mantovani A., Zaza G., Byrne C.D., et al. Nonalcoholic fatty liver disease increases risk of incident chronic kidney disease: A systematic review and meta-analysis. *Metabolism.* 2018;79:64-76.
71. Musso G., Gambino R., Tabibian J.H., et al. Association of non-alcoholic fatty liver disease with chronic kidney disease: a systematic review and meta-analysis. *PLoS Med.* 2014;11(7):e1001680.
72. Vaxillaire M., Cavalcanti-Proença C., Dechaume A., et al. The common P446L polymorphism in GCKR inversely modulates fasting glucose and triglyceride levels and reduces type 2 diabetes risk in the DESIR prospective general French population. *Diabetes.* 2008;57(8):2253-2257.
73. Peter A., Stefan N., Cegan A., et al. Hepatic glucokinase expression is associated with lipogenesis and fatty liver in humans. *Journal of Clinical Endocrinology and Metabolism.* 2011;96(7):E1126-E1130.
74. Gorman T., Hope D.C., Brownlie R., et al. Effect of high-fat diet on glucose homeostasis and gene expression in glucokinase knockout mice. *Diabetes Obes Metab.* 2008;10(10):885-897.
75. Gu Y., Mao Y., Li H., et al. Long-term renal changes in the liver-specific glucokinase knockout mouse: implications for renal disease in maturity-onset diabetes of the young 2. *Transl Res.* 2011;157(3): 111-116.
76. Targher G., Byrne C.D., Lonardo A., Zoppini G., Barbui C. Non-alcoholic fatty liver disease and risk of incident cardiovascular disease: A meta-analysis. *J Hepatol.* 2016;65(3):589-600.
77. Zhou Y.-Y., Zhou X.-D., Wu S.-J., et al. Nonalcoholic fatty liver disease contributes to subclinical atherosclerosis: A systematic review and meta-analysis. *Hepatology Communications.* 2018;2(4): 376-392.
78. Brouwers M.C.G.J., Simons N., Stehouwer C.D.A., Koek G.H., Schaper N.C., Isaacs A. Relationship between nonalcoholic fatty liver disease susceptibility genes and coronary artery disease. *Hepatology Communications.* 2019;3(4):587-596.
79. Willer C.J., Schmidt E.M., Sengupta S., et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013;45(11):1274-1283.
80. Wang Y., Viscarra J., Kim S.-J., Sul H.S. Transcriptional regulation of hepatic lipogenesis. *Nature reviews Molecular cell biology.* 2015;16(11):678-689.
81. Marston N.A., Giugliano R.P., Im K., et al. Association between triglyceride lowering and reduction of cardiovascular risk across multiple lipid-lowering therapeutic classes. *Circulation.* 2019;140(16): 1308-1317.
82. Allara E., Morani G., Carter P., et al. Genetic determinants of lipids and cardiovascular disease outcomes. *Circulation: Genomic and Precision Medicine.* 2019;12(12):e002711.
83. Do R., Willer C.J., Schmidt E.M., et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet.* 2013;45(11):1345-1352.
84. Li Y., Si S., Hou L., et al. Causal effect of sex hormone-binding globulin and testosterone on coronary heart disease: A multivariable and network Mendelian randomization analysis. *Int J Cardiol.* 2021.
85. Kurnaz-Gomleksiz O., Akadam-Teker B., Bugra Z., Omer B., Yilmaz-Aydogan H. Genetic polymorphisms of the SHBG gene can be the effect on SHBG and HDL-cholesterol levels in Coronary Heart Disease: a case-control study. *Mol Biol Rep.* 2019;46(4):4259-4269.
86. Alevizaki M., Saltiki K., Xita N., et al. The importance of the (TAAAA)n alleles at the SHBG gene promoter for the severity of coronary artery disease in postmenopausal women. *Menopause.* 2008;15(3):461-468.
87. Zhao J.V., Schooling C.M. Genetically Predicted Sex Hormone Binding Globulin and Ischemic Heart Disease: A Sex-Specific Mendelian Randomization Study. *Available at http://dxdoiorg/102139/ssrn3803355`.* 2021.

88. Sáez-López C., Salcedo-Allende M.T., Hernandez C., Simó-Servat O., Simó R., Selva D.M. Sex hormone-binding globulin expression correlates with acetyl-coenzyme a carboxylase and triglyceride content in human liver. *J Clin Endocrinol Metab.* 2019;104(5):1500-1507.
89. Delgado-Rodríguez M., Llorca J. Bias. *Journal of Epidemiology and Community Health.* 2004;58(8): 635-641.
90. Mansournia M.A., Jewell N.P., Greenland S. Case-control matching: effects, misconceptions, and recommendations. *Eur J Epidemiol.* 2018;33(1):5-14.
91. Lu H., Cole S.R., Platt R.W., Schisterman E.F. Revisiting Overadjustment Bias. *Epidemiology.* 2021;32(5).
92. Schisterman E.F., Cole S.R., Platt R.W. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology.* 2009;20(4):488-495.
93. Davey Smith G., Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet.* 2014;23(R1):R89-R98.
94. Hemani G., Bowden J., Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet.* 2018;27(R2):R195-R208.
95. Burgess S., Timpson N.J., Ebrahim S., Davey Smith G. Mendelian randomization: where are we now and where are we going? *Int J Epidemiol.* 2015;44(2):379-388.
96. Haycock P.C., Burgess S., Wade K.H., Bowden J., Relton C., Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr*. 2016;103(4):965-978.
97. Sanderson E., Spiller W., Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat Med.* 2021.
98. Smith G.D., Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol.* 2004;33(1):30-42.
99. Pierce B.L., Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol.* 2013;178(7):1177-1184.
100. Burgess S., Thompson S.G. *Mendelian randomization: Methods for using genetic variants in causal estimation.* London, UK: Chapman & Hall/CRC Press; 2015.
101. Gratten J., Visscher P.M. Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome Med.* 2016;8(1):78.
102. Manolio T.A., Collins F.S., Cox N.J., et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747-753

11

CHAPTER ELEVEN

Impact paragraph

Polycystic ovary syndrome (PCOS) is a condition that is characterized by an irregular menstrual cycle, hyperandrogenism and a polycystic ovarian morphology. In addition, women with PCOS are often affected by subfertility, obesity, type 2 diabetes, dyslipidaemia and coronary artery disease[1]. PCOS is a very common hormonal disorder, as approximately one in every ten premenopausal women is affected[2]. Despite extensive research thus far, the pathophysiology of PCOS remains poorly understood. Particularly in light of the long-term cardiometabolic complications, it is of importance to better unravel the pathophysiology of PCOS and identify potential therapeutic targets.

In this thesis we found that PCOS by itself does not seem to be causal in increasing the risk of coronary artery disease. Rather, a dysregulation of de novo lipogenesis – the process of converting glucose or fructose into lipids which occurs primarily in the liver and is a downstream consequence of obesity[3,4] – seems to be the common factor predisposing to an increased risk of PCOS and coronary artery disease. Moreover, we identified that de novo lipogenesis links to PCOS and other cardiometabolic disorders by decreasing serum sex hormone-binding globulin (SHBG) levels.

This chapter describes how the findings of this thesis may impact scientific research and clinical practice.

PCOS and hyperandrogenism: a more personalised approach?

The results described in this thesis suggest that not all women with PCOS are at risk for developing cardiometabolic disease. Previously, it has been observed that cardiometabolic features in women with PCOS cluster primarily in those women with hyperandrogenism[5]. According to the 2018 EHSRE guidelines for the diagnosis and management of women with PCOS, the definition of biochemical hyperandrogenism is based on the levels of free testosterone or the free androgen index[6]. Our studies suggest that a more extensive phenotyping of hyperandrogenism in women with PCOS, by additionally considering the determinants of free testosterone (i.e. serum SHBG and total testosterone), may guide clinicians in better understanding the primary pathway that contributes to the development of hyperandrogenism.

In light of the findings presented in this thesis, it is hypothesized that in particular women with PCOS and low serum SHBG levels are at risk of developing cardiometabolic complications. However, it will require further study to assess whether serum SHBG is a good indicator of individual cardiometabolic risk. Although we observe that de novo lipogenesis is one of the pathways that regulates serum SHBG levels, and as such serum SHBG may be an indicator of metabolic dysfunction, there are also many other

factors that contribute to the regulation of serum SHBG levels within an individual[7]. It would, therefore, be of interest to study the predictive value of serum SHBG as a biomarker of metabolic dysfunction and, by extension, as a prognostic marker for cardiometabolic risk in women with PCOS.

The current guidelines regarding the management of cardiometabolic dysfunction in women with PCOS advise screening for obesity, dyslipidaemia, hypertension, impaired glucose tolerance and lifestyle factors (i.e. lack of physical activity and smoking)[6]. Compared to these guidelines, serum SHBG levels may be an early biomarker of metabolic dysfunction, well before signs of dyslipidaemia, hypertension or impaired glucose tolerance are present. Although obesity will likely remain the most prognostic, non-invasive marker for cardiometabolic risk, recent genetic studies have identified that there are distinct adiposity clusters that associate with favourable and unfavourable metabolic effects[8]. The latter cluster was found to associate with higher non-alcoholic fatty liver disease (NAFLD) risk, lower serum SHBG levels, and higher PCOS risk[8]. Although it is likely that unfavourable adiposity represents the majority of obese patients seen in clinical practice, it would be of interest to study whether serum SHBG may aid in further delineating the individual cardiometabolic risk of women with PCOS.

Not only clinical practice, but also scientific research may benefit from a better phenotyping of women with hyperandrogenic-PCOS. Different (hyperandrogenic) phenotypes within PCOS may be characterized by vastly different underlying pathophysiological mechanisms[9]. A better characterization of PCOS patients in scientific research would likely allow researchers to gain a better understanding the pathophysiology in PCOS as a whole and within subgroups of PCOS patients.

De novo lipogenesis as a potential therapeutic target

In this thesis, dysregulation of DNL seems to predispose to cardiometabolic disorders including type 2 diabetes and coronary artery disease in both women and men. Nevertheless, we were unable to study the extent to which de novo lipogenesis contributes to the risk of cardiometabolic disease. Therefore, the findings in this thesis justify further research assessing whether de novo lipogenesis may be a therapeutic target through which to reduce the risk of cardiometabolic disease.

Several avenues through which de novo lipogenesis could be reduced, and consequently serum SHBG could be increased, are under investigation. First, as de novo lipogenesis is highly associated with obesity, lifestyle interventions that achieve weight-loss are likely to be a successful approach[3]. Indeed, intervention trials have

shown that a weight reduction of ~10 kg has been shown to reduce intrahepatic lipid (IHL) content (~1.69%) and increase serum SHBG levels (~26%)[10]. Second, thyroid receptor hormone beta agonists, such as resmetirom which has a liver-specific profile, mimic the beneficial effects of thyroid hormones on de novo lipogenesis[11,12]. In phase II clinical trials it has been observed that resmetirom treatment resulted in a relative reduction in IHL content by ~50% and an increase in serum SHBG levels by ~116%[13]. Third, acetyl-coenzyme A carboxylase (ACC) inhibitors, which prevent the conversion of acetyl-CoA to malonyl-CoA, i.e. the first step in de novo lipogenesis, are currently undergoing clinical trials. In these trials 16 weeks of treatment with a high-dose ACC inhibitor resulted in a relative reduction in IHL content by ~65%[14]. As of yet, the effects of ACC inhibitors on serum SHBG have not been reported.

Although the initial effects of the hitherto presented interventions on IHL content and serum SHBG are beneficial[13,14], the exact clinical relevance of reducing de novo lipogenesis on hard clinical end points is yet to be investigated. The clinical trials report promising results regarding the effect of these interventions on glucose metabolism and lipid profile[13,14]. However, specifically with regard to PCOS, the effect of reducing de novo lipogenesis remains uncertain. In chapter four, de novo lipogenesis in women ranged from 1.3% to 24.5%. Based on the data in this thesis, we can extrapolate that a 10 percent point decrease in de novo lipogenesis is expected to result in ~11 nmol/l increase in serum SHBG (chapter four). Relative to the average serum SHBG levels in women with PCOS this would be a 36% increase in serum SHBG (chapter three). Consequently, this would result in ~4 pmol/l or 19% decrease in serum free testosterone levels in women with PCOS. The effect of such interventions on the phenotype of women with PCOS deserves further investigation.

Future investigation into the role of SHBG in cardiometabolic disease

It is increasingly recognized that serum SHBG has a greater role in metabolic disorders than it has been given credit for. As indicated previously, serum SHBG may be a biomarker of metabolic dysfunction, in particular de novo lipogenesis. However, serum SHBG may also be involved in the pathophysiological processes of metabolic disease including PCOS and type 2 diabetes, and initial studies indicate it may also contribute to the risk of NAFLD and coronary artery disease[15,16]. It deserves further study to assess the extent to which serum SHBG acts as a causal factor in these disorders, and whether SHBG may be a potential future target for intervening in the cardiometabolic risk profile of patients.

Conclusion

In conclusion, this chapter illustrates ways in which the findings presented in this thesis may guide future research and, ultimately, clinical practice. In particular, the current findings further our understanding on the pathophysiology that underlies PCOS and cardiometabolic disease, and highlights the role of de novo lipogenesis and serum SHBG herein. This implicates serum SHBG as a potential screening tool and de novo lipogenesis as a potential therapeutic target that may, in the long term, help to improve the long-term cardiometabolic well-being in women with PCOS. This thesis also highlights the need for a more personalised approach in the study of PCOS, as this may help to further unravel the relationship between PCOS, cardiometabolic disease and de novo lipogenesis.

References

1. Azziz R., Carmina E., Chen Z., et al. Polycystic ovary syndrome. *Nature Reviews Disease Primers*. 2016;2(1):16057.
2. Wolf W.M., Wattick R.A., Kinkade O.N., Olfert M.D. Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity. *International Journal of Environmental Research and Public Health*. 2018;15(11):2589.
3. Smith G.I., Shankaran M., Yoshino M., et al. Insulin resistance drives hepatic de novo lipogenesis in nonalcoholic fatty liver disease. *J Clin Invest*. 2020;130(3):1453-1460.
4. Nguyen P., Leray V., Diez M., et al. Liver lipid metabolism. *Journal of Animal Physiology and Animal Nutrition*. 2008;92(3):272-283.
5. Daan N.M.P., Louwers Y.V., Koster M.P.H., et al. Cardiovascular and metabolic profiles amongst different polycystic ovary syndrome phenotypes: who is really at risk? *Fertil Steril*. 2014;102(5):1444-1451.e1443.
6. Teede H.J., Misso M.L., Costello M.F., et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Hum Reprod*. 2018;33(9):1602-1618.
7. Simó R., Sáez-López C., Barbosa-Desongles A., Hernández C., Selva D.M. Novel insights in SHBG regulation and clinical implications. *Trends in Endocrinology and Metabolism*. 2015;26(7):376-383.
8. Martin S., Cule M., Basty N., et al. Genetic Evidence for Different Adiposity Phenotypes and Their Opposing Influences on Ectopic Fat and Risk of Cardiometabolic Disease. *Diabetes*. 2021;70(8): 1843-1856.
9. Lizneva D., Suturina L., Walker W., Brakta S., Gavrilova-Jordan L., Azziz R. Criteria, prevalence, and phenotypes of polycystic ovary syndrome. *Fertil Steril*. 2016;106(1):6-15.
10. Telgenkamp I., Kusters Y.H.A.M., Schalkwijk C.G., et al. Contribution of Liver Fat to Weight Loss–Induced Changes in Serum Hepatokines: A Randomized Controlled Trial. *The Journal of Clinical Endocrinology & Metabolism*. 2019;104(7):2719-2727.
11. Sinha R.A., Bruinstroop E., Singh B.K., Yen P.M. Nonalcoholic Fatty Liver Disease and Hypercholesterolemia: Roles of Thyroid Hormones, Metabolites, and Agonists. *Thyroid*. 2019;29(9):1173-1191.
12. Selva D.M., Hammond G.L. Thyroid hormones act indirectly to increase sex hormone-binding globulin production by liver via hepatocyte nuclear factor-4alpha. *J Mol Endocrinol*. 2009;43(1):19-27.
13. Harrison S.A., Bashir M.R., Guy C.D., et al. Resmetirom (MGL-3196) for the treatment of non-alcoholic steatohepatitis: A multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet*. 2019;394(10213):2012-2024.
14. Calle R.A., Amin N.B., Carvajal-Gonzalez S., et al. ACC inhibitor alone or co-administered with a DGAT2 inhibitor in patients with non-alcoholic fatty liver disease: two parallel, placebo-controlled, randomized phase 2a trials. *Nat Med*. 2021.
15. Saez-Lopez C., Barbosa-Desongles A., Hernandez C., et al. Sex hormone-binding globulin reduction in metabolic disorders may play a role in NAFLD development. *Endocrinology*. 2017;158(3):545-559.
16. Zhao J.V., Schooling C.M. Genetically predicted sex hormone binding globulin and ischemic heart disease in men and women: a univariable and multivariable Mendelian randomization study. *Sci Rep*. 2021;11(1):23172.

A

ADDENDUM

Nederlandse samenvatting

Dankwoord

Curriculum vitea

Scientific output

Ontrafelen van de driehoeksrelatie tussen het polycysteus ovarium syndroom, cardio-metabole aandoeningen en de novo lipogenese

Het antwoord bevindt zich in de genen

Introductie

Polycysteus ovarium syndroom (PCOS) is een van de meest voorkomende hormonale aandoeningen onder vrouwen in de vruchtbare levensfase. Het is een aandoening die wordt gekenmerkt door een onregelmatige menstruatiecyclus, hoge waarden van mannelijke hormonen (androgenen) en blaasjes in de eierstokken (een polycysteus ovarium). Vrouwen met PCOS ontwikkelen op latere leeftijd vaker cardio-metabole ziekten (waaronder type 2 diabetes – ook wel bekend als ouderdomssuikerziekte – en hart- en vaatziekten) dan vrouwen zonder PCOS. Het is onduidelijk waarom vrouwen met PCOS meer risico hebben op het ontwikkelen van cardio-metabole ziekten. Enerzijds kan dit het gevolg zijn van PCOS zelf (een oorzakelijk verband), anderzijds kan het ook het gevolg zijn van een andere onderliggende factor die zowel leidt tot het ontwikkelen van PCOS als het ontwikkelen van cardio-metabole ziekten.

Eén van de factoren die zou kunnen bijdragen aan het ontstaan van zowel PCOS als cardio-metabole ziekten is de novo lipogenese. De novo lipogenese is een proces waarbij suikers worden omgezet in vetten, en dit vindt voornamelijk plaats in de lever. De novo lipogenese heeft een belangrijke bijdrage in het ontwikkelen van leververvetting, maar het kan eveneens ook allerlei andere negatieve gezondheids-gevolgen hebben.

In dit proefschrift hebben we geprobeerd om de relatie tussen PCOS en cardio-metabole ziekten te ontrafelen. Tevens hebben we geprobeerd om de invloed van de novo lipogenese op zowel PCOS als cardio-metabole ziekten te onderzoeken en de processen die hieraan ten grondslag liggen te bestuderen.

Onderzoeksmethoden

De relaties die we in dit proefschrift beoogde te onderzoeken kunnen op verschillende manieren bestudeerd worden. In dit proefschrift hebben we gebruik gemaakt van observationeel onderzoek en Mendeliaanse randomisatie studies om (oorzakelijke) verbanden te onderzoeken.

Bij observationeel onderzoek worden de deelnemers van het onderzoek geobserveerd, maar wordt er geen interventie uitgevoerd. Observationeel onderzoek levert een waardevolle bijdrage aan de wetenschap, maar op basis van observationeel onderzoek kunnen geen uitspraken worden gedaan over oorzakelijke verbanden.

Om te onderzoeken of een verband oorzakelijk is, hebben we in dit proefschrift de Mendeliaanse randomisatie onderzoekstechniek toegepast. Mendeliaanse randomisatie is gebaseerd op het idee dat er binnen een populatie een grote overeenkomst is in het genetisch profiel van mensen, maar ook een klein deel dat varieert. Als gevolg van deze variatie hebben sommige mensen genen die leiden tot een hoger risico op het ontwikkelen van een bepaalde ziekte (zoals bijvoorbeeld PCOS) terwijl andere mensen genen hebben die beschermen tegen het ontwikkelen van PCOS. Door deze genetische variatie te bestuderen in relatie tot PCOS en hart- en vaatziekten is het mogelijk om te onderzoeken of het hebben van PCOS leidt tot het ontwikkelen van hart- en vaatziekten (een oorzakelijk verband).

Belangrijkste bevindingen

Deel 1: De relatie tussen PCOS en (risicofactoren) van hart- en vaatziekten

In het eerste deel van dit proefschrift beoogde we te onderzoeken of PCOS leidt tot het ontwikkelen van hart- en vaatziekten.

In **hoofdstuk twee** van dit proefschrift zagen we dat genen die het risico op PCOS verhogen niet samengaan met een verhoogd risico op hart- en vaatziekten. Daarmee lijkt PCOS op zichzelf niet te leiden tot het ontwikkelen van hart- en vaatziekten. Desalniettemin hebben eerdere studies aangetoond dat vrouwen met PCOS toch vaker hart- en vaatziekten hebben dan vrouwen zonder PCOS. Er dient dus een andere verklarende factor te zijn voor dit verband. Vrouwen met PCOS worden vaak gekenmerkt door overgewicht of obesitas, en obesitas zou een mogelijk verklarende factor kunnen zijn die bijdraagt aan het ontstaan van zowel PCOS als hart- en vaatziekten. In hoofdstuk twee hebben we daarom ook onderzocht of genen die leiden tot obesitas ook samengaan met het ontwikkelen van PCOS en hart- en vaatziekten. De resultaten van deze analyses suggereren inderdaad dat obesitas een oorzakelijk verband heeft met zowel PCOS als hart- en vaatziekten.

Het mechanisme waarmee obesitas leidt tot het ontwikkelen van PCOS is waarschijnlijk het gevolg van diverse factoren. Te hoge waarden van mannelijke hormonen (androgenen) zou mogelijk een rol kunnen spelen. Daarom hebben we in **hoofdstuk**

drie de relatie tussen symptomen van PCOS en drie androgeen waarden onderzocht, specifiek de androgenen vrij testosteron, totaal testosteron en seks hormoon bindend globuline [SHBG] – het eiwit wat testosteron bindt en transporteert in het lichaam. In een grote groep vrouwen met PCOS zagen we dat lage SHBG waarden samengingen met een verstoorde vethuishouding, verstoorde suikerhuishouding en verhoogd lichaamsgewicht. Daarentegen gingen hoge totaal testosteron waarden vooral samen met gynaecologische kenmerken waaronder de hoeveelheid blaasjes in de eierstokken en anti-Müllerian hormoon (een weerspiegeling van de hoeveelheid eicellen). Dit suggereert dat de metabole problematiek die gezien wordt in PCOS vooral samengaat met lage SHBG waarden, en dat SHBG mogelijk een rol speelt in de relatie tussen obesitas en PCOS.

Deel 2: De relatie tussen de novo lipogenese en SHBG

Om nader te onderzoeken of SHBG inderdaad een rol speelt in de relatie tussen obesitas en PCOS, hebben we vervolgens onderzocht hoe obesitas leidt tot een verandering van SHBG waarden. Eerdere experimentele studies in muizen en cellen hebben laten zien dat de novo lipogenese één van de factoren is die bijdraagt aan de regulatie van de hoeveelheid SHBG dat wordt geproduceerd. Ook is bekend dat in mensen met obesitas het proces van de novo lipogenese verhoogd is. Daarmee zou ook in mensen de novo lipogenese wellicht één van de factoren kunnen zijn die bijdraagt aan de regulatie van SHBG en, daarmee, aan het risico op PCOS. Echter, tot nu toe was het effect van de novo lipogenese op SHBG enkel in muizen bestudeerd. Daarom hebben we in deel twee van dit proefschrift op diverse manieren bestudeerd of de novo lipogenese betrokken is bij de regulatie van SHBG in mensen.

In hoofdstuk vier hebben we in een groep mannen en vrouwen met obesitas en leververvetting gekeken of de hoeveelheid de novo lipogenese samengaat met SHBG waarden. In vrouwen zagen we inderdaad dat meer de novo lipogenese samengaat met lagere SHBG waarden. In mannen werd dit echter niet gezien. Dit kan een teken zijn dat dit proces in mannen minder of niet aanwezig is, of het kan een teken zijn dat de groep mannen die we hebben bestudeerd niet groot genoeg was om een verband te ontdekken.

Ondanks dat we zien dat de novo lipogenese samengaat met SHBG waarden in vrouwen, hoeft dit geen oorzakelijk verband aan te duiden. Daarom hebben we in hoofdstuk vijf onderzocht of de novo lipogenese een oorzakelijk effect heeft op SHBG. Dit hebben we onderzocht door patiënten met een zeldzame stofwisselingsziekte (glycogeen stapelingsziekte type 1a; GSD1a) te bestuderen. Deze patiënten hebben een genetische afwijking waardoor ze een verhoogde de novo lipogenese hebben, en daardoor zijn patiënten met GSD1a een goed model om het effect van de novo

lipogenese op SHBG te bestuderen. We zagen dat GSD1a patiënten lagere SHBG waarden hadden dan gezonde controles, wat suggereert dat het effect van de novo lipogenese op SHBG waarschijnlijk berust op een oorzakelijk verband.

In **hoofdstuk zes** hebben we de relatie tussen de novo lipogenese en SHBG bestudeerd middels genetica. Een eerder gepubliceerde grootschalige genetische studie heeft alle genen die betrokken zijn bij de regulatie van SHBG geïdentificeerd. In dit proefschrift hebben we al deze genen uitgebreid bestudeerd, waarbij het opviel dat er veel genen werden gevonden die betrokken zijn bij de regulatie van de novo lipogenese. Dat suggereert dat de novo lipogenese inderdaad een oorzakelijk verband lijkt te hebben met SHBG en betrokken is bij de regulatie van SHBG.

Deel 3: De relatie tussen de novo lipogenese, cardio-metabole ziekten en SHBG

Tot slot hebben we in het derde deel van dit proefschrift bestudeerd of de novo lipogenese betrokken is bij het ontwikkelen van diverse cardio-metabole ziekten waaronder PCOS, type 2 diabetes en hart- en vaatziekten. Tevens hebben we onderzocht of deze relaties via SHBG lopen.

In **hoofdstuk zes** hebben we onderzocht of Mendeliaanse randomisatie studies die eerder door andere onderzoekers zijn uitgevoerd ons informatie kunnen geven over de rol van SHBG in type 2 diabetes en PCOS. De bevindingen van deze eerdere studie suggereren dat SHBG een oorzakelijke relatie heeft met PCOS en type 2 diabetes. Bovendien zien we dat de relatie tussen SHBG en type 2 diabetes het gevolg is van SHBG zelf, terwijl de relatie tussen SHBG en PCOS via het effect van SHBG op vrij testosteron loopt.

In **hoofdstuk zeven** hebben we bestudeerd of het effect van leververvetting op type 2 diabetes via SHBG verloopt. Dit hebben we gedaan door een mediatie analyse uit te voeren. Daarmee bestuderen we of, en in hoeverre, het effect van leververvetting (wat het gevolg is van de novo lipogenese) op type 2 diabetes via SHBG loopt. In vrouwen blijkt ongeveer 51% van de relatie tussen leververvetting en type 2 diabetes verklaard te worden door SHBG. In mannen wordt 17% van de relatie tussen leververvetting en type 2 diabetes verklaard door SHBG. Het verschil in de bijdrage van SHBG in de relatie tussen leververvetting en type 2 diabetes in mannen en vrouwen is opvallend, maar de oorzaak voor dit verschil blijft helaas onduidelijk. In de toekomst zou verder onderzoek nodig zijn om deze geslachtsverschillen, die we ook in andere studies zagen (bijvoorbeeld hoofdstuk vier), nader te bestuderen.

Leververvetting is een belangrijke risico factor voor het ontwikkelen van hart- en vaatziekten. Of de novo lipogenese – wat een van de belangrijkste processen is dat bijdraagt aan het ontstaan van leververvetting – ook een oorzakelijk effect heeft op hart- en vaatziekten hebben we in hoofdstuk acht en negen onderzocht. In **hoofdstuk acht** hebben we daarvoor gebruik gemaakt van een variant in het *GCKR* gen. Een eerdere studie heeft aangetoond dat deze variant in het *GCKR* gen samen gaat met meer de novo lipogenese. In hoofdstuk acht hebben we daarom alle studies die kijken naar de relatie tussen deze variatie in het *GCKR* gen en hart- en vaatziekten bij elkaar gevoegd (een meta-analyse; waarmee het netto effect van alle studies bij elkaar opgeteld onderzocht wordt). In deze meta-analyse bleken mensen met de variant in het *GCKR* gen – wat leidde tot meer de novo lipogenese – vaker hart- en vaatziekten te hebben dan mensen met de variant in het *GCKR* gen dat samen ging met minder de novo lipogenese. Dit suggereert dat de novo lipogenese inderdaad oorzakelijk betrokken is bij het ontstaan van hart- en vaatziekten. *GCKR* is echter slechts één van de vele genen die bijdraagt aan de novo lipogenese. Daarom hebben we vervolgens in **hoofdstuk negen** op drie verschillende manieren diverse andere genen geïdentificeerd die ook effect hebben op de novo lipogenese. De gevonden genen – die naar alle waarschijnlijkheid leiden tot meer de novo lipogenese – gingen ook samen met meer hart- en vaatziekten. In conclusie suggereren de bevindingen in hoofdstuk acht en negen dat de novo lipogenese genen een oorzakelijk verband hebben met hart- en vaatziekten.

Conclusie

Concluderend hebben we in de studies die in dit proefschrift gebundeld zijn gevonden dat PCOS geen oorzakelijk verband lijkt te hebben met hart- en vaatziekten. Obesitas, en specifieker de novo lipogenese, lijkt een oorzakelijk verband te hebben met PCOS en andere cardio-metabole aandoeningen, waaronder hart- en vaatziekten en type 2 diabetes. SHBG is waarschijnlijk een van de eiwitten die voor een deel de relatie tussen de novo lipogenese en zowel PCOS als type 2 diabetes faciliteert.

Alhoewel dit proefschrift diverse (oorzakelijke) verbanden heeft onderzocht, dienen de resultaten nader bestudeerd te worden in vervolg onderzoek. Dit heeft meerdere redenen. Allereerst is PCOS een zeer divers ziektebeeld, waarbij patiënten worden gekenmerkt door een breed scala aan klachten. In dit proefschrift zien we geen oorzakelijk verband tussen PCOS en hart- en vaatziekten, echter zouden toekomstige studies nodig zijn om te onderzoeken of dit voor alle subtypes van PCOS patiënten geldt. Ten tweede hebben we in dit proefschrift gebruik gemaakt van genetica om oorzakelijke verbanden te bestuderen. De gouden standaard voor het bevestigen van een oorzakelijk verband is het uitvoeren van een interventie studie. Daarom zouden

idealiter deze relaties nader bestudeerd moeten worden in een interventie studie om nog meer bewijs voor een oorzakelijk verband tussen de novo lipogenese en cardio-metabole aandoeningen te vinden.

De bevindingen in dit proefschrift kunnen als startpunt dienen voor verder onderzoek en daardoor uiteindelijk een toevoeging zijn aan de huidige zorg richtlijnen. SHBG zou mogelijk, in de toekomst, een waardevolle toevoeging kunnen zijn op het protocol voor de screening voor cardio-metabole ziekten bij vrouwen met PCOS. De vroegtijdige opsporing van een verhoogd cardio-metabool risico kan een waardevolle bijdrage zijn voor deze patiëntengroep. Verder suggereren de bevindingen in dit proefschrift ook dat het de novo lipogenese een mogelijk aangrijpingspunt is voor het verbeteren van het cardio-metabole welzijn van mensen. Daarom zouden toekomstige studies kunnen onderzoeken of leefstijl interventies of medicamenten die de novo lipogenese verlagen op het lange termijn gunstig bijdragen aan het risico op de ontwikkelen van cardio-metabole aandoeningen.

Dankwoord

Tijd is een bijzonder fenomeen: als je het naar je zin hebt dan vliegt het voorbij. De afgelopen jaren zijn dan ook voorbij gegaan voordat ik het goed en wel doorhad. Dat is een teken van het plezier waarmee ik de afgelopen jaren heb mogen werken aan dit proefschrift, een teken van een hele leerzame periode en een teken van de gezelligheid van de vrienden, familie en collega's met wie ik deze ervaring heb mogen delen. In dit laatste hoofdstuk wil ik graag een groot woord van dank uitspreken aan iedereen die heeft bijgedragen aan de totstandkoming van dit proefschrift.

Allereerst aan mijn promotieteam:
Geachte **prof. dr. Brouwers**, beste Martijn. Ik prijs mezelf erg bevoorrecht dat ik mijn proefschrift heb mogen schrijven onder jouw supervisie. Ik heb veel geleerd van jouw kennis en kunde, werkethiek en enthousiasme voor het onderzoek. De hoeveelheid ideeën die je op willekeurig moment te binnen springen, en die je vervolgens met heel veel enthousiasme deelt, is bewonderenswaardig. Je benoemde ooit dat kwaliteit blijkt uit het werk dat iemand verzet en niet uit hoe hard iemand roept. Inmiddels is mij duidelijk dat dit typerend is voor jouw werkhouding, want alhoewel je heel veel kan praten – een werkoverleg dat op tijd werd afgerond was eerder een uitzondering dan de regel – ben je altijd kritisch over de onderzoeksresultaten en blijken je kwaliteiten vooral uit je werkethiek en het resultaat dat daaruit voortvloeit. Ondanks dat je altijd helder voor ogen hebt hoe je jouw onderzoekslijn wil voortzetten heb je mij de ruimte geboden om, op mijn eigen manier, mijzelf op wetenschappelijk vlak maar ook als persoon te ontwikkelen de afgelopen jaren. Ik wil je bedanken dat je me de kans(en) hebt geboden om dit promotietraject te mogen uitvoeren.
Geachte **dr. Valkenburg**, beste Olivier. In mijn promotieteam zorg jij voor het gynaecologisch perspectief en de klinische input. Tijdens de overlegmomenten bracht je een nuchtere blik op een wellicht wat internistische manier van overleggen. Daarin heb je me regelmatig herinnerd aan het belang van een klinische terugkoppeling in de manuscripten. En alhoewel de gynaecologie niet voor mij is weggelegd heb ik grote bewondering voor jouw vakgebied – niet tenminste omdat (tot mijn frustratie) PCOS mij nog altijd een mysterieus en onvoldoende begrepen ziekte is gebleven en het behandelen hiervan me daarom ook een grote uitdaging lijkt.
Geachte **prof. dr. Stehouwer**, beste Coen. Uw snelle en kritische blik op alle manuscripten was van grote waarden in mijn promotietraject. Terugkijkend op de overlegmomenten met het promotieteam staat vooral uw rustige observerende aanwezigheid op de voorgrond, waarbij u aanschouwt hoe het er voor staat en hulp aanbied waar nodig. Daarmee biedt u de ruimte en vrijheid om binnen het onderzoek en het promotietraject een eigen route uit te stippelen. Tijdens de schrijfcursus haalde u een gezegde van Blaise Pascal aan om te illustreren dat de moeilijkheid van het schrijven zit in de bondigheid van de tekst. Dit is een van de (overigens vele) goede

adviezen die ik in de toekomst (na dit dankwoord wat niet perse voldoet aan dit advies) zeker ga meenemen.

Achter iedere cijfer in dit proefschrift schuilen velen deelnemers die elk met hun eigen verhaal en achtergrond bereid waren om bij te dragen aan de ontwikkeling van kennis. Ik hoop dat de studies die in dit proefschrift beschreven staan, en die dankzij de vrijgevigheid van al deze deelnemers tot stand zijn gekomen, een heel klein steentje kunnen bijdragen aan meer kennis en begrip over het functioneren en disfunctioneren van het menselijk lichaam.

Geachte leden van de beoordelingscommissie, prof. dr. Spaanderman, prof. dr. Laven, dr. Meex, prof. dr. Snieder en Prof. dr. Zeegers, graag wil ik u bedanken voor uw tijd en bereidheid om mijn proefschrift te lezen en tijdens de verdediging daarover van gedachte te wisselen. Prof. dr. Lie Fong en dr. Peeters wil ik graag bedanken voor de bereidheid om plaats te nemen in de oppositie.

To all the co-authors of the various chapters, thank you for the collaboration and your valuable input in the writing of the manuscripts.

De afgelopen jaren is thuiswerken de norm geweest. Juist daardoor is mij nog duidelijker geworden hoe belangrijk collega's zijn.
Aan mijn kamergenoot Evelien, dat thuiswerkadvies is misschien toch nog ergens goed voor, want zonder dat hadden wij alleen maar gekletst en gelachen en was er van dit proefschrift nooit iets terecht gekomen. Dat 'klets-luik' tussen onze computerschermen was het beste idee ooit!
Armand en Evelien, het circulatielab is dankzij jullie de plek voor een goede debriefing of een gezellig gesprek. Bedankt dat ik daar regelmatig even langs mocht lopen.
Amée en Kim, jullie oprechte interesse, gezelligheid en positiviteit leidde ertoe dat we regelmatig in het patiëntenrestaurant te vinden waren om bij te kletsen onder het genot van een koffietje. Ik ben blij dat we ook post-UM tijd deze koffietjes, maar zeker ook de filmavonden en wandelingen voort zetten!
Nynke, jij hebt me wegwijs gemaakt in het onderzoek en op de afdeling. Ik ben dankbaar dat ik van jou heb mogen leren en jouw kennis en kunde als basis voor mijn onderzoek heb mogen gebruiken.
To all my colleagues who make the department such a "gezellige" place to work – April, Dija, Elena, Eline, Frank, Ine, Kai, Lise, Matthias, Margarita, Myrthe, Philippe, Rianneke, Tan Lai, Tate, Shunxin, Wenij, Xiaodi, Xiaofei, Zhewen and Yuri – the many (virtual) coffee's after the Friday meetings were always one of the highlights of the week, and are greatly missed!

Pandi thank you for the good conversations during the MRI measurements and statistics course. It's been a pleasure to learn more about your culture and experiences in the Netherlands.
Rachel, thuiswerken met mijn (bijna) buurvrouw was altijd een groot feest, zeker (tja belofte maakt schuld) onder het genot van Britney Spears, de Spice Girls en kerstliedjes.

Aan de senior onderzoekers op de afdeling, **Casper, Marleen, Kristiaan** en **Boy** het onderwerp van mijn onderzoek lag soms ver verwijderd van jullie vakgebied. Desalniettemin waren jullie altijd bereid om daarover met mij van gedachte te wisselen, en ik heb bovendien veel van jullie veel mogen leren over experimenteel onderzoek en statistiek. Ik wil ik jullie bedanken voor het warme welkom op de afdeling en dat ik de afgelopen jaren van al jullie kennis en ervaring heb mogen leren.
Jean, **Katrien** en **Margee**, vanwege het thuiswerken hebben we elkaar niet zo vaak getroffen op het lab, maar ook jullie wil ik bedanken voor het warme welkom op de afdeling en het geduld waarmee jullie mijn vragen over de (voor mij magische) werkzaamheden op het lab beantwoorden. **Judith** en **Marjo**, de mooie SHBG resultaten in dit boekje zijn te danken aan jullie zorgvuldige metingen waarvoor mijn dank groot is! **Petra** en **Vicky**, bedankt voor de gezellige gesprekken op de gang en de bereidwilligheid waarmee jullie altijd klaar staan om te helpen.
Hanny en **Marian**, ook jullie wil ik graag bedanken voor jullie vriendelijkheid op weg naar het werkoverleg en de hulp in het naast elkaar leggen van de drukke agenda's van mijn promotoren.

Mijn nieuwe collega's in het Elkerliek, **Arthur, Lisa, Marieke** en alle **internisten** bedankt voor het warme welkom en ondersteuning terwijl ik mijn weg vindt in deze klinische wereld. **Prof. dr. Koopmans** en **drs. Kox** wil ik graag bedanken voor de kans om met de opleiding Interne Geneeskunde te mogen starten.

Lieve **Eef**, ik ben erg blij dat je ooit verdwaald bent geraakt hier in het "verre zuiden", want daardoor hebben we de afgelopen jaren heel veel avonturen kunnen delen. Van de studiedagen in de bieb, de etentjes in de avond tot de vakanties in Italië en de werk-feestjes samen (hoe we van 'oh trouwens nog één ding, en dan gaan we echt werken' tot een afgerond proefschrift zijn gekomen begrijp ik nog steeds niet). Dank je wel voor jouw vriendschap, voor het luisterend oor wat je altijd biedt, als ik even mijn hart wilde luchten of als ik een monoloog hield over SHBG; voor het raadzame advies en je wijze woorden; maar bovenal voor alle gezelligheid. Tot slot, we hebben veel geleerd de afgelopen jaren, maar daarvan is het allerbelangrijkste (en ik citeer jouw wijze woorden): "thee is altijd een goed idee" – dus op naar nog vele theetjes!

To my friends on the other side of the North sea: Sresha, what would we have done without voice messages! It's telling that we have only spent four years in the same country and 14 years on completely different continents and yet when we do see one another it's as if no time has passed at all. Thank you for being only a phone call away – which will then of course result in several hours of laughter. Rosi, from the chocolate-bar in Boston to here, thank you for your catch-up skype-baking calls. Michaela your friendship and calm presence whenever I'm in London means a lot to me.

Aan alle vrienden van iets dichter bij: Carolien, van 4-jarige vriendinnetjes in tutu's tot samen studeren in Maastricht (gelukkig wel zonder de tutu's) – het is altijd één groot, gezellig avontuur. Op naar nog vele auto-karaoke sessies, goede gesprekken, boerenkoolstamppotten en kampeerweekenden. Jorinde, Anna en Isabel, onze wekelijkse etentjes zijn zowel culinair als qua gezelligheid een hoogtepunt in mijn week. Bedankt voor de vele bordspelletjes, goede gesprekken tot laat op de avond en de vele heuvels die we hebben 'getrotseerd'. Stephanie en Michelle toch nog eens wadlopen proberen? Het moet toch ooit lukken! Tessa, dankzij onze stedentripjes heb ik allerlei nieuwe plekken in Nederland leren kennen (met Venlo als hoogtepunt natuurlijk). Gelukkig zijn er nog vele plekken te ontdekken, dus nog een hoop gezellige dagen te gaan! Janine, Astrid, Wendy, Pauline, Vivienne, en Lina, de gezelligheid van A-KO avonturen in de bieb, de stad en het ziekenhuis zetten we voort ook na die tijd, en ik hoop dat we dat nog lang volhouden!

Aan mijn behandelteam in Eindhoven, dankzij jullie ondersteuning op de achtergrond ben ik toch net iets minder intensief bezig mijn de welbekende schommelingen en heb ik net iets meer tijd over voor alle leuke dingen in het leven. De controles bij jullie zijn na al die ervaringsjaren vooral een moment om even te delen hoe het ervoor staat met mijn proefschrift, de opleiding of privé… en zijn de uploads gelukkig slechts bijzaak. Ik ben soms (iets te) eigenwijs, maar wil jullie bedanken voor de ruimte die jullie mij daarin geven, en de fijne ondersteuning al die jaren.

Lieve opa's en oma's – lieve opa en oma Bosmans, bedankt voor alle kaarsjes die jullie altijd voor ons hebben laten branden als we een belangrijke dag hadden. Oma ik ben heel erg dankbaar en blij dat ik samen met jouw de afronding van dit proefschrift mag vieren. Lieve opa en oma Simons, ik weet dat u trots zou zijn op dit proefschrift, maar zeker ook dat uw kleindochter in Maastricht is beland. Het Limburgse dialect is me nog steeds niet eigen, maar wellicht dat het diep verstopt zit in de genen en nog eens naar boven komt.

Lieve Puck, Jort en Melle – ik ben stiekem toch altijd heel trots als ik mag zeggen dat ik ben opgegroeid in een grote bende van gezelligheid. Met z'n 4^{e} is er altijd iemand die je met twee benen op de grond houdt, iemand met wie je even kan sparren, maar

bovenal is er altijd veel gezelligheid en plezier. Onze famous four weekenden, waarin we menig (studenten)huis betraden met een hele berg aan luchtbedden, slaapzakken en rumoerigheid is een traditie de we hopelijk nog lang voortzetten. Puck, in heel veel dingen zijn we hetzelfde, in sommige totaal niet. Juist daardoor kan jij mij goed een spiegel voorhouden. Jort, ik ken weinig mensen die zo'n grote intrinsieke drang naar leren, onderzoeken en uitpluizen hebben als jij. Je kan het zo gek niet bedenken of jij hebt er over nagedacht. Ik zou dan ook geen beter persoon weten om naast me te staan als paranimf en ben heel blij dat je dat met zoveel enthousiasme en zorgvuldigheid wil doen. Melle, je bent dan misschien wel de benjamin van ons vier, maar van jou leer ik om te genieten van het leven, en dat is misschien wel de belangrijkste vaardigheid in het leven.

De appel valt niet ver van de boom. Lieve papa en mama, ik wil jullie bedanken voor de warme thuisbasis die jullie mij, en de rest van het gezin altijd bieden. Thuis is voor mij nooit een dorp of huis geweest, maar de plek waar ik door jullie met warme armen ontvangen wordt, waar dan ook ter wereld. Mama bij jou kunnen we altijd terecht voor een goed gesprek, of gewoon een fijne knuffel. Papa jij steunt alle keuzes die ik maak en laat door jouw rationele zorgzaamheid altijd je liefde voor het gezin blijken. Bedankt voor jullie onvoorwaardelijke steun en liefde.

Curriculum vitae

Pomme Simons was born on July 11th 1994 in Best, the Netherlands, and raised in a close-knit family. With her family, she moved to England where she followed part of her secondary education at ACS Egham International School in. Her VWO studies were completed at 2College Durendael in Oisterwijk, the Netherlands. In 2012, Pomme moved to Utrecht to start her bachelor's degree at University College Utrecht. She focussed her bachelor's studies on biomedical sciences, (health) psychology and linguistics. As part of her bachelor's degree she spent half a year studying Law, Medicine and Public Policy at Boston College, USA, where she was able to compare the ethics of Dutch versus American health care. In 2015 she was accepted to the master's degree Physician-Clinical Researcher in Maastricht. After successfully obtaining her medical degree in 2019, Pomme started her PhD research at the department of Internal Medicine, Maastricht University/Maastricht University Medical Centre under the supervision of prof. dr. Martijn Brouwers, prof. dr. Coen Stehouwer and dr. Olivier Valkenburg. During this time she focussed on unravelling the relationships between polycystic ovary syndrome and cardiometabolic disease and the role of de novo lipogenesis herein. As of January 1st 2022 Pomme is working as a resident Internal Medicine at Elkerliek hospital, Helmond.

Scientific output

List of publications

Simons PIHG, Valkenburg O, Telgenkamp I, van der Waaij KM, de Groot DM, Veeraiah P, Bons JAP, Derks TGJ, Schalkwijk CG, Schrauwen-Hinderling VB, Stehouwer CDA, Brouwers MCGJ. *Serum sex hormone-binding globulin levels are reduced and inversely associated with intrahepatic lipid content and saturated fatty acid fraction in adult patients with glycogen storage disease type 1a.* Journal of Endocrinological Investigation. 2022; online ahead of print.

Simons PIHG, Cornelissen M, Valkenburg O, Onland-Moret CN, van der Schouw YT, Stehouwer CAD, Burgess S, Brouwers MCGJ. *Causal association between polycystic ovary syndrome and coronary artery disease: a Mendelian randomization study.* Clinical Endocrinology. 2022; 96 (4): 599-604.

Simons PIHG, Valkenburg O, Stehouwer CDA, Brouwers MCGJ. *Sex hormone–binding globulin: biomarker and hepatokine?* Trends in Endocrinology and Metabolism. 2021; 32 (8): 544-553.

Simons PIHG, Valkenburg O, Bons JAP, Stehouwer CDA, Brouwers MCGJ. *The relationships of sex hormone-binding globulin, total testosterone, androstenedione and free testosterone with metabolic and reproductive features of polycystic ovary syndrome.* Endocrinology, Diabetes & Metabolism. 2021;n/a(n/a):e00267.

Simons PIHG, Valkenburg O, Telgenkamp I, van der Waaij KM, de Groot DM, Veeraiah P, Bons JAP, Taskinen M-R, Borén J, Schrauwen P, Rutten JHW, Cassiman D, Schalkwijk CG, Stehouwer CDA, Schrauwen-Hinderling VB, Hodson L, Brouwers MCGJ. *Relationship between de novo lipogenesis and serum sex hormone binding globulin in humans.* Clinical Endocrinology. 2021;95(1):101-106.

Simons N, Veeraiah P, Simons PIHG, Schaper NC, Kooi ME, Schrauwen-Hinderling VB, Feskens EJM, van der Ploeg EMC, Van den Eynde MDG, Schalkwijk CG, Stehouwer CDA, Brouwers MCGJ. *Effects of fructose restriction on liver steatosis (FRUITLESS); a double-blind randomized controlled trial.* The American Journal of Clinical Nutrition. 2020;113(2):391-400.

Simons PIHG*, Simons N*, Stehouwer CDA, Schalkwijk CG, Schaper NC, Brouwers MCGJ*. *Association of common gene variants in glucokinase regulatory protein with cardiorenal disease: A systematic review and meta-analysis.* PloS One. 2018;13(10):e0206174.
* Contributed equally

Submitted

Simons PIHG, Valkenburg O, van de Waarenburg MPH, van Greevenbroek MMJ, Kooi ME, Jansen JFA, Schalkwijk CG, Stehouwer CDA, Brouwers MCGJ. Serum seks hormone-binding globulin mediates the association between intrahepatic lipid content and type 2 diabetes: The Maastricht Study.

Simons PIHG, Valkenburg O, Stehouwer CDA, Brouwers MCGJ. Association between de novo lipogenesis susceptibility genes and coronary artery disease.

Poster presentations

Dutch Endocrine Meeting 2020 (Noordwijkerhout, the Netherlands): *Relationship between de novo lipogenesis and serum sex hormone-binding globulin levels in humans.*

European Association for the Study of Diabetes 2020 (online): *Relationship between de novo lipogenesis and serum sex hormone-binding globulin levels in humans.*

Oral presentations

Annual Dutch diabetes Research Meeting 2020 (online): *Relationship between de novo lipogenesis and serum sex hormone-binding globulin levels in humans.*

Dutch Epidemiological Conference 2021 (online): *Causal relationship between polycystic ovary syndrome and coronary artery disease: a mendelian randomization study*

EASD NAFLD study group, 2021 (online): A*ssociation between de novo lipogenesis susceptibility genes and coronary artery disease*

Annual Dutch Diabetes Research Meeting 2022 (online): *Association between de novo lipogenesis susceptibility genes and coronary artery disease*