Medical University of Bialystok

Faculty of Medicine with the Division of Dentistry
and Division of Medical Education in English

Doctoral dissertation in medical science

THE USE OF POLYGENIC RISK SCORES FOR TYPE 2 DIABETES IN PREDICTION OF METABOLIC CHANGES

Luis Felipe Padilla Martínez

Supervised by **Prof. dr hab. n. med. Adam Jacek Kretowski**
Clinical Research Center
Medical University of Bialystok

Bialystok, 2022

This research was conducted within the project which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754432 and the Polish Ministry of Science and Higher Education, from financial resources for science in 2018-2023 granted for the implementation of an international co-financed project.

This dissertation prepared with thanks to supervisors

Prof. dr hab. n. med. Adam Jacek Kretowski,

Dr. Lukasz Szczerbinski

& dr hab. Mirosław Kwaśniewski

as well as to my Parents, my brother, family and friends

Index

I. Abbreviations

ADA – American Diabetes Association

AUC – Area Under the Curve

BMI – Body mass index

CARDIA – Coronary Artery Risk Development in Young Adults

Chol – Total cholesterol

CHR – Chromosome

CRF – Clinical Risk Factors

DIAGRAM – Diabetes Genetics Replication And Meta-analysis

FDR – False discovery rate

FFM – Fat-free mass

FG – Fasting glucose

FI – Fasting insulin.

FM – Fat mass

GoDARTS – Genetics of Diabetes Audit and Research

GRS – Genetic Risk Score

GWAS – Genome-Wide Association Studies

Hba1c – Glycated hemoglobin

HDL – High-density lipoprotein

HLA – Human Leukocyte Antigens

ICD-10-CM – International Classification of Diseases, 10th Revision, Clinical Modification

IPAQ – International Physical Activity Questionnaire

IR – Interquartile range

LDL– Low-density lipoproteins

MAF – Minor allele frequency

MM – Muscle mass

MODY – Maturity onset diabetes of the young

NICE – National Institute for Health and Care Excellence

OGTT – Oral glucose tolerance test

PolRed – Polish Registry of Diabetes

PRS – Polygenic risk scores

SAT – Subcutaneous adipose tissue

SD – Standard Deviation

SF – Subcutaneous fat

SNP – Single Nucleotide Polymorphism

T1D – Type 1 Diabetes

T1DCG – Type 1 Diabetes Genetics Consortium

T2D – Type 2 diabetes

TG – Triglycerides

VAT – Visceral adipose tissue

VF – Visceral fat

WHO – World Health Organization

WTCCC – Welcome Trust Case Control Consortium

II. Tables and Figures list

Figures

Tables

I. Abstract

Prediabetes is an intermediate state of dysglycemia during which glycemic parameters are above normal levels but below the type 2 diabetes (T2D) threshold. It is well documented that prediabetes is a risk factor for progression to diabetes and cardiovascular disease. Recent studies have led to considerable advances in the identification of genetic variants associated with type 1 diabetes (T1D) and T2D. An approach for converting genetic data to a predictive measure of disease susceptibility is to add the risk effects of loci into a polygenic risk score (PRS).

The main objective of this research was to create a type 2 diabetes predictive polygenic risk score (T2D PRS) and obesity polygenic risk score (Obesity PRS) and find associations between these PRSs versus changes over time (Δ) in metabolic parameters related to T2D in Polish population.

For the present study, 446 prediabetic subjects (54.9% of females, median age at baseline: 42.5 yrs., median BMI at baseline: 26.9) have been selected from the Polish Registry of Diabetes study maintained by the Department of Endocrinology, Diabetes, and Internal Medicine, Medical University of Bialystok. All subjects who were included underwent follow-up exams five years after the initial exam.

In order to build a T2D PRS that can be accurate, the development of a systematic review of the most recent PRSs for different forms of diabetes with their advantages and disadvantages was done. Three PRS that discriminate between T1D patients and healthy people were identified, one that discriminate between T1D and T2D, two that discriminate between T1D and monogenic diabetes, and eight PRSs that discriminate between T2D patients and healthy people. After gathering and comparing all the information, genetic polymorphisms determined in studied patients were selected to build a T2D PRS (68 SNPs) and an obesity PRS (21 SNPs). Subsequently, 17 metabolic parameters were measured, and compared at baseline and after five years using statistical analysis. Finally, the associations between the two PRSs and the change in the metabolic traits were assessed. After a multiple linear regression with adjustment for age, sex, and BMI at a nominal significance of ($P < 0.05$) and adjustment for multiple testing, the T2D PRS was found to have a positive association with the change of fat mass (Δ FM) ($p = 0.025$). Meanwhile, the obesity PRS was also positively associated with Δ FM ($p = 0.023$) and Δ 2-hour glucose ($p = 0.034$). The comparison of genotype frequencies showed that

the AA genotype of *MTCH2* (rs10838738) is significantly associated with Δ glucose and Δ 2-hour insulin. Our findings suggest that prediabetic individuals with a higher risk for T2D experience increased Δ FM, and those with a higher risk of obesity experience increased Δ FM and Δ two-hour postprandial glucose. The associations found in this research could be a helpful tool for identifying individuals with an increased risk of worsening of the metabolic state.

II. Abstract in Polish

Stan przedcukrzycowy to pośredni stan dysglikemii, w którym parametry glikemii są powyżej normy, ale poniżej progu dla cukrzycy typu 2 (T2D). Jest dobrze udokumentowane, że stan przedcukrzycowy jest czynnikiem ryzyka progresji do cukrzycy i chorób układu krążenia. Ostatnie badania doprowadziły do znacznych postępów w identyfikacji wariantów genetycznych związanych z cukrzycą typu 1 (T1D) i T2D. W celu praktycznego zastosowania danych genetycznych do przewidywania ryzyka rozwoju choroby wykorzystuje się połączony efekt wielu genów tworząc poligenowe wskaźniki ryzyka (PRS, Polygenic Risk Score).

Głównym celem obecnego badania było stworzenie predykcyjnego wielogenowego wskaźnika cukrzycy typu 2 (T2D PRS) i wielogenowego wskaźnika ryzyka otyłości (Obesity PRS) oraz znalezienie w polskiej populacji związku między tymi PRS a zmianami w czasie (Δ) parametrów metabolicznych związanych z T2D.

Do niniejszego badania z Polskiego Rejestru Cukrzycy prowadzonego przez Klinikę Endokrynologii, Diabetologii i Chorób Wewnętrznych Uniwersytetu Medycznego w Białymstoku wybrano 446 pacjentów w stanie przedcukrzycowym (54,9% kobiet, mediana wieku na początku badania: 42,5 roku, mediana BMI na początku badania: 26,9). Wszyscy badani, którzy zostali uwzględnieni, zostali ponownie przebadani po okresie 5 lat.

W celu zaprojektowania PRS T2D/PRS otyłości, wykonano systematyczny przegląd najnowszych publikacji dotyczących PRS dla różnych postaci cukrzycy wraz z ich zaletami i wadami. Zidentyfikowano trzy PRS, które odróżniają pacjentów z T1D od osób zdrowych, jeden, który odróżnia T1D od T2D, 2, który odróżnia T1D od cukrzycy monogenowej i 8 PRS, który odróżnia pacjentów z T2D od osób zdrowych. Po zebraniu i porównaniu wszystkich informacji określono polimorfizmy genetyczne występujące u pacjentów w celu zbudowania PRS T2D (68 SNP) i PRS otyłości (21 SNP). Następnie zmierzono 17 parametrów metabolicznych i porównano je na początku i po pięciu latach przy użyciu analizy statystycznej. Na koniec oceniono związek między dwoma PRS i zmianą cech metabolicznych. Po wielokrotnej regresji liniowej z korektą ze względu na wiek, płeć i BMI przy nominalnej istotności ($P < 0,05$) i korektą na wielokrotne testy wykazano, że T2D PRS ma dodatni związek ze zmianą masy tłuszczowej (Δ FM) ($p = 0,025$). Zaobserwowano, że PRS otyłości koreluje ze zmianą masy tłuszczowej Δ FM (p

= 0,023) i zmianą wartości glikemii w czasie 120 (p = 0,034). Porównanie częstości występowania genotypów wykazało, że genotyp AA *MTCH2* (rs10838738) jest istotnie związany z Δ glukozy i Δ insuliny w czasie 120. Nasze wyniki sugerują, że wśród osób z prediabetes badany wielogenowy wskaźnik cukrzycy typu 2 (T2D PRS) koreluje z ryzykiem przyrostu tłuszczowej masy ciała, a wielogenowy wskaźnik ryzyka otyłości (Obesity PRS) jest dobrym predyktorem zwiększenia tłuszczowej masy ciała i wzrostu glikemii w 2 godzinie OGTT w trakcie 5-letniej obserwacji. Wyniki tych badań mogą sugerować, że analizowane wskaźniki ryzyka cukrzycy typu 2/otyłości mogą być użytecznym narzędziem do identyfikacji osób o zwiększonym ryzyku pogorszenia stanu metabolicznego.

III. Introduction

1. Diabetes

Diabetes mellitus is a complex and heterogeneous group of chronic metabolic diseases characterized by hyperglycemia, now recognized as one of the most critical public health challenges of the 21st century [1]. The World Health Organization [2] estimated that diabetes was the seventh leading cause of death in 2016, being the direct cause of 1.6 million deaths. In 2014, 8.5% adults of 18 years old and older developed diabetes. In Europe, 1 in 11 adults are living with diabetes (Figure 1), and the number of adults with diabetes is expected to reach 67 million by 2030[3].

Diabetes is a chronic disease that occurs when high blood sugar levels result from the body's inability to produce or make enough of the hormone insulin. The pancreas produces a hormone called insulin, which is essential for the body to function. Insulin is needed for the metabolism of carbohydrates, proteins, and fat, and it allows glucose from the bloodstream to enter a cell's interior, where it is converted to energy. When insufficient insulin or cells are not responding to insulin, the result is hyperglycemia (high blood glucose levels), indicating diabetes [3]. Diabetes can result in serious health complications if not well managed, including cardiovascular diseases, nerve damage, kidney diseases, limb amputation, and eye problems affecting the retina (including blindness). However, if diabetic management is well handled, these complications can be delayed or even prevented altogether. An insulin deficit left unchecked for a long time can also cause damage to many of the patient's organs, making the patient ill and even dying [3].

Diabetes can be controlled and its consequences prevented or delayed through diet, physical activity, medication, and regular assessment and treatment of complications [2]. There are three most common types of diabetes. Type 1 diabetes (T1D) occurs predominantly in people < 30 years old and is generally thought to be precipitated by immune-associated destruction of insulin-producing pancreatic beta cells, leading to insulin deficiency and requiring exogenous insulin supplement [4]. Type 2 diabetes (T2D), the most common type of diabetes, is a progressive metabolic disease characterized by insulin resistance [5] and eventual functional failure of pancreatic beta cells [6,7]. Maturity-onset diabetes of the young (MODY) is a monogenic form of diabetes showing an autosomal dominant mode of

inheritance. It accounts for 1-5% of all diabetic forms of young and is specified by anomalous pancreatic beta-cell activity [8–10].

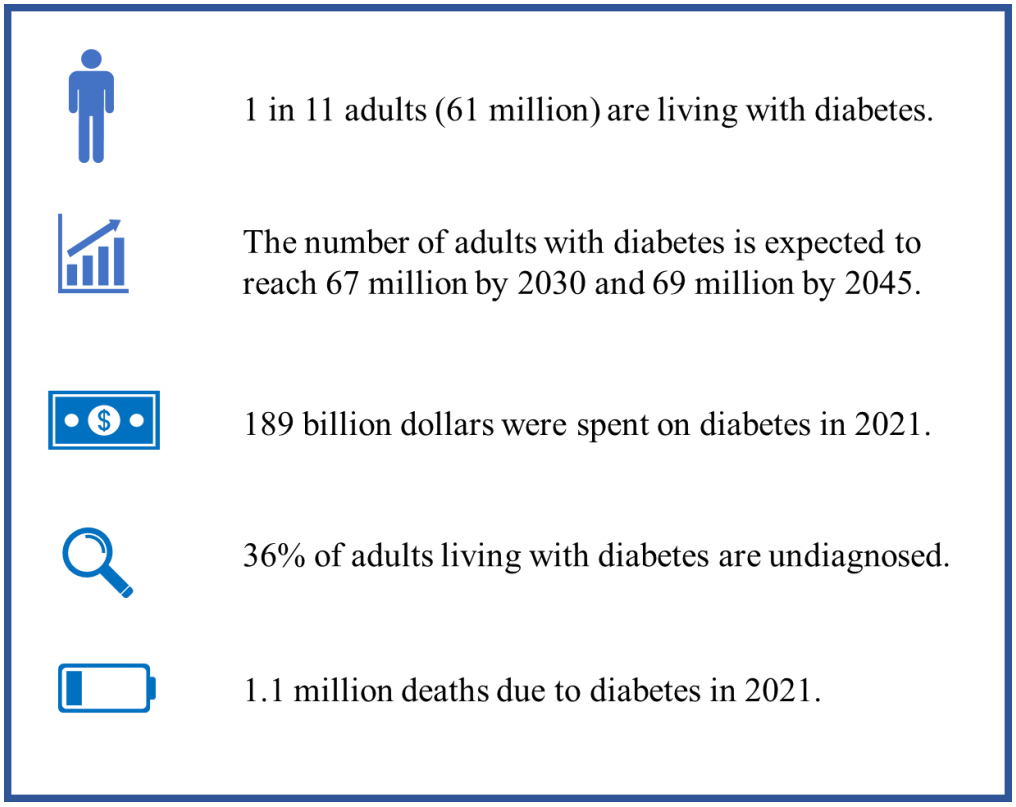

Figure 1. Statistics of Diabetes in Europe 2021[3]

2. Type 1 Diabetes

T1D is a chronic disease in which the immune system attacks the insulin-producing pancreatic beta-cells. This results in a lack of insulin and elevated blood glucose levels[11]. When the body's immune system starts to produce autoantibodies against beta-cells, the person eventually develops T1D and produces little to no insulin[3]. More than 90% of patients newly diagnosed with T1D have measurable autoantibodies against specific beta-cell proteins, including insulin, glutamate decarboxylase, islet antigen 2, zinc transporter 8, and tetraspanin-7 [12].

Different factors such as diet, genetic background, environment, beta-cell stress, and immune phenotype increase the development of autoimmunity and beta-cell loss in clinical T1D [13]. T1D has a substantial heritable component, estimated to be between 65 to 88% [14,15]. Genes in the HLA region confer 50% of the genetic risk of T1D. The genes in this complex

are categorized into two major classes: class I and class II. Class-I HLA presents antigen peptide found within the cell to CD8 positive (cytotoxic T cells). In contrast, Class-II HLA presents antigen peptide found outside the cell to CD4 positive (helper T cells)[16]. Over 60 common non-HLA T1D risk variants across the genome have been identified in linkage and genome-wide association studies (GWAS) [17,18]. Over the past two decades, there has been an explosion of knowledge about T1D, including the immune characteristics of the disease, as well as its incidence, genetics, and clinical burden. There have been many interventions to preserve beta cells and several methods to improve disease management. However, despite this increased knowledge, there are still many gaps in our understanding of T1D and our ability to manage the disease and its complications[3].

3. Type 2 Diabetes

T2D remains a significant clinical burden worldwide. T2D is costly, and affects individuals, health care systems, and economies [19]. T2D affects 6.28% of the world's population and is the most common form of diabetes, accounting for more than 90% of all diagnosed cases of diabetes worldwide [2,3]. The most crucial feature of T2D pathogenesis is insulin resistance, where tissues are not responding correctly to physiological insulin secretion. With the onset of insulin resistance, insulin is less effective and prompts an increase in pancreatic production. Over time, this can lead to the failure of the pancreatic beta cells and the development of overt T2D[3]. After the onset, T2D for many years can be asymptomatic; however, when symptoms are already present, they are usually less pronounced than in T1D. The beginning of T2D is impossible to pinpoint, and many people who have it go undiagnosed for an extended period (even up to half or a third of people with T2D). When the disease is asymptomatic for a long time, complications such as retinopathy, neuropathy, heart disease, and even stroke can occur as a first manifestation of the disease leading to the diagnosis of diabetes [3,20,21]. Many variables can increase the risk of T2D, such as age, obesity, having a family history of diabetes, or being of a certain ethnicity. The pathogenesis of the disease is a combination of environmental triggers and genetic predispositions [3,22].

More than 400 genetic loci have been discovered to be associated with diabetes risk by multiple studies [23–25]. While lifestyle and drug interventions can play a part in slowing down the progression of diabetes development, much research is still being done to determine those who will develop T2D at some point in their lives. There is some skepticism regarding

the practical use of these genetic variants in personal risk prediction for T2D due to the relatively weak effect size of single genetic variants and the fact that the environment is the main cause of the development of T2D [26,27]. However, there are ongoing efforts to explore the clinical utility of polygenic risk scores, combining the effects of multiple genetic variants.

Current treatments for T2D have been incapable of stopping the development of T2D and complications [28,29]. One of the reasons may be the heterogeneity of the disease and the fact that the one-fits-all approach for diabetes prevention and treatment does not work the same way for all patients [30,31]. The implications of wrong diagnosis, coding or classification affect optimal treatment regimen and cause inappropriate financial and psychological impact in such patients. Patients with the correct diagnosis of disease, with identification of the etiology of the disease, achieve significant improvements in their glycemic control [32,33].

4. Prediabetes

Prediabetes is an intermediate state of hyperglycemia with glycemic parameters above normal but below the diabetes threshold [34] and it affects 7.3% of the world's population [35,36]. Approximately 25% of people who have prediabetes will develop full T2D in 3-5 years, and up to 70% of people who have prediabetes will develop T2D during their life [37,38]. Many different factors can cause diabetes and prediabetes. Lifestyle, genetic, and environmental factors all can play a part. The primary cause is obesity; in fact even 80-85% of cases of diabetes and prediabetes can be mediated by excessive body mass [39,40]. While prediabetes can lead to T2D, itself has negative health consequences. Clear links between cardiovascular disease, metabolic syndrome, and prediabetes have emerged in recent years. Nevertheless, the pathophysiological defects seen in prediabetes can be managed by lifestyle modifications in most patients [41,42]. In addition to the complications associated with the condition, differentiating prediabetes from diabetes is supported by the International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM) [36,43]. Different organizations have different criteria for prediabetes, and care needs to be taken when describing prevalence and incidence statistics (Table 1).

Table 1. Prediabetes diagnostic criteria[44]

Diagnostic criteria	**WHO**[2]	**ADA**[45]	**NICE**[46]
HbA1c	Not recommended for diagnosis	39 – 47 mmol/mol	42 – 47 mmol/mol
2-hour glucose during OGTT	7.8 – 11 mmol/L	7.8 – 11 mmol/L	7.8 – 11 mmol/L
Fasting plasma glucose	6.1 – 6.9 mmol/L	5.6 – 6.9 mmol/L	6.1 – 6.9 mmol/L

5. Obesity

Obesity is a chronic disease, a global pandemic, and a significant risk factor for other conditions such as T2D, heart disease, and cancer [47,48]. Over 1 billion people will be obese worldwide by 2030, according to new data presented in the World Obesity Atlas 2022 [48]. Obesity, like all chronic diseases, has a different range of determinants such as genetics, biology, healthcare access, mental health, sociocultural factors, diet, economics, and environment [48]. Changes in appetite, satiety, metabolism, amount of body fat, and hormone balance are all caused by obesity. These changes do not always go away with weight loss and can last many years. Prevention is critical in countries where the obesity trend is just beginning. The global obesity epidemic is driving the rapid increase in the prevalence of T2D[49]. It is known that obesity, prediabetes, and insulin resistance are highly related [50–52], nevertheless in the last years it have been discovered the highly potential causal effect of obesity on prediabetes and insulin resistance and the key role of adipose tissue in insulin resistance[53].

While the environment has been a major factor in increasing obesity rates, genetic factors also play a key role in the development of the disease [54,55]. Hundreds of genes have been identified through GWAS to be connected to obesity [55–57], though they only influence around 5% of the chance of someone being obese [55,58]. Their low influence may be because there are still unknown interactions between genes, the environment, and other epigenetic factors [55,59]. Many genes connected to obesity are involved in energy-regulating processes, such as glucose metabolism and circadian rhythm.

6. Genomics

Genomics medicine is focused on understanding an individual's biology based on their genetic code [60]. Genome technology is transforming healthcare, enabling more genes to be sequenced in less time and at a lower cost. Today, physicians and scientists have an unprecedented ability to discover genes, unravel molecular signaling pathways abstractly, and find new targets for biomarkers and therapy [61]. Doing so, they can predict a patient likelihood of developing a disease in the future, which can help improve the health care sector by preventing unnecessary concerns and preempting therapies for people who are considered at higher risk [62]. Thanks to genomics and GWAS, scientist have been able to identify genes associated with a particular disease, searching for single nucleotide polymorphism (SNPs). GWAS aims to determine genotype-phenotype associations by testing for differences in allele frequencies of genetic variants between individuals with similar ancestry but different phenotypes. GWAS results have many applications, such as to gain insight into the underlying biology of phenotypes, estimate their heritability, and find potential relationships between genetic risk factors and health outcomes[63]. The data generated from GWAS is being used more and more to predict metabolic diseases [64,65]. If genomic data will be used across healthcare, specialists must understand the potential risk associated with interpreting genomic data to ensure its safety and benefit to patients.

7. Type 2 Diabetes Genomics

The advent of genotyping and sequencing technologies has contributed to the discovery of many genetic variants contributing to T2D pathogenetic complexity. Likewise, the generation of genome-wide variation data has become common for predicting metabolic diseases [64,65]. T2D has well-established risk loci and likely contains many genetic determinants with effects too small to be detected at genome-wide levels of statistical significance [66]. This demonstrates that all common variants across the genome explain a much higher proportion of heritability (50% or more) in many complex traits than can be seen using only a small subset of significant SNPs [67]. These advances provide opportunities to determine the utility of genomic regions in predicting treatment responses [68]. The number of studies combining phenotypic and genetic variables to predict diabetes risk has increased recently and show generally promising results [62,69].

Many loci on the genome have been shown to increase the risk of T2D. Many genetic components may not be seen at high significance levels on a larger scale but can be seen when looking at all common variants across the entire genome [66]. This suggests that all common variants in the genome explain a higher proportion of heritability in many complex traits than can be seen based on only a small subset of significant SNPs [67].

8. Polygenic Risk Scores

An approach for converting genetic data to a predictive measure of disease susceptibility is to add the risk effects of these loci into a single score called polygenic risk score (PRS) [70,71]. Based on the largest GWAS studies, the PRS combines multiple alleles an individual carries that are considered risk alleles for a disease. The number of risk alleles an individual carries is added together, and then weighted by the size of each allele effect (log of odds ratio for binary traits or beta coefficient for continuous traits). The result is one overall score indicating an individual's likelihood of developing a disease or possessing a particular trait [72].

Despite only explain a small fraction of trait variation, the correlation between PRSs and an individual's highest likelihood of a trait has made them a very popular tool in biomedical research. They may be used in clinical practice in patients with a higher likelihood of suffering from the disease, for example, in early stages of disease. They could be used to help with diagnosis, suggest treatment options, to determine shared etiology between diseases, and more [72,73]. PRS studies generally use cohorts that are fairly similar together (such as linked to ethnicity), which is one of the limitations of these studies [74]. To solve this problem, calibration, validation, and optimization of the PRS is needed for every study cohort to ensure that the results are not fitted [60].

9. Polygenic Risk Scores in Diabetes

Over the past few decades, there have been many extensive genetic studies that have looked at the risk of developing T2D or T1D across multiple sites in the genome [75,76]. There are over 400 different genetic signals on T2D risk identified [77], and over 50 loci influencing T1D risk [78]. There are many reasons why genetic testing for diabetes risk is not part of the standard care, but some of the main ones are the cost of genotyping, lack of education of

healthcare providers in precision medicine utilizing genetic testing, and still ongoing efforts to improve their predictive power [13]. There is an increased rate of T2D in the public today, and it is one of the biggest health concerns [79]. Although obesity is the strongest predictor of T2D, it is also known that heritability of T2D is between 26-69% depending on age of onset, thus motivating the search for genetic predictors for T2D [80–82]. This encourages the search for genetic markers that will predict T2D and create a numeric index of risk: a PRS based on many genotyped variants [83]. The PRS encourages decision support for diagnosis, and they are reliable when discriminating diabetes subtypes[84–86].

IV. Objectives

1. Systematic review of studies comparing the accuracy of polygenic risk scores developed for T1D and T2D.

2. Evaluation of the T2D PRS and the obesity PRS in predicting changes in clinical parameters related to prediabetes and metabolic complications over time.

3. Evaluation of the association of selected T2D and obesity SNPs genotypes with changes in clinical parameters related to prediabetes and metabolic complications.

V. Materials and Methods

1. Systematic review: Search Strategies and Study Selection

The databases for literature searches were Web of Science, Scopus and PubMed. The keywords for the databases search were: T1D, T2D, and monogenic polygenic risk score studies discovered between 2000 and September 2019. Search terms were "type 1", "diabetes", "genetic risk score" "polygenic risk score", "type 2", "mature-onset diabetes in young adults", and all possible combination of these terms. Publications excluded during the screening phase were (1) articles only available as abstracts, (2) risk assessments developed before 2000, and (3) non-English publications.

2. Data Collection Process

The details collected from the full text and additional information included the first author of the study, the year the study was published, the DOI when available, the ethnic background of the participants, the country in which the study took place, the data set used in the study, and if validation sets were used. The number of patients and controls, method of genotyping/sequencing, and the specific panels and numbers of genes used for genotyping were also collected. Additionally, the numbers of SNPs used to create a PRS were noted, as well as the clinical risk factors used and the AUC each had.

3. Synthesis of Review Results

The AUC was considered to compare and assess the accuracy of the PRS. The AUC was split into three categories based on the subtype of diabetes to differentiate between them. The first group included a T1D PRS comparison. The second group had a T2D PRS comparison. The third group included a T1D PRS comparison used to discriminate T1D vs. T2D and T1D vs. monogenic diabetes.

4. Observational study: Study Design and Participants

The data was collected within the Polish Registry of Diabetes (PolReD) study (formerly known as the 1000PLUS cohort), conducted at the Clinical Research Centre of the Medical University of Bialystok, Poland, and comprises patients with follow-up data, enrolled in the study between 2009 and 2012. In total, 446 subjects who were prediabetic but did not have a diagnosis of T2D at baseline were selected from the overall population for this study. The PolReD study design has been described in detail [87,88]. Before the study began, all participants signed an informed consent form. The ethics committee of the Medical University of Bialystok originally approved this study (RI-002/436/2019). Patients at risk of developing diabetes (prediabetes), defined as impaired fasting glucose, impaired glucose tolerance, or both [89], were excluded if they had any recent surgery, infection, cardiovascular disease, or severe illness. Those included underwent a follow-up exam five years after the initial exam.

5. Sample Collection and Body Composition Measurements

Participants had to fast overnight and not do much physical activity the day before their tests. Their blood was taken from whole blood samples in two visits: baseline (visit 1) and follow-up after five years (visit 2). The participant's weight, body mass index (BMI), and other anthropometric measurements were taken using standardized procedures. Biochemical measurements, including plasma glucose, serum triglycerides (TG), total cholesterol, high-density lipoprotein (HDL), and low-density lipoprotein (LDL) concentrations, were performed by the colorimetric method with Cobas c111 (Roche Diagnostics, Basel, Switzerland). Insulin concentrations were measured in the serum using an immunoradiometric assay kit (DIAsource ImmunoAssays SA, Belgium). Glycated hemoglobin (HbA1C) was measured by the high-performance liquid chromatography method (D-10 Hemoglobin Testing System, Bio-Rad Laboratories Inc., Hercules, CA, USA; Bio-Rad, Marnes-la-Coquette, France). The colorimetric method measured fasting glucose concentration and glucose concentration at two hours in the plasma. The fat-free mass (FFM), fat mass (FM), muscle mass (MM), visceral fat (VF), subcutaneous fat (SF), and the ratio of visceral adipose tissue to subcutaneous adipose tissue (VAT/SAT) were measured using a Maltron body fat analyzer (Maltron BioScan 920-2, Maltron International Ltd., United Kingdom). Physical activity was measured using the International Physical Activity Questionnaire (IPAQ). Visit 1 and 2 used the same method for all measurements.

6. Genotyping

DNA was extracted from the peripheral blood leukocytes using the classical salting out method. The SNP genotyping was done with TaqMan SNP technology from a ready-to-use human assay library (Applied Biosystems, MA, USA) using a high-throughput genotyping system, OpenArray (Life Technologies, CA, USA). A sample without DNA was used as a negative control to help detect any contamination in the system.

7. Polygenic Risk Score Analysis

Two PRSs were constructed, one for T2D and the other for obesity, the approach to building them was by summing the number of risk alleles carried by each individual, weighted by the effect size estimates from well-established genome-wide associations selected from the Type 2 Diabetes Knowledge Portal [90]. Due to the limited availability of SNPs on our genotyping platform, we could include only a subset of the known genome-wide significant loci for T2D and obesity, resulting in a T2D PRS of 68 SNPs and an obesity PRS of 21 SNPs. The analysis and calculations were done in R (version 4.1.0) [91].

8. Statistical analysis

The mean ± standard deviation (SD) or median (interquartile range) are reported for continuous normally, or non-normally distributed traits. Normality was assessed using the Shapiro-Wilk test. This analysis revealed that the studied parameters did not follow a normal distribution. Consequently, nonparametric tests were used for the statistical analysis between groups. The Wilcoxon signed-rank test was used to compare variables at baseline and follow-up. The change (Δ) in time (T2 minus T1) of each metabolic parameter was obtained.

To check if the genotypes' frequencies had a statistically significant effect, they were compared to different metabolic parameters in a series of tests. Statistically significant differences between groups, determined by genotypes, were estimated using the Kruskal-Wallis test. A post-hoc analysis was performed by applying the Wilcoxon rank-sum test for all pairwise comparisons to discover which genotypes caused the particular test to be significant. Multiple linear regression with adjustment for age and sex was used to test the association between the PRSs and baseline metabolic parameters.

After that, another multiple linear regression with adjustments for age, sex, and BMI was used to test the association between the PRSs and the changes in metabolic parameters between baseline and follow-up. β coefficients were presented as an incremental increase or decrease in the trait per the SD of the tested PRS. For all the tests described in this section, the *p*-values were adjusted to <0.05 using the false discovery rate correction for multiple comparisons. All calculations were prepared in R (version 4.1.0) [91].

VI. Results

1. Systematic Review: Selection of Studies

A total of 14 studies were selected for the systematic review after screening and evaluating 62 articles retrieved from PubMed, Scopus, and Web of Science. The studies have different genes, and genotyping strategies in their data sets, and panels. (Figure 2).

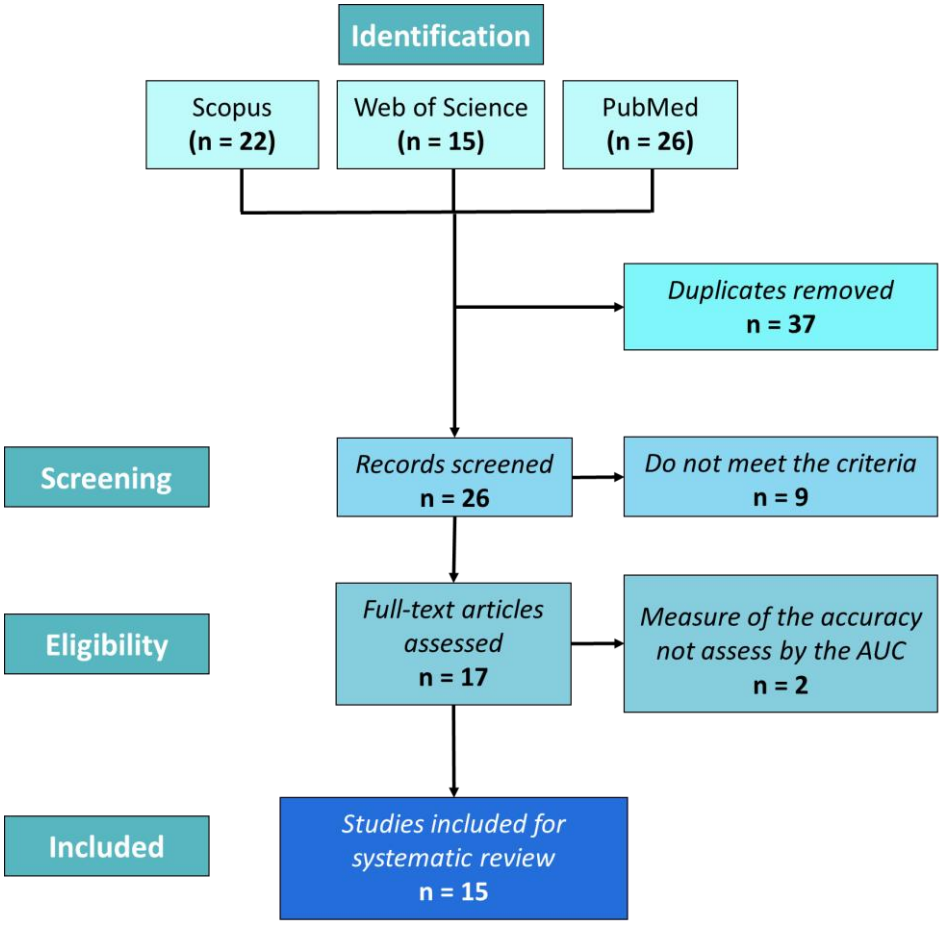

Figure 2. Study selection flow diagram[92]

Six studies reviewed all the possible PRS for T1D [93–98], and eight other studies reviewed PRS for T2D [83,99–105] (Table 2). Most of the studies were conducted on Caucasian populations, but some conducted studies on Hispanic, African-American, Asian-American and Iranian populations. Despite the Iranian cohort, all of the studies included a large number of patients and control subjects.

Table 2. The studies selected for systematic review[92]

Study	Year	Country / Ethnicity	Patients	Controls
Studies describing Type 1 Diabetes Polygenic Risk-Score				
Winkler, C. **[95]**	2014	Caucasian	4,574	1,207
Oram, R. **[93]**	2015	Caucasian	n = 1,938	
Patel, K. **[94]**	2016	Caucasian	1,963	805
Perry, D. **[96]**	2018	Caucasian, Hispanic, African-America, and Asian-American	627	423
Sharp, S **[97]**	2019	Caucasian	6,670	9,416
Yaghootkar, H. **[98]**	2019	Iranian	121	6
Studies describing Type 2 Diabetes Polygenic Risk-Score				
Weedon, M. **[100]**	2006	British	2,409	3,669
Lango, A. **[101]**	2008	Scotland	2,309	2,598
Lyssenko, V. **[102]**	2008	Finland	2,201	16,630
Meigs, J. **[103]**	2008	European ancestry in USA	n = 2,776	
Chatterjee, N. **[104]**	2013	Caucasian	8,130	38,987
Vassy, J. **[105]**	2014	European ancestry in USA	5,941	5,942.
Läll, K. **[83]**	2016	Estonia	1,181	9,092
Khera, A.V. **[99]**	2018	British	26,676	120,280

There are many datasets that were used in the studies: T1DGC [106], WTCCC [107], UFDI, Iranian Hospitals [35], UK hospital [45], GoDARTS [108], MPP [109], BPS [110], Framingham Offspring Study [111], Voight [112], CARDIA [113], The Estonian Biobank [114] and the UK Biobank [115] (Table 3).

The panel of genes used to build the PRS are also different. For T1D PRS, the studies used either the panel of genes from T1DGC (*n* = 4) [106], 1000 genomes project (*n* = 4) [116] or the Immunobase.org on October 2017 (*n* = 1). For T2D PRS, the studies used either specific genes from previous studies (*n* = 5), different versions of the DIAGRAM Consortium panel of genes (*n* = 2)[117], or the 1000 genomes project (n =1) [116]. Finally, the platforms used for genotyping also differ. Most of the studies used modified TaqMan assays (*n* = 5) [95,96,100,101,105], different versions of Affymetrix microarrays (*n* =5) [93,94,97,99,105] and Illumina technology (n = 3)[74,85,118]. One study used KASPar genotyping[102], another the iPLEX technology[103], and another one used next-generation sequencing unspecified[98].

Table 3. Data set source, panel of genes used and genotyping strategies[92]

Study	Year	Data set	Panel of genes	Platform
Studies describing Type 1 Diabetes Polygenic Risk-Score				
Winkler, C. **[95]**	2014	T1DGC	T1DGC	TaqMan 5´nuclease assay
Oram, R. **[93]**	2015	WTCCC	1000 genomes and T1DGC	Affymetrix 500K SNP chip
Patel, K. **[94]**	2016	WTCCC	1000 genomes and T1DGC	Affymetrix 500K SNP chip
Perry, D. **[96]**	2018	University of Florida diabetes institute (UFDI)	Immunobase.org October 2017	Taqman SNP genotyping array
Sharp, S. **[97]**	2019	T1DGC	1000 genomes	Affymetrix Axiom Array
Yaghootkar, H. **[98]**	2019	Imam Reza Hospital and Children's Medical Centre in Iran	1000 genomes and T1DGC	Targeted next-generation sequencing (Unspecified)
Studies describing Type 2 Diabetes Polygenic Risk-Score				
Weedon, M. **[100]**	2006	UK	KCNK11, PPARG, TCF7L2.	Modified TaqMan
Lango, A. **[101]**	2008	GoDARTS	Frayling [119] and Zeggini [120]	Modified TaqMan
Lyssenko, V. **[102]**	2008	Malmö Preventive Project (MPP) & Botnia Prospective Study (BPS).	Gloyn [121], Grant [122], Saxena [123], Frayling [119], Scott [124], Sladek [125]	Allele-specific (KASPar)
Meigs, J. **[103]**	2008	The Framingham Offspring Study	Saxena [123], Zeggini [120]	iPLEX technology
Chatterjee, N. **[104]**	2013	Voight [112]	Voight [112]	Illumina Omni 2.5M Platform
Vassy, J. **[105]**	2014	The Framingham Offspring Study & CARDIA	DIAGRAMv3	Taqman, Illumina technology, Affymetrix 6.0, llumina 370
Läll, K. **[83]**	2016	The Estonian Biobank	DIAGRAM Consortium	Illumina Human OmniExpress, Illumina Cardio-MetaboChip
Khera, A.V. **[99]**	2018	The UK Biobank	1000 genome phase 3 version 5 (Linkage disequilibrium panel)	Affymetrix UK BiLEVE Axiom array

2. Polygenic Risk Score for T1D prediction.

Winkler developed a multivariable logistic regression model to estimate PRS, including 40 non-HLA genes SNPs, significantly improving the risk score with an AUC of 0.87 compared to the control [95]. Oram and colleagues [93] used a log-additive model to develop a PRS model to discriminate between patients versus controls for T1D. They applied a 30 SNP T1D-PRS to a sample of cases of T1D versus controls. They found the T1D-PRS was highly discriminant, with an AUC of 0.88 (Table 4).

The majority of genetics studies on T1D are limited to Caucasian cohorts. However, Perry investigated the hypothesis that ethnicity would be necessary for evaluating genetic risk markers previously identified in Caucasian cohorts [96]. They apply the PRS used by Oram [94] to Hispanic Caucasian, African American and Asian American populations. The Hispanic Caucasian PRS was highly discriminant with an AUC of 0.90. The PRS for Asian Americans was also highly discriminant with an AUC of 0.92, the analysis indicated that this PRS could discriminate T1D subjects from controls in a small cohort for subjects of Asian Americans, but larger studies are required to validate and extend these findings. The African Americans obtained a less discriminant PRS with an AUC of 0.75; notable risk differences were observed for 3 SNPs: SH2B3, CTRB1/2, and GAB3 in this population [96]. The most recent update for T1D-PRS includes 67 SNPs and accounts for interactions between 18 HLA DR-DQ combinations. This risk score identifies individuals with T1D with an AUC of 0.92 [97] (Table 4).

Table 4. Comparison of the accuracy of T1D PRS assessed by the AUC[92].

Year	Autor	PRS	SNPs	AUC PRS	Ethnicity
2014	Winkler, C. [95]	T1D	41	0.87	Caucasian
2015	Oram, R. [93]	T1D	30	0.88	Caucasian
2018	Perry, D. [96]	T1D	32	0.86	Caucasian
2018	Perry, D. [96]	T1D	32	0.90	Caucasian Hispanic
2018	Perry, D. [96]	T1D	32	0.75	African American
2018	Perry, D. [96]	T1D	32	0.92	Asian American
2019	Sharp, S. [97]	T1D	67	0.93	Caucasian

3. Polygenic risk scores for T2D prediction

Before the first GWAS for T2D, a study was published describing three genetic variants associated with T2D and comparing the combined risk of these variants and the risk of genetic testing using AUC ratings. The AUC was 0.58, a value above 0.50, indicating no discrimination but lower than that observed in clinical trials [100]. Two years later, Lango and colleagues examined the PGR with 16 SNPs, with an AUC of 0.789 for predicting diabetes incidence, ; adding PRS to the clinical factors had little effect on performance and pushed the AUC down to 0.80 [101] (Table 5).

In a different study, a16 SNPs PGR (scores adjusted for age, sex, family history, BMI, blood pressure, triglycerides, and fasting glucose) predicted diabetes incidence with an AUC of 0.740, after adding PRS to clinical risk factors (CRF), an AUC of 0.750 had little effect on the ability to predict T2D [102]. In the same year, Meigs estimated a PRS with 18 SNPs with an AUC of 0.534 for new-onset diabetes. With an expanded clinical model including age, sex, family history, BMI, glucose levels, cholesterol levels, and triglyceride levels, the AUC was 0.90. Adding genetic data to these two PRSs increased the AUC to 0.58 and 0.910, respectively [103]. Another study was performed using a PRS with 22 SNPs, with an AUC of 0.570 adjusted for age, sex, and family history, adding the PRS, the AUC increased to 0.740 [104]. The PRS with 62 SNPs showed an improved AUC for T2D prediction, with an AUC of 0.72, and a score of 0.91 after adding other important clinical factors [105]. Khera [99] used a different method, including 7 million variants, generating a PRS with an AUC of 0.73 (Table 5).

Table 5. Comparison of the accuracy of T2D PRS assessed by the AUC[92].

Year	Autor	PRS	SNPs	AUC CRF	AUC PRS + CRF	Diff	Clinical risk factors	Ethnicity
2006	Weedon, M. [100]	T2D	3	-	0.580	-	-	Caucasian
2008	Lango, A. [101]	T2D	18	0.780	0.800	0.020	Age, BMI, sex	Caucasian
2008	Lyssenko, V. [102]	T2D	16	0.740	0.750	0.010	Age, sex, family history, BMI, blood pressure, triglycerides, fasting plasma glucose	Caucasian
2008	Meigs, J. [103]	T2D	18	0.534	0.581	0.047	Age, sex	Caucasian
2008	Meigs J. [103]	T2D	18	0.595	0.615	0.020	Sex, age, family history	Caucasian
2008	Meigs, J. [103]	T2D	18	0.900	0.910	0.010	Age, sex, family history, BMI, glucose, cholesterol, triglycerides	Caucasian
2013	Chatterjee, N. [104]	T2D	22	0.570	0.740	0.170	Age, sex, family history	Caucasian
2014	Vassy, J. [105]	T2D	62	0.698	0.726	0.028	Age, sex	Caucasian, USA population
2014	Vassy, J. [105]	T2D	62	0.903	0.906	0.003	Sex, parental T2D, BMI, blood pressure, HDL cholesterol, triglyceride levels, age	Caucasian, USA population
2016	Läll, K. [83]	T2D-double weighted	1000	0.699	0.74	0.042	Sex, age	Caucasian
2016	Läll, K. [83]	T2D-dw	1000	0.718	0.767	0.049	Sex, age, BMI	Caucasian
2016	Läll, K. [83]	T2D-dw	1000	0.777	0.79	0.012	Sex, age, BMI, history of hypertension, and vegetable consumption	Caucasian
2018	Khera, A.V.[99]	T2D	7M	0.66	0.73	0.070	Sex, age	Caucasian

4. Polygenic risk scores for different diabetes subtypes

Oram and his team were the first group to use a PRS to differentiate between T1D and T2D [93]. They used 30 SNPs, including HLA and non-HLA loci, and got an accuracy of 0.88. The AUC for the PRS using the top 9 SNPs was 0.873 (table 6).

A group of researchers, led by Patel [29], used 30 SPNs to create a T1D PRS. The PRS was used to differentiate monogenic diabetes from T1D. The AUC of the PRS was found to be highly discriminant between the two disease states, being 0.87. A study by Yaghootkar and his team [35] provided the first evidence that the T1D PRS proposed by Oram [28], could help to distinguish monogenic diabetes from T1D in an Iranian population. AUC analysis showed that T1D-PRS in the non-European cohort strongly discriminated between monogenic and T1D with a score of 0.898, which was similar to the ability of the same PRS in the European cohort (Table 6).

Table 6. Comparison of accuracy of T1D PRS, to discriminate diabetes subtypes, assessed by the AUC[92].

Year	Autor	PRS	SNPs	AUC PRS	Ethnicity
2015	Oram, R. [93]	T1D vs. T2D	30	0.88	Caucasian
2015	Oram, R. [93]	T1D vs. T2D	9	0.87	Caucasian
2016	Patel, K. [94]	T1D vs. MODY	30	0.87	Caucasian
2019	Yaghootkar, H. [98]	T1D vs Monogenic diabetes	9	0.90	Iranian

5. Observational study: Subjects Characteristics

The table below shows the details of the 446 participants. About half of the participants were women, and the average age was 42.54 years. The median BMI was 26.87 kg/m^2, and median fasting insulin and glucose levels were 9.62 uU/ml and 93 mg/dL, respectively, indicative of a population at risk for diabetes (Table 7).

Table 7. Demographic characteristics and baseline measurements of studied participants from PolRed cohort

Characteristics and measurements	**All participants (n = 446)**
Female [n]	245 (54.9%)
Age (years)	42.54 (30.33, 55.73)
BMI (kg/m^2)	26.87 (24.04, 30.85)
FG (mg/dl)	101 (95, 110)
FI (uU/ml)	10.78 (8.50, 14.75)

6. Comparison of Metabolic Parameters at Baseline and Follow-up

Table 8 compares the prediabetic population's metabolic parameters at baseline and follow-up. The variables that were statistically significant (highest to lowest) after adjustment for multiple testing are: are fasting glucose, SF, FM, IPAQ, FFM, HbA1c, VF, two-hour glucose during OGTT (2-h glucose), VAT/SAT ratio, BMI, fasting insulin, MM, 2-hour insulin, and Chol. *P*-values of <0.05 are in bold and reflect significance after adjustment for multiple testing.

Table 8. Description and comparison of metabolic variables in the prediabetic cohort between baseline and the follow-up

Metabolic Parameter	**Baseline Median (IR)**	**Follow-up Median (IR)**	***P§***
BMI (kg/m^2)	26.87 (24.04, 30.85)	27.51 (24.36, 31.780)	**7.64E-13**
FFM (kg)	53.92 (48.74, 61.52)	50.74 (44.55, 62.16)	**1.13E-21**
FM (kg)	23.25 (20.06, 28.23)	26.82 (21.22, 36.14)	**1.73E-24**
MM (kg)	24.98 (21.17, 30.3)	24.85 (21.15, 32.58)	**0.0010**
VF (cm^3)	82.50 (65, 101)	112 (72.25, 152)	**3.94E-15**
SF (cm^3)	145.50 (117, 184)	249 (180.25, 318.75)	**1.32E-45**
VAT/SAT ratio	0.54 (0.44, 0.64)	0.42 (0.33, 0.56)	**3.82E-13**
IPAQ (min/week)	1344 (240, 4306)	5368 (2530, 11546)	**3.90E-24**
Fasting glucose (mg/dl)	101 (95, 110)	109 (98, 121)	**8.10E-46**
2-hour glucose (mg/dl)	103 (86, 114)	112 (94.25, 122.75)	**1.50E-14**
HbA1c (%)	5.40 (5.10, 5.70)	5.50 (5.20, 5.80)	**3.82E-15**
Fasting insulin (uU/ml)	10.78 (8.50, 14.75)	11.87 (9.40, 15.67)	**0.0006**
2-hour insulin (uU/ml)	29.98 (17.17, 53.84)	34.11 (22.23, 49.89)	**0.0154**
Chol (mg/dL)	188 (165, 221)	194 (169, 220)	0.0276
TG (mg/dL)	91 (67.25, 133.00)	96.90 (71.25, 143.00)	0.0918
HDL (mg/dL)	59.70 (50.42, 68.00)	57.5 (47, 70)	0.2003
LDL(mg/dL)	109.60 (83.25, 137.40)	108.80 (87.40, 138.30)	0.0760

7. Construction of Polygenic Scores for T2D

In Tables 9 and 10, the genetic variants included in the PRSs are described, resulting in a T2D PRS of 68 SNPs and an obesity PRS of 21 SNPs. The mean for the T2D PRS in the prediabetic cohort was 1.03 (range: 0.23–1.64) with an SD of 0.30 (Figure 3A). The mean for the obesity PRS in the prediabetic cohort was 1.37 (range: 0.45–2.24) with an SD of 0.35 (Figure 3B).

Table 9. Genetic variants included in the T2D PRS

SNP	Locus	CHR	Pos	Alt Allel	β	Odd Rati	MAF	P-value	N
rs3101336	*NEGR1*	1	72751185	C	0.018	1.018	0.450	3.07E-02	939912
rs2568958	*NEGR1*	1	72765116	A	0.022	1.022	0.438	1.64E-04	1082380
rs2815752	*NEGR1*	1	72812440	A	0.022	1.022	0.388	4.05E-04	1084120
rs10913469	*SEC16B*	1	177913519	C	0.032	1.033	0.163	1.74E-12	1073630
rs340874	*PROX1*	1	214159256	C	0.047	1.048	0.475	6.11E-32	1081640
rs2605100	*LYPLAL1*	1	219644224	G	0.030	1.031	0.213	3.75E-07	1090350
rs12022722	*LYPLAL1*	1	219651133	T	0.020	1.020	0.413	7.57E-06	917427
rs2820464	*LYPLAL1*	1	219693220	A	-0.029	0.972	0.175	4.82E-10	927501
rs2785980	*LYPLAL1*	1	219700519	C	-0.031	0.969	0.175	1.30E-13	1039880
rs4846567	*SLC30A10*	1	219750717	T	-0.033	0.967	0.125	1.17E-15	1078840
rs6548238	*TMEM18*	2	634905	C	0.054	1.055	0.113	1.66E-21	1083220
rs7561317	*TMEM18*	2	644953	G	0.056	1.058	0.238	2.67E-22	1023140
rs780094	*GCKR*	2	27741237	C	0.053	1.055	0.488	2.05E-34	1082280
rs13389219	*GRB14*	2	165528876	T	-0.054	0.947	0.400	5.98E-35	1080230
rs7607980	*GRB14*	2	165551201	C	-0.068	0.935	0.138	1.75E-19	851502
rs1801282	*PPARG*	3	12393125	G	-0.022	0.978	0.263	4.85E-02	18252
rs11708067	*ADCY5*	3	123065778	G	-0.072	0.930	0.100	1.06E-44	756464
rs11920090	*SLC2A2*	3	170717521	A	-0.027	0.974	0.188	1.47E-03	952836
rs4402960	*IGF2BP2*	3	185511687	T	0.096	1.101	0.425	4.15E-122	1088690
rs7647305	*ETV5*	3	185834290	C	0.026	1.027	0.213	7.70E-07	965163
rs10938397	*GNPDA2*	4	45182527	G	0.038	1.039	0.388	5.11E-17	1048050
rs10946398	*CDKAL1*	6	20661034	C	0.113	1.120	0.375	5.97E-144	1085220
rs2844479	*AIF1*	6	31572956	C	0.026	1.027	0.500	4.49E-09	1078460
rs2260000	*PRRC2A*	6	31593476	G	0.020	1.020	0.375	2.71E-05	930840
rs1077393	*BAG6*	6	31610529	G	0.019	1.019	0.500	1.41E-04	1081460
rs2191349	*DGKB*	7	15064309	T	0.051	1.053	0.488	1.23E-41	1081160
rs4607517	*YKT6*	7	44235668	A	0.032	1.032	0.088	2.65E-10	1082060
rs4731702	*KLF14*	7	130433384	T	-0.037	0.964	0.475	1.09E-20	1075480
rs972283	*KLF14*	7	130466854	G	0.035	1.036	0.463	5.05E-20	1067780
rs13266634	*SLC30A8*	8	118184783	T	-0.091	0.913	0.388	2.45E-124	1181190

rs11558471	*SLC30A8*	8	118185733	G	-0.088	0.916	0.313	8.78E-106	1082770
rs7034200	*GLIS3*	9	4289050	A	0.040	1.041	0.488	1.20E-23	1070410
rs10811661	*GLIS3*	9	22134094	C	-0.140	0.869	0.113	2.88E-146	930908
rs1111875	*HHEX*	10	94462882	T	-0.093	0.912	0.388	7.30E-95	938378
rs5015480	*HHEX*	10	94465559	T	-0.089	0.915	0.388	1.64E-128	1069430
rs7923837	*HHEX*	10	94481917	A	-0.094	0.911	0.350	1.62E-75	888616
rs10885122	*ADRA2A*	10	113042093	G	0.024	1.024	0.125	1.03E-05	1082070
rs7901695	*TCF7L2*	10	114754088	C	0.207	1.230	0.313	5.82E-226	209644
rs4506565	*TCF7L2*	10	114756041	T	0.216	1.241	0.313	6.14E-229	208956
rs7903146	*TCF7L2*	10	114758349	T	0.235	1.265	0.275	2.64E-97	249463
rs5215	*KCNJ11*	11	17408630	T	-0.056	0.945	0.363	5.74E-52	1192080
rs4074134	*BDNF*	11	27647285	T	-0.023	0.978	0.150	1.59E-05	937741
rs4923461	*BDNF*	11	27656910	G	-0.019	0.981	0.150	3.05E-05	931716
rs925946	*BDNF*	11	27667202	G	0.000	1.000	0.363	2.44E-05	951966
rs10501087	*BDNF*	11	27670108	C	-0.023	0.977	0.150	2.81E-06	934,105
rs6265	*BDNF*	11	27679916	T	-0.022	0.978	0.138	9.00E-07	1169040
rs10835211	*BDNF*	11	27701365	A	0.022	1.103	0.213	1.83E-02	818173
rs11605924	*CRY2*	11	45873091	C	-0.021	0.979	0.450	4.96E-08	964253
rs10838738	*MTCH2*	11	47663049	G	0.013	1.013	0.363	5.75E-03	1085580
rs174550	*FADS1*	11	61571478	C	-0.027	0.973	0.350	1.20E-10	1019910
rs10830963	*MTNR1B*	11	92708710	G	0.061	1.063	0.225	7.36E-43	1059540
rs7138803	*BCDIN3D*	12	50247468	A	0.023	1.023	0.425	1.35E-09	1089630
rs8049439	*ATXN2L*	16	28837515	C	0.019	1.019	0.413	2.76E-05	1191610
rs4788102	*SH2B1*	16	28873398	A	0.022	1.022	0.413	8.70E-06	933067
rs7498665	*SH2B1*	16	28883241	G	0.022	1.023	0.400	1.86E-06	1024050
rs6499640	*FTO*	16	53769677	A	0.027	1.027	0.425	5.97E-08	1080780
rs8050136	*FTO*	16	53816275	A	0.094	1.098	0.488	8.61E-102	1090360
rs3751812	*FTO*	16	53818460	T	0.096	1.100	0.488	2.47E-76	935802
rs9939609	*FTO*	16	53820527	A	0.092	1.096	0.488	5.44E-131	1063990
rs7190492	*FTO*	16	53828752	G	0.048	1.049	0.325	8.44E-18	937716
rs8044769	*FTO*	16	53839135	C	0.061	1.062	0.438	2.73E-31	937714
rs633265	*MC4R*	18	57831468	T	0.039	1.040	0.413	1.05E-18	931300
rs1350341	*MC4R*	18	57842533	A	0.038	1.039	0.400	5.37E-18	926826
rs17782313	*MC4R*	18	57851097	C	0.055	1.057	0.238	1.46E-35	1088670
rs12970134	*MC4R*	18	57884750	A	0.050	1.051	0.300	2.83E-28	1089600
rs29941	*KCTD15*	19	34309532	G	0.001	1.001	0.338	1.36E-04	1082030
rs11084753	*KCTD16*	19	34322137	G	0.011	1.011	0.363	2.61E-03	1047670
rs2287019	*QPCTL*	19	46202172	T	0.021	1.021	0.213	1.00E-04	1149250

CHR = Chromosome, Pos = position based on human genome 19, MAF = minor allele frequency, P value = statistical significant associations of SNPS in GWAS for Obesity[90], N = Effective sample size from all the dataset where the P-value was significant.

Table 10. Genetic variants included in the Obesity PRS

SNP	Locus	CHR	Pos	Alt Allele	β	Odd Ratio	MAF	P-value	N
rs7561317	*TMEM18*	2	644953	G	0.141	1.151	0.2451	1.38E-04	11743
rs10938397	*GNPDA2*	4	45182527	G	0.099	1.104	0.4549	2.60E-04	11743
rs7903146	*TCF7L2*	10	114758349	T	-0.101	0.904	0.2280	3.16E-03	11743
rs4074134	*BDNF*	11	27647285	T	-0.163	0.850	0.2280	7.37E-06	11743
rs4923461	*BDNF*	11	27656910	G	-0.163	0.850	0.2232	6.98E-06	11743
rs925946	*BDNF*	11	27667202	G	-0.088	0.915	0.2610	1.88E-03	11743
rs10501087	*BDNF*	11	27670108	C	-0.162	0.850	0.2220	8.07E-06	11743
rs6265	*BDNF*	11	27679916	T	-0.164	0.849	0.2061	1.17E-05	11743
rs10835211	*BDNF*	11	27701365	A	0.098	1.103	0.1915	2.04E-03	11743
rs8049439	*ATXN2L*	16	28837515	C	0.087	1.091	0.4305	1.39E-03	11743
rs4788102	*SH2B1*	16	28873398	A	0.094	1.099	0.4256	5.71E-04	11743
rs7498665	*SH2B1*	16	28883241	G	0.094	1.099	0.4207	5.65E-04	11743
rs8050136	*FTO*	16	53816275	A	0.205	1.228	0.4915	9.23E-14	11743
rs3751812	*FTO*	16	53818460	T	0.206	1.229	0.4902	6.42E-14	11743
rs9939609	*FTO*	16	53820527	A	0.207	1.231	0.4890	5.07E-14	11743
rs7190492	*FTO*	16	53828752	G	0.126	1.134	0.3232	5.28E-06	11743
rs8044769	*FTO*	16	53839135	C	0.162	1.176	0.4195	1.76E-09	11743
rs9921518	*IRX3*	16	54494424	G	-0.105	0.900	0.2488	6.11E-03	11743
rs17782313	*MC4R*	18	57851097	C	0.102	1.107	0.1768	3.42E-03	11743
rs12970134	*MC4R*	18	57884750	A	0.099	1.104	0.2134	3.66E-03	11743
rs29941	*KCTD15*	19	34309532	G	0.061	1.063	0.3073	2.77E-02	11743

CHR = Chromosome, Pos = position based on human genome 19, MAF = minor allele frequency, P value = statistical significant associations of SNPS in GWAS for Obesity[90] ,N = Effective sample size, DataSet: FinnGen 2018 GWAS: European ancestry[126].

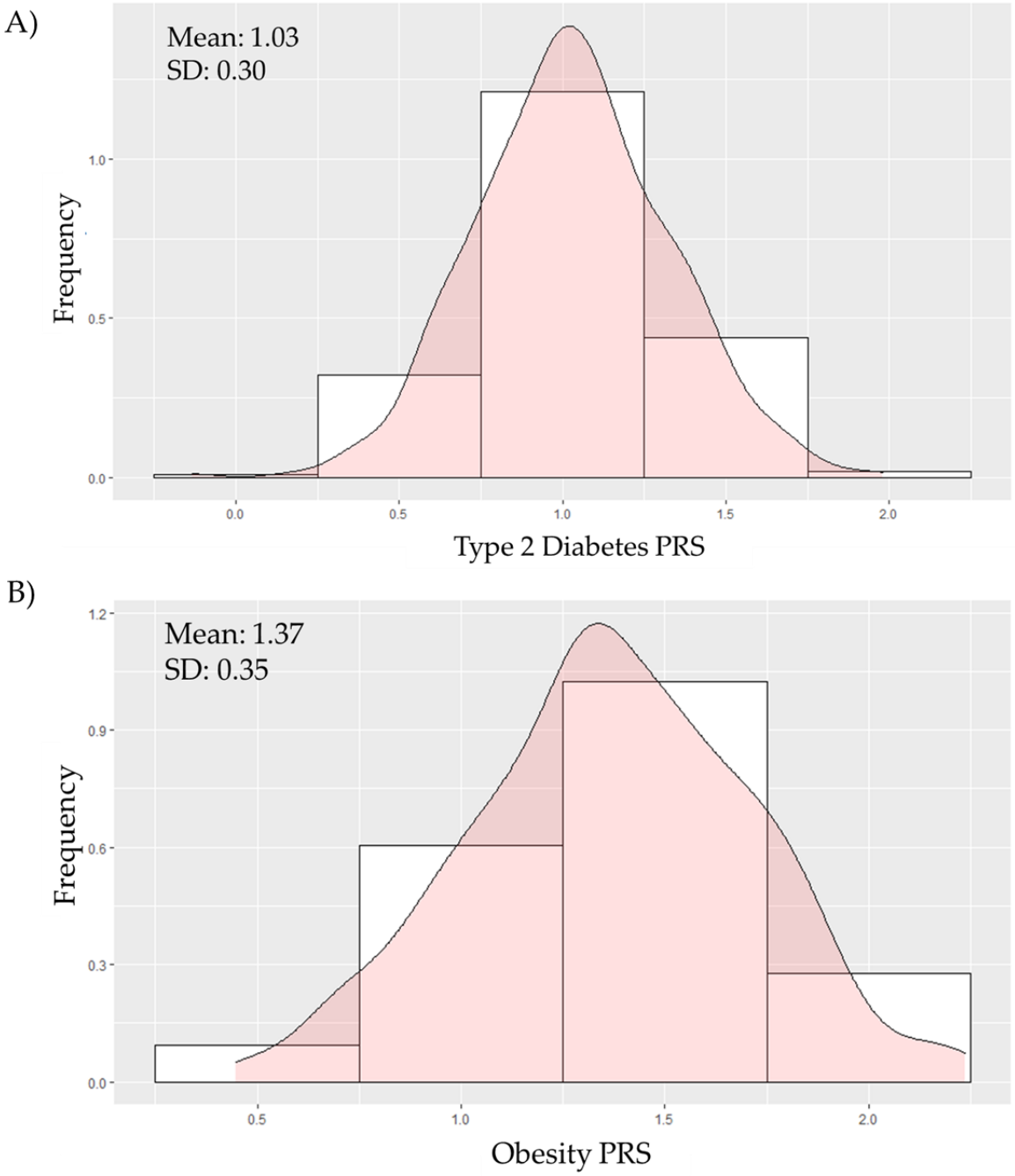

Figure 3. Distribution of polygenic risk scores (PRSs) for T2D (A) and obesity (B) across 446 prediabetic individuals in PolRed.

8. Association Between T2D PRS and Baseline Metabolic Parameters

Table 11 shows the associations between T2D PRS and metabolic parameters at baseline. At a nominal significance of P<0.05 after adjustment for multiple testing, the T2D PRS was associated with BMI, FFM, FM, fasting glucose, and fasting insulin. After testing, the T2D PRS was not significantly associated with MM, VF, SF, VAT/SAT ratio, IPAQ, 2-hour glucose, HbA1c, 2-hour insulin, Chol, TG, HDL, or LDL (Table 10). β values are reported per SD of PRS. *P*-values of <0.05 are in bold and reflect significance after adjustment for multiple testing.

Table 11. Association of T2D PRS with the baseline metabolic parameters in PolRed.

Metabolic Parameter	**β (95% CI)**	***P§***
BMI (kg/m^2)	0.0066 (-0.0020, 0.0112)	**0.0053**
FFM (kg)	0.0029 (-0.0001, 0.0056)	**0.0394**
FM (kg)	0.0028 (-0.0003, 0.0053)	**0.0293**
MM (kg)	-0.0045 (-0.0091, 0.0001)	0.0535
VF (cm^3)	-0.0004 (-0.0009, 0.0001)	0.0662
SF (cm^3)	-0.0003 (-0.0007, 0.0001)	0.1360
VAT/SAT ratio	-0.0544 (-0.1436, 0.0348)	0.2310
IPAQ (min/week)	0.0002 (-0.0002, 0.0005)	0.4450
Fasting glucose (mg/dl)	0.0024 (-0.0001, 0.0047)	**0.0451**
2-hour glucose (mg/dl)	0.0001 (-0.0008, 0.0009)	0.8990
HbA1c (%)	-0.0171 (-0.0629, 0.0288)	0.4650
Fasting insulin (uU/ml)	0.0033 (-0.0003, 0.0063)	**0.0320**
2-hour insulin (uU/ml)	-0.0001 (-0.0008, 0.0006)	0.7740
Chol (mg/dL)	0.3426 (-0.8705, 1.5558)	0.5790
TG (mg/dL)	-0.0683 (-0.3109, 0.1744)	0.5800
HDL (mg/dL)	-0.3415 (-1.5546, 0.8716)	0.5800
LDL (mg/dL)	-0.3428 (-1.556, 0.8704)	0.5790

9. Association Between Obesity PRS and Baseline Metabolic parameters

The associations between obesity PRS and baseline metabolic parameters are shown in Table 12. After adjustment for multiple testing, obesity PRS was associated with BMI with nominal significance (P<0.05). Since the beta coefficient was positive, each SD increase in PRS was associated with a rise of 0.0141 kg/m2 in BMI. After testing, no significant associations were found between obesity PRS and FFM, FM, MM, VF, SF, VAT/SAT ratio, IPAQ, Fasting glucose, 2-hour glucose, HbA1c, fasting insulin, 2-hour insulin, Chol, TG, HDL and LDL(Table 12). β values are reported per SD of PRS. P§ values of 0.05 are in bold and reflect significance after adjustment for multiple testing.

Table 12. Association of Obesity PRS with the baseline metabolic parameters in PolRed

Metabolic Parameter	**β (95% CI)**	***P§***
BMI (kg/m^2)	0.0141 (-0.0028, 0.0255)	**0.0145**
FFM (kg)	0.0027 (-0.0259 , 0.0312)	0.8539
FM (kg)	0.0043 (-0.0039 , 0.0126)	0.3016
MM (kg)	-0.0057 (-0.0522 , 0.0409)	0.8111
VF (cm^3)	0.0013 (-0.0006 , 0.0032)	0.1773
SF (cm^3)	-0.0008 (-0.0020 , 0.0005)	0.2489
VAT/SAT ratio	-0.2228 (-0.4980 , 0.0523)	0.1121
IPAQ (min/week)	-0.0001 (-0.0006 , 0.0003)	0.5656
Fasting glucose (mg/dl)	-0.0033 (-0.0073 , 0.0006)	0.0978
2-hour glucose (mg/dl)	0.0003 (-0.0012 , 0.0019)	0.6857
HbA1c (%)	0.0371 (-0.0256 , 0.0998)	0.2449
Fasting insulin (uU/ml)	0.0017 (-0.0029 , 0.0063)	0.4560
2-hour insulin (uU/ml)	0.0004 (-0.0009 , 0.0017)	0.5303
Chol (mg/dL)	0.8328 (-0.5718 , 2.2374)	0.2445
TG (mg/dL)	-0.1660 (-0.4469 , 0.1149)	0.2460
HDL (mg/dL)	-0.8329 (-2.2374 , 0.5716)	0.2444
LDL(mg/dL)	-0.8327 (-2.2373 ,0.5720)	0.2445

10. Association Between T2D PRS and Changes in Metabolic Parameters

Table 13 shows the associations between the T2D PRS and the changes in metabolic parameters after follow-up. The T2D PRS was associated with Δ FM at a nominal significance of P<0.05 after adjustment for multiple testing. As the beta coefficient is positive, for every SD increasing in the PRS, Δ FM will increase 0.0049 kg. After the testing, a significant association was not found between T2D PRS and Δ in FFM, MM, VF, SF, VAT/SAT ratio, IPAQ, Fasting glucose, 2-hour glucose, HbA1c, Fasting insulin, 2-hour insulin, Chol, TG, HDL and LDL(Table 13). β values are reported per SD of PRS. P§ values of 0.05 are in bold and reflect significance after adjustment for multiple testing.

Table 13. Association of T2D PRS with changes in metabolic parameters after follow-ups with the prediabetic cohort in PolRed

Metabolic Parameter	β (95% CI)	*P§*
Δ FFM (kg)	0.0017 (-0.0029, 0.0063)	0.462
Δ FM (kg)	0.0049 (-0.0006, 0.0091)	**0.025**
Δ MM (kg)	0.0001 (-0,0004, 0.0002)	0.548
Δ VF (cm^3)	0.0001 (-0.0009, 0.0012)	0.802
Δ SF (cm^3)	0.0001 (-0.0004, 0.0006)	0.738
Δ VAT/SAT ratio	-0.0369 (-0.1955, 0.1216)	0.647
Δ IPAQ (min/week)	0.0001 (-0,0004, 0.0002)	0.269
Δ Fasting glucose (mg/dl)	-0.0010 (-0.0034, 0.0013)	0.394
Δ 2-hour glucose (mg/dl)	-0.0004 (-0.0015, 0.0007)	0.467
Δ HbA1c (%)	0.0492 (-0.0242, 0.1226)	0.188
Δ Fasting insulin (uU/ml)	0.0011 (-0.0024, 0.0045)	0.548
Δ 2-hour insulin (uU/ml)	0.0002 (-0.008, 0.0012)	0.650
Δ Chol (mg/dL)	0.0063 (-0.0133, 0.0259)	0.531
Δ TG (mg/dL)	-0.0013 (-0.0053, 0.0026)	0.507
Δ HDL (mg/dL)	-0.0066 (-0.0261, 0.0130)	0.511
Δ LDL(mg/dL)	-0.0065 (-0.0261, 0.0132)	0.518
Δ FFM (kg)	0.0017 (-0.0029, 0.0063)	0.462

11. Association Between Obesity PRS and Changes in Metabolic Parameters

Table 14 summarizes the associations between the obesity PRS and changes in metabolic parameters after the follow-up. A high obesity PRS is associated with Δ FM and Δ 2-hour glucose at a nominal significance of $P<0.05$ after adjustment for multiple testing. As their beta coefficients are positive, for every increase in SD in the PRS, Δ FM increased by 0.0056 kg, and Δ 2-hour glucose increased by 0.0013 mg/dl. After testing, obesity PRS was not significantly associated with Δ FFM, MM, VF, SF, VAT/SAT ratio, IPAQ, fasting glucose, HbA1c, Fasting insulin, 2-hour insulin, Chol, TG, HDL, or LDL(Table 14). β values are reported per SD of PRS. P§ values of 0.05 are in bold and reflect significance after adjustment for multiple testing.

Table 14. Association of obesity PRS with changes in metabolic parameters after follow-up in PolRed

Metabolic Parameter	β (95% CI)	*P§*
Δ FFM (kg)	0.0021 (-0.0032, 0.0074)	0.4383
Δ FM (kg)	0.0056 (-0.0008, 0.0105)	**0.0231**
Δ MM (kg)	0.0002 (-0.0002, 0.0005)	0.3184
Δ VF (cm^3)	0.0002 (-0.0002, 0.0005)	0.7600
Δ SF (cm^3)	0.0001 (-0.006, 0.0006)	0.8850
Δ VAT/SAT ratio	0.0205 (-0.1619, 0.2029)	0.8252
Δ IPAQ (min/week)	0.0001 (-0.0004, 0.0002)	0.4108
Δ Fasting glucose (mg/dl)	0.0020 (-0.0007, 0.0047)	0.1446
Δ 2-hour glucose (mg/dl)	0.0013 (-0.0001, 0.0026)	**0.0341**
Δ HbA1c (%)	0.0014 (-0.0830, 0.0859)	0.9732
Δ Fasting insulin (uU/ml)	-0.0024 (-0.0064, 0.0016)	0.2316
Δ 2-hour insulin (uU/ml)	0.0007 (-0.0004, 0.0018)	0.2273
Δ Chol (mg/dL)	0.0168 (-0.0057, 0.0394)	0.1434
Δ TG (mg/dL)	-0.0036 (-0.0081, 0.0009)	0.1167
Δ HDL (mg/dL)	-0.0158 (-0.0383, 0.0067)	0.1685
Δ LDL(mg/dL)	-0.0164 (-0.0390, 0.0061)	0.1533
Δ FFM (kg)	0.0021 (-0.0032, 0.0074)	0.4383

12. Association of Genotypes Frequencies with Changes in Metabolic Parameters

Tests were run to check if the genotypes' frequencies significantly affected all 17 metabolic parameters. All the SNPs included in the T2D PRSs and obesity PRSs (69 SNPs in total) were analyzed, and the metabolic parameters were stratified by investigated genotypes with a significant association or a tendency (Table 15). No significant deviation from the Hardy-Weinberg equilibrium was reported for any investigated SNPs.

It was observed that AA genotype carriers of rs10838738 presented statistically significantly difference at Δ 2-hour glucose and Δ 2-hour insulin (Table 15). No other significant differences were observed between the different genotypes; however, a tendency toward a lower Δ FM and Δ VF was noticed in GG genotype carriers of rs2260000. Between carriers of investigated genetic variants in rs7647305, a trend in Δ 2-hour glucose and Δ IPAQ was seen. Another tendency toward a lower Δ 2-hour glucose in GG genotype carriers of rs29941 was noticed (Table 15).

Table 15. Description and comparison of the prediabetic cohort participants stratified by *rs10838738, rs2260000, rs7647305,* and *rs29941* genotypes.

rs10838738	**A/A**	**A/G**	**G/G**	***P§***
N	143	229	83	
Δ 2-hour glucose (mg/dl)	19 (3, 35)	13 (-6.75, 33)	8 (-8, 23.5)	**0.017**
Δ 2-hour insulin (uU/ml)	5.88 (-5.85, 21.34)	3.74 (-10.25, 22.49)	-1.68 (-26.02, 10.02)	**0.001**
rs2260000	**A/A**	**A/G**	**G/G**	***P§***
N	151	223	81	
Δ FM (kg)	5.27 (0.50, 10.17)	4.87 (1.32, 8.96)	2.40 (-1.92, 6.45)	0.051
Δ VF (cm3)	35.75 (-8.25, 73)	29 (-6, 67)	8 (-29.50, 53.50)	0.068
rs7647305	**C/C**	**C/T**	**T/T**	***P§***
N	309	126	20	
Δ 2-hour glucose (mg/dl)	13.3 (-1, 35)	10.5 (-11.8, 25)	9 (-30, 27.5)	0.077
Δ IPAQ (min/week)	2712.5 (-604.9, 6962.2)	2601 (-23.3, 9537.3)	6463 (2470, 12901)	0.077
rs29941	**A/A**	**A/G**	**G/G**	***P§***
N	38	204	213	
Δ 2-hour glucose (mg/dl)	20.5 (12, 33)	15 (0.75, 36.25)	9.50 (-9.50, 26.25)	0.068

The AA genotype carriers of rs10838738 presented significantly greater Δ 2-hour glucose (Figure 4A) than the AG and GG genotypes. The Δ in 2-hour insulin (Figure 4B) had a significantly smaller difference when comparing the GG genotype and the AG and AA genotypes. By analyzing the differences between the rs2260000 genotypes, we observed that the GG genotype carriers presented a significantly lower Δ FM (Figure 5A) and Δ VF (Figure 5B) compared to the AA and AG genotype carriers. The CC genotype carriers of rs7647305 presented significantly greater Δ in 2-hour glucose (Figure 6A) than CT genotypes. The Δ IPAQ (Figure 6B) had a significantly lower difference when comparing the CC genotype to the TT genotype. The differences between rs29941 genotypes show that AG genotype carriers presented significantly greater Δ in 2-hour glucose (Figure 7) than GG genotype carriers did.

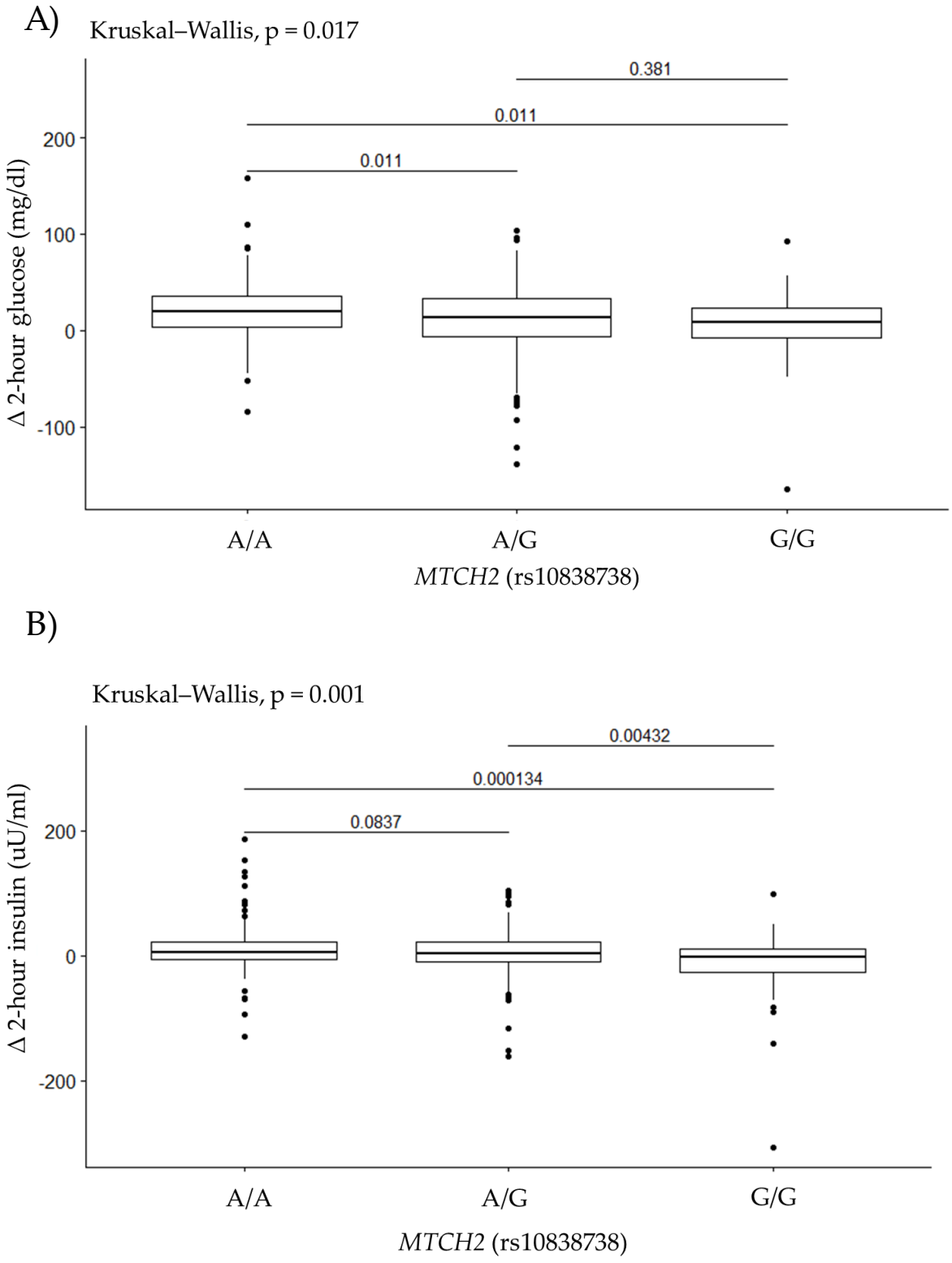

Figure 4. Association of genotype rs10838738 with (**A**) Δ 2-hour glucose (mg/dl) and (**B**) Δ 2-hour insulin (uU/ml).

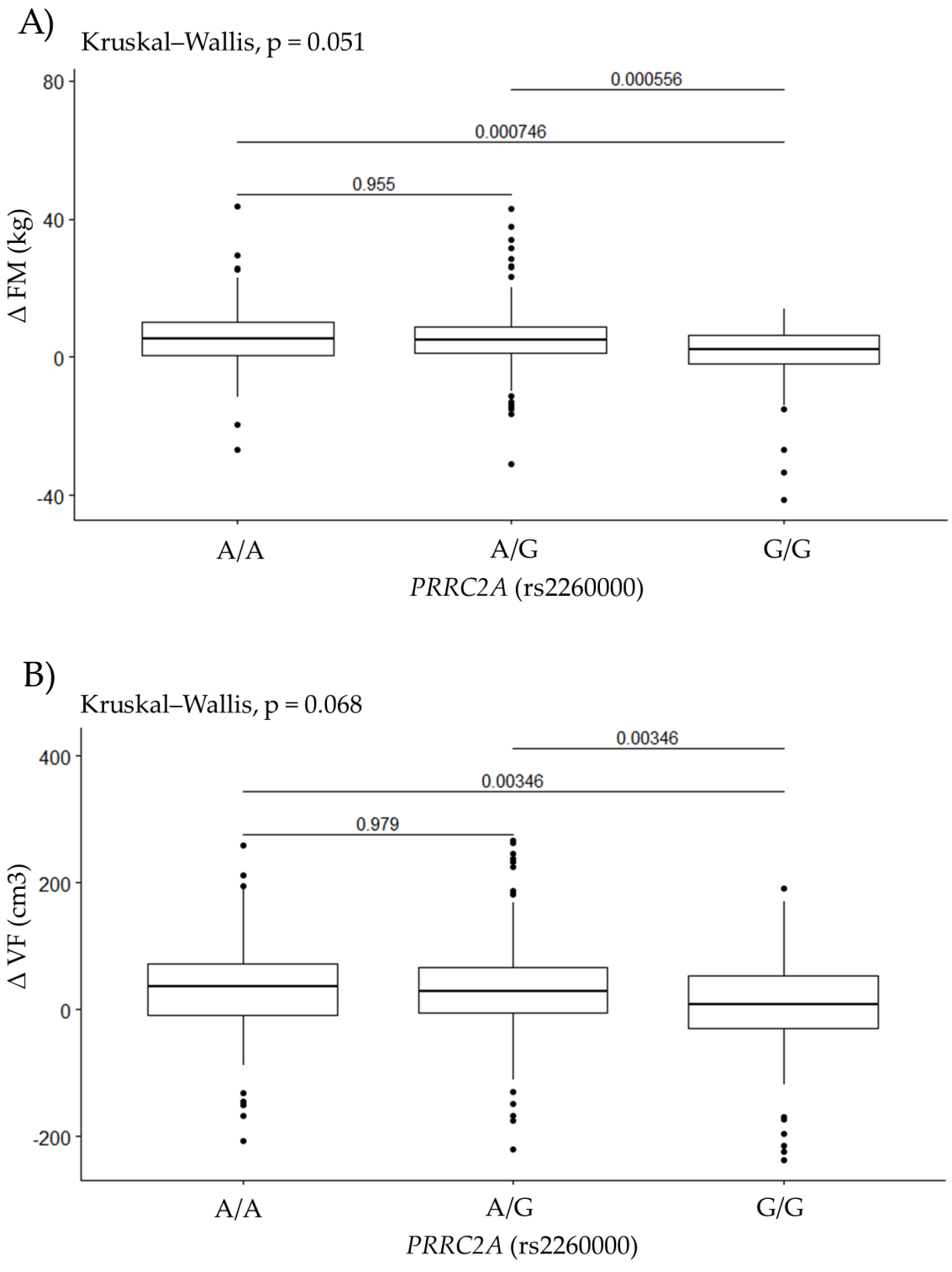

Figure 5. Association of genotype rs2260000 with (**A**) Δ FM (kg) and (**B**) Δ VF (cm3).

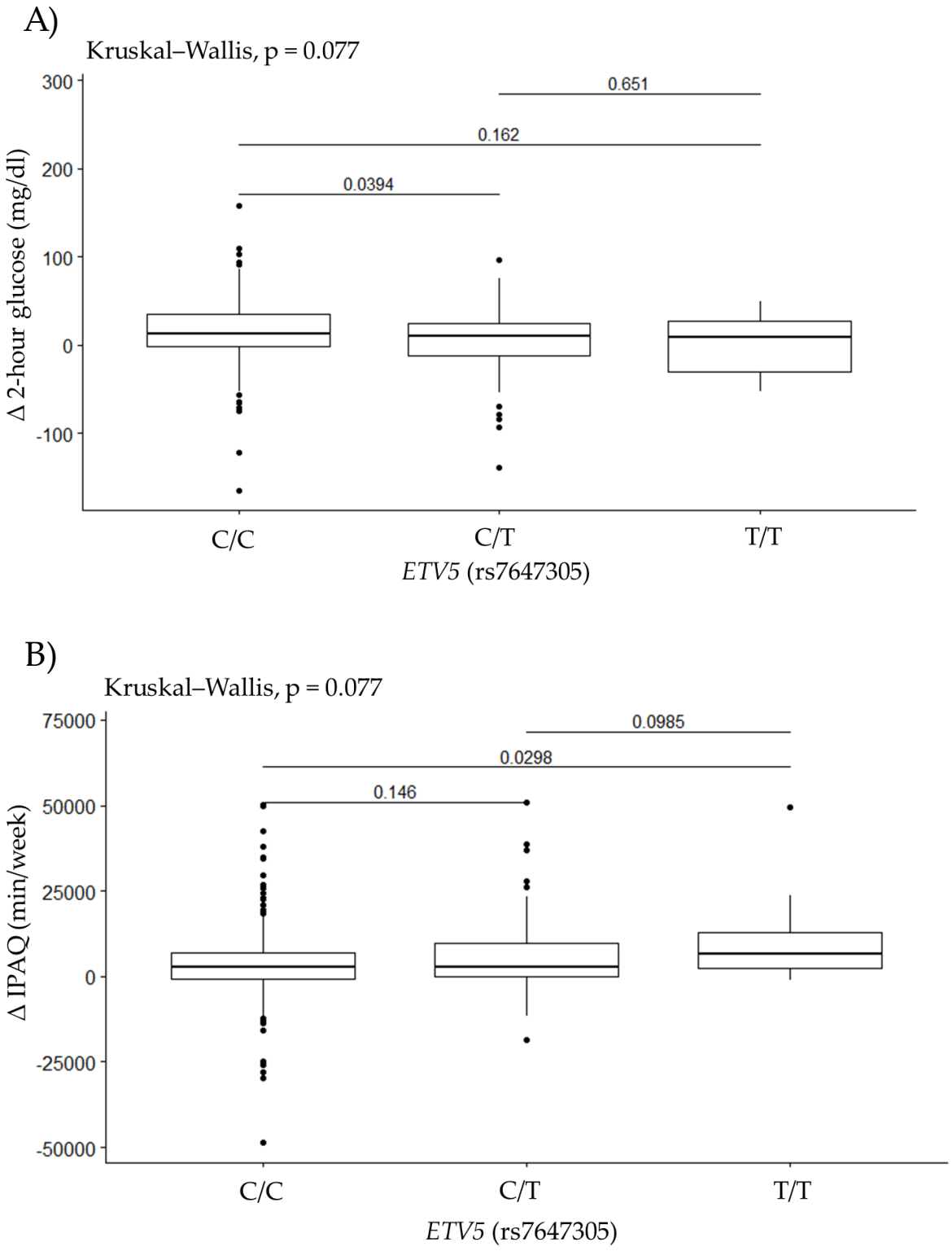

Figure 6. Association of genotype rs7647305 with (**A**) Δ 2-hour glucose (mg/dl) and (**B**) Δ IPAQ (min/week).

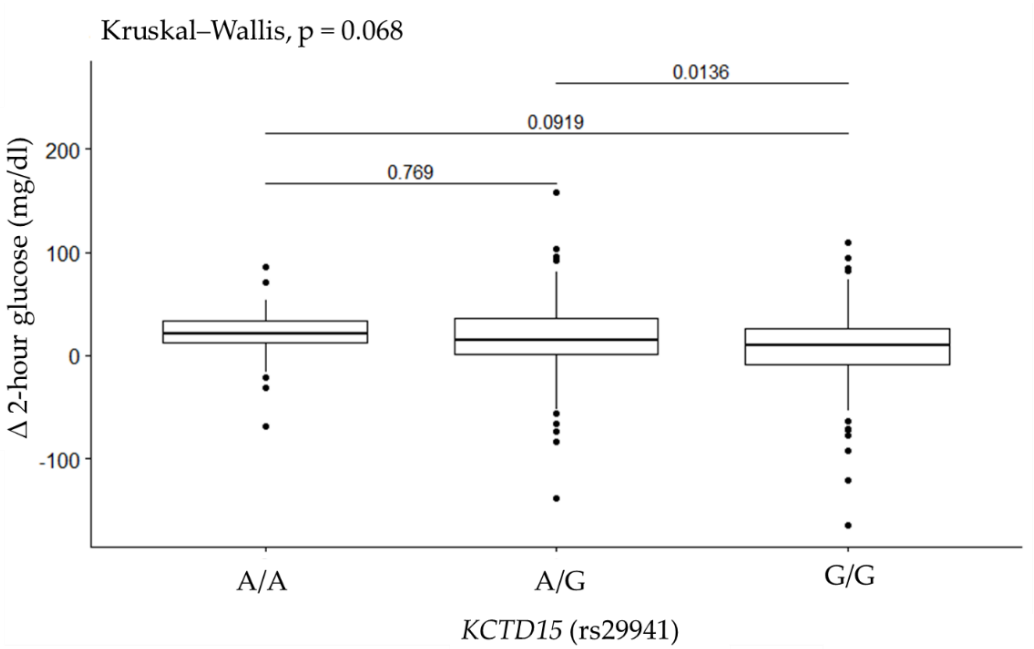

Figure 7. Association of genotype rs29941 with Δ 2-hour glucose (mg/dl)

VII. Discussion

1. Systematic review

The PRS support the diagnosis of the disease; they are consistent throughout life; thus they could be an effective tool to determine whether a particular patient has any form of diabetes. Thanks to them, it is less difficult to predict the risk of prediabetes [127]. As previously noted in the analysis of T1D PRS in **table 4**, except for the cohort of African-Americans [96], the AUC values were more significant than 0.80, implying that all PRSs had sufficient sensitivity and specificity to be able to differentiate patients with T1D. Genetic factors contribute significantly to the prevention of T1D by providing a reliable risk score. PRS for T1D can diagnose young diabetic patients in European cohorts requiring insulin therapy, which is essential for accurately classifying patients when clinical factors are misdiagnosed. On **table 6,** the use of PRS as a tool to distinguish diabetes subtypes as an additional advantage is shown. The latest studies have shown that T1D PRS are excellent at distinguishing patients with T1D from T2D[93] and Monogenic diabetes [94]. T1D PRS that has been validated for a Caucasian population might be possible to use in other ethnicities. This is being researched and if true could be a good option [98]. Proper diagnosis of T1D could come about as a result of using the PRS and could help lead to lifestyle changes and medications to slow down the progression of T1D. After a thorough review, it was found that the PRSs need improvements to be useful in a real-world application. As shown in **Table 5**, all of the studies in the last decade have said that clinical risk factors can predict T2D well and that there is almost no improvement when adding the PRS. Therefore more research needs to be done to fully understand how the PRS can be useful in diagnosing T2D.

There are several obstacles to overcome for the PRSs to be taking part in clinical diagnosis, the first one is the lack of innovation in the generation of PRS for T2D. The primary goal of developing a PRS is to predict who will get T2D[128,129], which can be improved by using newer, more optimized equations in the logistic regression portion of the formula[130]. There are two ways to build a PRS model: tree-based and logistic regression based[131,132]. Regression-based methods use either polynomial parametric regression or non-parametric regression to create a line of data input and output. Tree-based methods use a binary split rule to create a correlation between data input and output[133–135]. Both of these methods have their flaws, but the tree-based is the preferred method because it is more accurate and has been used in risk prediction for cardiovascular diseases[136–139]. Combined with GWAS

data, machine learning techniques can improve polygenic trait prediction [140]. Another problem that can arise from population heterogeneity being underestimated is overfitting. The PRS must be calibrated, validated, and optimized for each separate population before it can be proven that the PRS is not fitting the data too well, producing results that are inaccurate[141]. To avoid producing an overfit prediction model, the point of reference is to use data outside of the sample set[26,142]. Most PRS created using Caucasian GWAS show bias due to allele drift compared to other ethnic groups, even when utilizing the same variants [75,143]. Therefore, it is necessary to develop generalized risk prediction methods and include more diverse participants in risk score studies[75]. It might be necessary to adapt an existing T2D PRS validated in Caucasian participants to other ethnicities to avoid overfitting and obtaining false positive results. The last problem taking our attention is the environmental effect on genetic studies. Gene-environment interactions (GxE) refer to the fact that the effect of genes on a disease can be different in different environments[144]. In most GWAS, it is assumed that there are no GxE interactions, which could mean that clinically significant risk factors might go undetected[145]. The environmental impact can be a bias in developing a T2D PRS. There are benefits to analyzing GxE interactions, such as discovering new loci of disease susceptibility [146–148]. These interactions have been proven to be identified well by PRS approaches[149,150].

2. Polygenic risk scores for T2D and obesity

Studies have shown that patients with prediabetes may develop coronary artery disease [151] and diastolic heart failure [152] before overt T2D. However, prediabetes can be managed by changing the patient's habits [153]. Therefore, it is important to identify patients with prediabetes and take appropriate measures to optimize glycemic control [42]. From the PolReD study, 446 subjects who were prediabetic but did not have a diagnosis of T2D at baseline, were selected. The subjects' characteristics at baseline, are described in **table 7** and give us an insight into the general information of the population. The median age of the prediabetes cohort was 42.54 years, which is consistent with previous findings showing that prediabetes is more common in middle age (25-44 years) and diabetes is more common in the 45-60 year age group[154]. It shows that the prevalence of diabetes increases with age. The median BMI was 26.87 kg/m^2, placing the cohort in the overweight zone ($25 < BMI < 30$)[155]. The median values of fasting glucose and fasting insulin are 101 mg/dl and 10.78 uU/ml, respectively. According to Polish Diabetes Association and American Diabetes Association, prediabetes has been defined by the presence of impaired fasting glucose (IFG)

and/or impaired glucose tolerance (IGT) and/or HbA1C (5.7-6.4%). IFG is defined as fasting plasma glucose (FPG) levels from 100 to 125 mg/dL (from 5.6 to 6.9 mmol/l) and IGT as 2-h plasma glucose (PG) levels during 75-g OGTT from 140 to 199 mg/dL.

After comparing the metabolic parameters at baseline and follow-up, using the Wilcoxon signed-rank test for paired samples, 13 of the 17 parameters were statistically significant after adjustment for multiple testing, as seen in **table 8**. The increase of the values on the metabolic parameters is expected as studies have shown that with aging, metabolic disturbances progress, especially when they are not treated, like in the case of our cohort [156]. From the 13 statically significant metabolic parameters, FFM, MM, and VAT/SAT ratio decrease their median values, and BMI, FM, VF, SF, IPAQ, Fasting glucose, 2-hour glucose, HbA1c, Fasting insulin and 2-hour insulin, increase their median values. The four parameters that are not statistically significant were Chol, TG, HDL and LDL. When comparing the values of the cohort to the guidelines of lipid profile, we can describe them as optimal for Chol (125 to 200mg/dL), TG (less than 150 mg/dL) and HDL (40 to 60 mg/dL), and above optimal (100 to 129 mg/dL) for LDL [157]. When these parameters' values are imbalanced with the normal ones, the patient can suffer from dyslipidemia. This condition can result from diet, tobacco exposure, or genetics and can lead to cardiovascular disease with severe complications [158]. Lipid profile in prediabetes has been of research relevance recently. In contrast with our data, several studies on the Asian population conclude that prediabetics had a deranged lipid profile compared to normal healthy subjects [159,160]. This gives an exciting lead to continue studying lipids and their association with prediabetes in European cohorts.

An approach for converting genetic data to a predictive measure of disease susceptibility is to add the risk effects of loci into a PRS. Following the guidelines published in 2020 [73], the construction of PRSs for T2D and obesity using genome-wide significant variants found in GWAS for T2D and obesity in PolRed's data was achieved. The methodology followed gave, as a result a weighted PRSs, which have been described as optimal compared to not weighted PRSs [161,162]. The normal distribution of the two PRSs in the prediabetic cohort was confirmed, and shown in two graphics with their mean and standard deviation, as seen in **figure 3**. It is also demonstrated that combining individual variants into a PRS can provide more information about T2D vulnerability patterns. Prediabetes and obesity are global epidemics with rapidly increasing mortality and morbidity. Obesity is a significant factor in the development of T2D; thus, there is a close relationship between them [47].

Multiple linear regression with adjustment for age and sex was done to find the associations between T2D PRS and baseline metabolic parameters (**Table 11**). After adjustment for multiple testing, associations were found between T2D PRS and BMI, FFM, FM, Fasting glucose and 2-hour glucose. The association found with BMI matches the findings of a study in a European cohort. Applying logistic regression to calculate odds and hazard ratios, the predictive effect of BMI for T2D incidence was found[163]. Similar results have been found when analyzing a study cohort from the United States[164]. A direct association between FFM and the risk of developing T2D hasn't been described; nevertheless, the association between FFM and T2D remission in males was described at the begging of the year[165]. A research in 2021 described a FM association with the risk of developing T2D in childhood, concluding that an increase in childhood FM was more strongly associated with increased adult T2D risk than an increase in weight, independent of childhood height[166]. Overall, the correlation of prediabetes and T2D with adiposity in adults is a hot topic; the most up-to-date research indicated that compared with people without diabetes, adults with prediabetes and T2D had significantly higher percentages of total fat. Furthermore, as the disease progresses, fat mass decreases in T2D patients [167]. As described before, the values of fasting glucose were in the range to be described as a prediabetic cohort; thus, their association with the risk of developing T2D was expected. It has been shown that the results included in **table 11** are in accordance with other studies published; nevertheless, their replication and validation on European cohorts may be needed. Multiple linear regression with adjustment for age and sex was done to find the associations between obesity PRS and baseline metabolic parameters (**Table 12**). After adjustment for multiple testing, the only association was between the obesity PRS and BMI. The relationship between BMI and obesity has been widely researched, and the values of BMI described in table 7 are in accordance with values that classify our cohort as overweight [48,168]; therefore, the association was expected.

The associations of a T2D PRS and changes over time in metabolic parameters in the prediabetic cohort of Polish Caucasians were done. Individuals with high T2D scores showed increased Δ FM. With each SD increase of PRS, Δ FM increased by 0.0049 kg (**Table 13**). The results are consistent with previously reported data [169,170]. In 2021, an association between weight change and FM change with prediabetes was found in African Americans and European Americans [171]. Using linear regression models, they discovered that Δ FM was a significant predictor of progression in the prediabetes cohort. In addition, a study on the association of prediabetes and T2D with adipose tissue in adults was reported in 2022.

Based on multivariate linear regression models, the researchers found an association between adults with prediabetes and increased FM[167]. In contrast to the results, a recently published study found associations of Δ FFM with risk for T2D in a Hispanic cohort [172]. Another study that focus its research on the genetic variations in the gene *KCNJ11*, included in the T2D PRS. *KCNJ11* has been associated with prediabetes in an Asian population[173]; the research of it in an extensive study will be of great interest to understand its disease risk predictor power.

For the second PRS, high obesity PRS was associated with Δ FM and Δ 2-hour glucose. For each SD increase in PRS, ΔFM and Δ 2-hour glucose increased by 0.0056 kg and 0.0013 mg/dL, respectively (**Table 14**). Obesity has become a significant problem due the increased number of patients and some metabolic complications. The *FTO* gene was the first gene identified to link FM and obesity in humans[174,175]. Several MC4R gene variants have been associated with FM, weight, and obesity risk [176,177]. Of the 21 SNPs included in our study, 7 had loci on these genes; therefore, the associations shown in this study are consistent with previous publications. In terms of 120-minute blood glucose changes, 2-hour postprandial blood glucose levels are associated with the development of metabolic diseases such as obesity, T2D, and cardiovascular disease[178] in patients with T2D and prediabetes [179]. Both of our PRSs were associated with changes in FM, which may be related to the 20 SNP overlap between T2D and obesity scores.

3. Genotypes frequencies

It was observed that *MTCH2* rs10838738 was significantly associated with metabolic parameters in the study group (**Table 15**). From the data, it was found that carriers of the rs10838738 homozygous AA genotype exhibited more significant changes at 2-hour glucose. In contrast, the homozygous GG genotype exhibited significantly smaller changes at 2-hour insulin (**Figure 4**). In contrast to the results, a previous study showed that the homozygous genotype for the risk allele G was not significantly associated with 120-minute blood glucose change but with higher BMI [180]. The SNP *MTCH2* (rs10838738) has been reported to affect gene expression in visceral adipose tissue [181]. However, as shown in our study and some previous studies[182], there is little evidence that this SNP is significantly associated with diabetes. Extensive research is required to discover and demonstrate the impact of this SNP on T2D evolution. We observed a trend for *PRRC2A* rs2260000 according to the metabolic parameters of the study group (**Table 15**). The alternative allele GG genotype

homozygous carriers slightly varied in FM and VF (**Figure 5**). No data have been published on the association between *PRRC2A* rs2260000 and T2D. Nonetheless, recent studies have linked variants in the PRRC2A gene with obesity[183] and T1D risk [184]. Furthermore, consistent with our results, current data show a relationship between this gene and human adipocytes isolated from VF[185,186]. Since 2020, articles researching the function of *PRRC2A* have been published; therefore, trends in research can be observed. The next trend observed was on the gene *ETV5* rs7647305 according to the metabolic parameters of the study group (**Table 15**). We found that homozygous carriers of the alternative allele CC genotype had very large changes in 2-hour glucose and very small changes in IPAQ (**Figure 6**). The association between this gene and T2D has been described previously. One study found an association between *ETV5* and hypertension [187], and the researchers noted that SNPs that predict the development of hypertension could also predict T2D According to linear model analysis, a recent study found that the *ETV5* affects B cell dysfunction and pathophysiology in T2D [188]. The association of this SNP with metabolic parameters has rarely been studied. A recent work revealed a critical role for *ETV5* in regulating insulin secretion[189], while another study highlighted the importance of studying this gene to compare sedentary behavior with physical activity[190]. The last trend observed was for *KCTD15* rs29941 according to the metabolic parameters of the study group (**Table 15**). From the data, we found that carriers of the AG genotype exhibited significant glucose changes at 120 (**Figure 7**). The *KCTD15* gene has previously been associated with the risk of obesity [191] and T2D[192,193]. In particular, *KCTD15* rs29941 was significantly associated with fasting blood glucose [191] [181] and the risk of insulin resistance [192].

Among the strengths of the Ph.D. work dissertation is the fact that it was one of the first studies to research the associations of the changes in several metabolic parameters with T2D and obesity PRSs in Polish population. Additionally, its strengths are that it was based on a relatively large population. Although many risk factors for T2D have already been identified, early markers of transitioning from normal to prediabetes aren't yet identified[179]. This study could help explain why this is the case. Associations between PRSs and changes in metabolic traits related to T2D show how the genetic information of patients with prediabetes can be used to prevent the disease. The work also shows how the clinical environment can use the data of patients with prediabetes to prevent complications of metabolic syndrome, such as heart disease, obesity, and hypertension.

Nevertheless, there are also limitations in the study. One is the lack of long-term human studies to analyze the change in two-hour postprandial glucose and its association with the different genotypes described in our research. Additionally, a better understanding of the lack of association between other Δ metabolic parameters and PRSs needs to be researched. Only Caucasian participants were included in the study; further consideration of different ethnicities is necessary. The effect size estimates used to create the PRSs were based on data from European ancestry. This study didn't consider the possibility of different effect sizes for diverse populations. As a result, the data shouldn't be trusted without being replicated in additional multi-ethnic populations.

VIII. Conclusions

1. Out of the 14 studies identified in the systematic review that developed PRSs, 11 were used to differentiate patients from controls, and three were used to discriminate between T1D and diabetes subtypes. These PRSs were assessed for accuracy using the AUC metric regardless of the data source, the panel of genes used, and genotyping strategies. To better predict diabetes, the use of PRS that combines clinical, environmental, and genetic interactions must be used. Creating a pipeline that translates findings into actual evidence is the first step in demonstrating PRS's clinical validity.

2. Two PRSs were created, T2D PRS was made from 68 SNPs, and the obesity PRS included 21 SNPs. There is an overlap between genes implicated in the risk of developing T2D and those associated with the risk of obesity. The Δ FM is associated with T2D and obesity PRSs in a prediabetic cohort. The Δ glucose at 120 min is associated with obesity's PRS. The findings are consistent with recent results demonstrating that an increase in the change of FM and obesity are closely related to insulin resistance and abnormalities in glucose metabolism and, therefore, T2D risk[194,195].

3. The AA genotype carriers of the gene *MTCH2* (rs10838738*)* were significantly higher in Δ 2-hour glucose and Δ 2-hour insulin. The results may have practical clinical implications if confirmed in larger populations and among different ethnic groups. The associations found in this project could be considered a pilot study for producing a powerful tool for identifying individuals with an increased risk of complications at diagnosis.

IX. Bibliography

1. Zimmet, P.Z.; Magliano, D.J.; Herman, W.H.; Shaw, J.E. Diabetes: A 21st Century Challenge. *Lancet Diabetes Endocrinol.* **2014**, *2*, 56–64, doi:10.1016/s2213-8587(13)70112-8.
2. WHO Retrieved from Https://Www.Who.Int/News-Room/Fact-Sheets/Detail/Diabetes.
3. Home; Resources; diabetes, L. with; Acknowledgement; FAQs; Contact; Policy, P. IDF Diabetes Atlas 2021 | IDF Diabetes Atlas.
4. Jeffrey A. Bluestone, G.E., Kevan Herold Genetics, Pathogenesis and Clinical Interventions in Type\hspace0.167em1 Diabetes. *Nature* **2010**, *464*, 1293–1300, doi:10.1038/nature08933.
5. Kahn, et al., Steven E. Pathophysiology and Treatment of Type 2 Diabetes: Perspectives on the Past, Present, and Future. *The Lancet* **2014**, *383*, 1068–1083, doi:10.1016/s0140-6736(13)62154-6.
6. Boyle, J.P.; Thompson, T.J.; Gregg, E.W.; Barker, L.E.; Williamson, D.F. Projection of the Year 2050 Burden of Diabetes in the US Adult Population: Dynamic Modeling of Incidence, Mortality, and Prediabetes Prevalence. *Popul. Health Metr.* **2010**, *8*, doi:10.1186/1478-7954-8-29.
7. Whiting, D.R.; Guariguata, L.; Weil, C.; Shaw, J. IDF Diabetes Atlas: Global Estimates of the Prevalence of Diabetes for 2011 and 2030. *Diabetes Res. Clin. Pract.* **2011**, *94*, 311–321, doi:10.1016/j.diabres.2011.10.029.
8. Owen, K. Maturity-Onset Diabetes of the Young: From Clinical Description to Molecular Genetic Characterization. *Best Pract. Res. Clin. Endocrinol. Metab.* **2001**, *15*, 309–323, doi:10.1053/beem.2001.0148.
9. Gloyn, S.E., A.L.; Shepherd, M.; Howell, R.T.; Parry, E.M.; Jefferson, A.; Levy, E.R.; Hattersley, A.T. Maturity-Onset Diabetes of the Young Caused by a Balanced Translocation Where the 20q12 Break Point Results in Disruption Upstream of the Coding Region of Hepatocyte Nuclear Factor-4 (HNF4A) Gene. *Diabetes* **2002**, *51*, 2329–2333, doi:10.2337/diabetes.51.7.2329.
10. Stride, A.H.A. Different Genes, Different Diabetes: Lessons from Maturity Onset Diabetes of the Young. . Ann Med ,. *Ann Med* **2002**, *34*, 207-16.

11. DiMeglio, L.A.; Evans-Molina, C.; Oram, R.A. Type 1 Diabetes. *Lancet Lond. Engl.* **2018**, *391*, 2449–2462, doi:10.1016/S0140-6736(18)31320-5.
12. McLaughlin, K.A.; Richardson, C.C.; Ravishankar, A.; Brigatti, C.; Liberati, D.; Lampasona, V.; Piemonti, L.; Morgan, D.; Feltbower, R.G.; Christie, M.R. Identification of Tetraspanin-7 as a Target of Autoantibodies in Type 1 Diabetes. *Diabetes* **2016**, *65*, 1690–1698, doi:10.2337/db15-1058.
13. Sharp; Weedon, M.N.; Hagopian, W.A.; Oram, R.A. Clinical and Research Uses of Genetic Risk Scores in Type 1 Diabetes. *Curr. Opin. Genet. Dev.* **2018**, *50*, 96–102, doi:10.1016/j.gde.2018.03.009.
14. Redondo, M.J.; Jeffrey, J.; Fain, P.R.; Eisenbarth, G.S.; Orban, T. Concordance for Islet Autoimmunity among Monozygotic Twins. *N. Engl. J. Med.* **2008**, *359*, 2849–2850, doi:10.1056/nejmc0805398.
15. Kuo, C.-F.; Chou, I.-J.; Grainge, M.; Luo, S.-F.; See, L.-C.; Yu, K.-H.; Zhang, W.; Doherty, M.; Valdes, A. Familial Aggregation and Heritability of Type 1 Diabetes Mellitus and Coaggregation of Chronic Diseases in Affected Families. *Clin. Epidemiol.* **2018**, *Volume 10*, 1447–1455, doi:10.2147/clep.s172207.
16. Gale, E.A.M. Type 1 Diabetes Mellitus (Revision Number 38). In *Diapedia*; Diapedia.org, 2014.
17. Barrett, J.C.; Clayton, and D.G.; Concannon, P.; Akolkar, B.; Cooper, J.D.; Erlich, H.A.; Julier, C.; Morahan, G.; Nerup, J.; Nierras, C.; et al. Genome-Wide Association Study and Meta-Analysis Find That over 40 Loci Affect Risk of Type 1 Diabetes. *Nat. Genet.* **2009**, *41*, 703–707, doi:10.1038/ng.381.
18. Pociot, F. Type 1 Diabetes Genome-Wide Association Studies: Not to Be Lost in Translation. *Clin. Transl. Immunol.* **2017**, *6*, e162, doi:10.1038/cti.2017.51.
19. Galaviz, K.I.; Narayan, K.M.V.; Lobelo, F.; Weber, M.B. Lifestyle and the Prevention of Type 2 Diabetes: A Status Report. *Am. J. Lifestyle Med.* **2018**, *12*, 4–20, doi:10.1177/1559827615619159.
20. King, P.; Peacock, I.; Donnelly, R. The UK Prospective Diabetes Study (UKPDS): Clinical and Therapeutic Implications for Type 2 Diabetes. *Br. J. Clin. Pharmacol.* **1999**, *48*, 643–648, doi:10.1046/j.1365-2125.1999.00092.x.

21. Gregg, E.W.; Li, Y.; Wang, J.; Burrows, N.R.; Ali, M.K.; Rolka, D.; Williams, D.E.; Geiss, L. Changes in Diabetes-Related Complications in the United States, 1990-2010. *N. Engl. J. Med.* **2014**, *370*, 1514–1523, doi:10.1056/NEJMoa1310799.
22. Murea, M.; Ma, L.; Freedman, B.I. Genetic and Environmental Factors Associated with Type 2 Diabetes and Diabetic Vascular Complications. *Rev. Diabet. Stud.* **2012**, *9*, 6–22, doi:10.1900/rds.2012.9.6.
23. Morris, A. Large-Scale Association Analysis Provides Insights into the Genetic Architecture and Pathophysiology of Type 2 Diabetes. *Nat. Genet.* **2012**, *44*, 981–990, doi:10.1038/ng.2383.
24. Fuchsberger, C.; Flannick, J.; Teslovich, T.M.; Mahajan, A.; Agarwala, V.; Gaulton, K.J.; Ma, C.; Fontanillas, P.; Moutsianas, L.; McCarthy, D.J.; et al. The Genetic Architecture of Type 2 Diabetes. *Nature* **2016**, *536*, 41–47, doi:10.1038/nature18642.
25. Scott, R.A.; Scott, L.J.; Mägi, R.; Marullo, L.; Gaulton, K.J.; Kaakinen, M.; Pervjakova, N.; Pers, T.H.; Johnson, A.D.; Eicher, J.D.; et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* **2017**, *66*, 2888–2902, doi:10.2337/db16-1253.
26. Wray, N.; Yang, J.; Hayes, B.J.; Price, A.L.; Goddard, M.E.; Visscher, P.M. Pitfalls of Predicting Complex Traits from SNPs. *Nat. Rev. Genet.* **2013**, *14*, 507–515, doi:10.1038/nrg3457.
27. Lyssenko, V.; Laakso, M. Genetic Screening for the Risk of Type 2 Diabetes: Worthless or Valuable? *Diabetes Care* **2013**, *36*, S120–S126, doi:10.2337/dcs13-2009.
28. Wu, Y.; Ding, Y.; Tanaka, Y.; Zhang, W. Risk Factors Contributing to Type 2 Diabetes and Recent Advances in the Treatment and Prevention. *Int. J. Med. Sci.* **2014**, *11*, 1185–1200, doi:10.7150/ijms.10001.
29. American Diabetes Association 2. Classification and Diagnosis of Diabetes: *Standards of Medical Care in Diabetes—2020*. *Diabetes Care* **2020**, *43*, S14–S31, doi:10.2337/dc20-S002.
30. Faerch, K.; Hulman, A.; P.J. Solomon, T. Heterogeneity of Pre-Diabetes and Type 2 Diabetes: Implications for Prediction, Prevention and Treatment

Responsiveness. *Curr. Diabetes Rev.* **2015**, *12*, 30–41, doi:10.2174/1573399811666150416122903.

31. Ilonen, J.; Lempainen, J.; Veijola, R. The Heterogeneous Pathogenesis of Type 1 Diabetes Mellitus. *Nat. Rev. Endocrinol.* **2019**, *15*, 635–650, doi:10.1038/s41574-019-0254-y.
32. Seidu, S.; Davies, M.J.; Mostafa, S.; Lusignan, S. de; Khunti, K. Prevalence and Characteristics in Coding, Classification and Diagnosis of Diabetes in Primary Care. *Postgrad. Med. J.* **2013**, *90*, 13–17, doi:10.1136/postgradmedj-2013-132068.
33. Stone, M.A.; Camosso-Stefinovic, J.; Wilkinson, J.; Lusignan, S.D.; Hattersley, A.T.; Khunti, K. Incorrect and Incomplete Coding and Classification of Diabetes: A Systematic Review. *Diabet. Med.* **2009**, *27*, 491–497, doi:10.1111/j.1464-5491.2009.02920.x.
34. Bansal, N. Prediabetes Diagnosis and Treatment: A Review. *World J. Diabetes* **2015**, *6*, 296, doi:10.4239/wjd.v6.i2.296.
35. Khan, M.A.B.; Hashim, M.J.; King, J.K.; Govender, R.D.; Mustafa, H.; Al Kaabi, J. Epidemiology of Type 2 Diabetes - Global Burden of Disease and Forecasted Trends. *J. Epidemiol. Glob. Health* **2020**, *10*, 107–111, doi:10.2991/jegh.k.191028.001.
36. Hostalek, U. Global Epidemiology of Prediabetes - Present and Future Perspectives. *Clin. Diabetes Endocrinol.* **2019**, *5*, 5, doi:10.1186/s40842-019-0080-0.
37. Souza, C.F. de; Gross, J.L.; Gerchman, F.; Leitão, C.B. [Prediabetes: diagnosis, evaluation of chronic complications, and treatment]. *Arq. Bras. Endocrinol. Metabol.* **2012**, *56*, 275–284, doi:10.1590/s0004-27302012000500001.
38. Huang, Y.; Cai, X.; Mai, W.; Li, M.; Hu, Y. Association between Prediabetes and Risk of Cardiovascular Disease and All Cause Mortality: Systematic Review and Meta-Analysis. *BMJ* **2016**, *355*, i5953, doi:10.1136/bmj.i5953.
39. Khaodhiar, L.; Cummings, S.; Apovian, C.M. Treating Diabetes and Prediabetes by Focusing on Obesity Management. *Curr. Diab. Rep.* **2009**, *9*, 348–354.

40. Public Health England's Health Profiles 2014 Published Available online: https://www.gov.uk/government/news/public-health-englands-health-profiles-2014-published (accessed on 15 August 2022).
41. Eckel, R.H.; Alberti, K.; Grundy, S.M.; Zimmet, P.Z. The Metabolic Syndrome. *The Lancet* **2010**, *375*, 181–183, doi:10.1016/S0140-6736(09)61794-3.
42. Zand, A. Prediabetes: Why Should We Care? *Mech. ACTION* **2018**, 9.
43. The Web's Free 2022 ICD-10-CM/PCS Medical Coding Reference Available online: https://www.icd10data.com/ (accessed on 21 July 2022).
44. Sherwood, Z. Prediabetes: Definition, Diagnostic Criteria and Management. *22*, 4.
45. Diagnosis | ADA Available online: https://diabetes.org/diabetes/a1c/diagnosis (accessed on 15 August 2022).
46. Overview | Type 2 Diabetes: Prevention in People at High Risk | Guidance | NICE Available online: https://www.nice.org.uk/guidance/PH38/ (accessed on 15 August 2022).
47. Sheikhpour, M.; Abolfathi, H.; Khatami, S.; Meshkani, R.; Barghi, T.S. The Interaction between Gene Profile and Obesity in Type 2 Diabetes: A Review. *Obes. Med.* **2020**, *18*, 100197, doi:10.1016/j.obmed.2020.100197.
48. World Obesity Atlas 2022 Available online: https://www.worldobesity.org/resources/resource-library/world-obesity-atlas-2022 (accessed on 15 August 2022).
49. Poirier, P.; Giles, T.D.; Bray, G.A.; Hong, Y.; Stern, J.S.; Pi-Sunyer, F.X.; Eckel, R.H. Obesity and Cardiovascular Disease: Pathophysiology, Evaluation, and Effect of Weight Loss. *Circulation* **2006**, *113*, 898–918, doi:10.1161/CIRCULATIONAHA.106.171016.
50. Neeland, I.J.; Turer, A.T.; Ayers, C.R.; Powell-Wiley, T.M.; Vega, G.L.; Farzaneh-Far, R.; Grundy, S.M.; Khera, A.; McGuire, D.K.; de Lemos, J.A. Dysfunctional Adiposity and the Risk of Prediabetes and Type 2 Diabetes in Obese Adults. *JAMA* **2012**, *308*, 1150–1159, doi:10.1001/2012.jama.11132.
51. Goran, M.I.; Lane, C.; Toledo-Corral, C.; Weigensberg, M.J. Persistence of Pre-Diabetes in Overweight and Obese Hispanic Children : Association With

Progressive Insulin Resistance, Poor β-Cell Function, and Increasing Visceral Fat. *Diabetes* **2008**, *57*, 3007–3012, doi:10.2337/db08-0445.

52. Martinez, K.E.; Tucker, L.A.; Bailey, B.W.; LeCheminant, J.D. Expanded Normal Weight Obesity and Insulin Resistance in US Adults of the National Health and Nutrition Examination Survey. *J. Diabetes Res.* **2017**, *2017*, 1–8, doi:10.1155/2017/9502643.
53. Miao, Z.; Alvarez, M.; Ko, A.; Bhagat, Y.; Rahmani, E.; Jew, B.; Heinonen, S.; Muñoz-Hernandez, L.L.; Herrera-Hernandez, M.; Aguilar-Salinas, C.; et al. The Causal Effect of Obesity on Prediabetes and Insulin Resistance Reveals the Important Role of Adipose Tissue in Insulin Resistance. *PLOS Genet.* **2020**, *16*, e1009018, doi:10.1371/journal.pgen.1009018.
54. Kasuga, M. [Genetic factor for diabetes and obesity]. *Nihon Rinsho Jpn. J. Clin. Med.* **2010**, *68 Suppl 8*, 359–363.
55. Lin, X.; Li, H. Obesity: Epidemiology, Pathophysiology, and Therapeutics. *Front. Endocrinol.* **2021**, *12*, 706978, doi:10.3389/fendo.2021.706978.
56. Srinivasan, S.; Chen, L.; Todd, J.; Divers, J.; Gidding, S.; Chernausek, S.; Gubitosi-Klug, R.A.; Kelsey, M.M.; Shah, R.; Black, M.H.; et al. The First Genome-Wide Association Study for Type 2 Diabetes in Youth: The Progress in Diabetes Genetics in Youth (ProDiGY) Consortium. *Diabetes* **2021**, *70*, 996–1005, doi:10.2337/db20-0443.
57. Chen, J.; Sun, M.; Adeyemo, A.; Pirie, F.; Carstensen, T.; Pomilla, C.; Doumatey, A.P.; Chen, G.; Young, E.H.; Sandhu, M.; et al. Genome-Wide Association Study of Type 2 Diabetes in Africa. *Diabetologia* **2019**, *62*, 1204–1211, doi:10.1007/s00125-019-4880-7.
58. Bogardus, C. Missing Heritability and GWAS Utility. *Obes. Silver Spring Md* **2009**, *17*, 209–210, doi:10.1038/oby.2008.613.
59. Wang, T.; Xu, M.; Bi, Y.; Ning, G. Interplay between Diet and Genetic Susceptibility in Obesity and Related Traits. *Front. Med.* **2018**, *12*, 601–607, doi:10.1007/s11684-018-0648-6.
60. Vega, F.M.D.L.; Bustamante, C.D. Polygenic Risk Scores: A Biased Prediction? *Genome Med.* **2018**, *10*, doi:10.1186/s13073-018-0610-x.

61. Josephs, K.S.; Berner, A.; George, A.; Scott, R.H.; Firth, H.V.; Tatton-Brown, K. Genomics: The Power, Potential and Pitfalls of the New Technologies and How They Are Transforming Healthcare. *Clin. Med.* **2019**, *19*, 269–272, doi:10.7861/clinmedicine.19-4-269.
62. Khera, A.V.; Chaffin, M.; Aragam, K.G.; Haas, M.E.; Roselli, C.; Choi, S.H.; Natarajan, P.; Lander, E.S.; Lubitz, S.A.; Ellinor, P.T.; et al. Genome-Wide Polygenic Scores for Common Diseases Identify Individuals with Risk Equivalent to Monogenic Mutations. *Nat. Genet.* **2018**, *50*, 1219–1224, doi:10.1038/s41588-018-0183-z.
63. Uffelmann, E.; Huang, Q.Q.; Munung, N.S.; de Vries, J.; Okada, Y.; Martin, A.R.; Martin, H.C.; Lappalainen, T.; Posthuma, D. Genome-Wide Association Studies. *Nat. Rev. Methods Primer* **2021**, *1*, 1–21, doi:10.1038/s43586-021-00056-9.
64. Bycroft, C.; Freeman, C.; Petkova, D.; Band, G.; Elliott, L.T.; Sharp, K.; Motyer, A.; Vukcevic, D.; Delaneau, O.; O'Connell, J.; et al. The UK Biobank Resource with Deep Phenotyping and Genomic Data. *Nature* **2018**, *562*, 203–209, doi:10.1038/s41586-018-0579-z.
65. Evangelou, E.; Warren, and H.R.; Mosen-Ansorena, D.; Mifsud, B.; Pazoki, R.; Gao, H.; Ntritsos, G.; Dimou, N.; Cabrera, C.P.; Karaman, I.; et al. Genetic Analysis of over 1 Million People Identifies 535 New Loci Associated with Blood Pressure Traits. *Nat. Genet.* **2018**, *50*, 1412–1425, doi:10.1038/s41588-018-0205-x.
66. Zeng, P. Statistical Analysis for Genome-Wide Association Study. *J. Biomed. Res.* **2015**, doi:10.7555/jbr.29.20140007.
67. Golan, D.; Lander, E.S.; Rosset, S. Measuring Missing Heritability: Inferring the Contribution of Common Variants. *Proc. Natl. Acad. Sci.* **2014**, *111*, E5272–E5281, doi:10.1073/pnas.1419064111.
68. Lyssenko, V.; Bianchi, C.; Del Prato, S. Personalized Therapy by Phenotype and Genotype. *Diabetes Care* **2016**, *39*, S127–S136, doi:10.2337/dcS15-3002.
69. Ahlqvist, E.; Storm, P.; Käräjämäki, A.; Martinell, M.; Dorkhan, M.; Carlsson, A.; Vikman, P.; Prasad, R.B.; Aly, D.M.; Almgren, P.; et al. Novel Subgroups of Adult-Onset Diabetes and Their Association with Outcomes: A Data-Driven

Cluster Analysis of Six Variables. *Lancet Diabetes Endocrinol.* **2018**, *6*, 361–369, doi:10.1016/S2213-8587(18)30051-2.

70. Yang, J.; Benyamin, B.; McEvoy, B.P.; Gordon, S.; Henders, A.K.; Nyholt, D.R.; Madden, P.A.; Heath, A.C.; Martin, N.G.; Montgomery, G.W.; et al. Common SNPs Explain a Large Proportion of the Heritability for Human Height. *Nat. Genet.* **2010**, *42*, 565–569, doi:10.1038/ng.608.
71. Bailey, J.N.C.; Igo, R.P. Genetic Risk Scores. *Curr. Protoc. Hum. Genet.* **2016**, *91*, 1291–1299, doi:10.1002/cphg.20.
72. Lewis, C.M.; Vassos, E. Polygenic Risk Scores: From Research Tools to Clinical Instruments. *Genome Med.* **2020**, *12*, 44, doi:10.1186/s13073-020-00742-5.
73. Choi, S.W.; Mak, T.S.-H.; O'Reilly, P.F. Tutorial: A Guide to Performing Polygenic Risk Score Analyses. *Nat. Protoc.* **2020**, *15*, 2759–2772, doi:10.1038/s41596-020-0353-1.
74. Lall, K.; Mägi, R.; Morris, A.; Metspalu, A.; Fischer, K. Personalized Risk Prediction for Type 2 Diabetes: The Potential of Genetic Risk Scores. *Genet. Med.* **2016**, *19*, 322–329, doi:10.1038/gim.2016.103.
75. Kolb, H.; Martin, S. Environmental/Lifestyle Factors in the Pathogenesis and Prevention of Type 2 Diabetes. *BMC Med.* **2017**, *15*, doi:10.1186/s12916-017-0901-x.
76. Sanna, S.; Zuydam, N.R. van; Mahajan, A.; Kurilshikov, A.; Vila, A.V.; Võsa, U.; Mujagic, Z.; Masclee, A.A.M.; Jonkers, D.M.A.E.; Oosting, M.; et al. Causal Relationships among the Gut Microbiome, Short-Chain Fatty Acids and Metabolic Diseases. *Nat. Genet.* **2019**, *51*, 600–605, doi:10.1038/s41588-019-0350-x.
77. Mahajan, A.; Taliun, D.; Thurner, M.; Robertson, N.R.; Torres, J.M.; Rayner, N.W.; Payne, A.J.; Steinthorsdottir, V.; Scott, R.A.; Grarup, N.; et al. Fine-Mapping Type 2 Diabetes Loci to Single-Variant Resolution Using High-Density Imputation and Islet-Specific Epigenome Maps. *Nat. Genet.* **2018**, *50*, 1505–1513, doi:10.1038/s41588-018-0241-6.
78. Onengut-Gumuscu, S.; Chen, and W.-M.; Burren, O.; Cooper, N.J.; Quinlan, A.R.; Mychaleckyj, J.C.; Farber, E.; Bonnie, J.K.; Szpak, M.; Schofield, E.; et

al. Fine Mapping of Type 1 Diabetes Susceptibility Loci and Evidence for Colocalization of Causal Variants with Lymphoid Gene Enhancers. *Nat. Genet.* **2015**, *47*, 381–386, doi:10.1038/ng.3245.

79. Wareham, N.J.; Herman, W.H. The Clinical and Public Health Challenges of Diabetes Prevention: A Search for Sustainable Solutions. *PLOS Med.* **2016**, *13*, e1002097, doi:10.1371/journal.pmed.1002097.
80. Poulsen, P.; Kyvik, K.O.; Vaag, A.; Beck-Nielsen, H. Heritability of Type II (Non-Insulin-Dependent) Diabetes Mellitus and Abnormal Glucose Tolerance - a Population-Based Twin Study. *Diabetologia* **1999**, *42*, 139–145, doi:10.1007/s001250051131.
81. Tuomilehto, J.; Lindström, J.; Eriksson, J.G.; Valle, T.T.; Hämäläinen, H.; Ilanne-Parikka, P.; Keinänen-Kiukaanniemi, S.; Laakso, M.; Louheranta, A.; Rastas, M.; et al. Prevention of Type 2 Diabetes Mellitus by Changes in Lifestyle among Subjects with Impaired Glucose Tolerance. *N. Engl. J. Med.* **2001**, *344*, 1343–1350, doi:10.1056/nejm200105033441801.
82. Almgren, P. Heritability and Familiality of Type 2 Diabetes and Related Quantitative Traits in the Botnia Study. *Diabetologia* **2011**, *54*, 2811–2819, doi:10.1007/s00125-011-2267-5.
83. Lall, K.; Mägi, R.; Morris, A.; Metspalu, A.; Fischer, K. Personalized Risk Prediction for Type 2 Diabetes: The Potential of Genetic Risk Scores. *Genet. Med.* **2016**, *19*, 322–329, doi:10.1038/gim.2016.103.
84. Meigs, J.B.; Shrader, P.; Sullivan, L.M.; McAteer, J.B.; Fox, C.S.; Dupuis, J.; Manning, A.K.; Florez, J.C.; Wilson, P.W.F.; D\textquotesingleAgostino, R.B.; et al. Genotype Score in Addition to Common Risk Factors for Prediction of Type 2 Diabetes. *N. Engl. J. Med.* **2008**, *359*, 2208–2219, doi:10.1056/nejmoa0804742.
85. Chatterjee, N.; Wheeler, B.; Sampson, J.; Hartge, P.; Chanock, S.J.; Park, J.-H. Projecting the Performance of Risk Prediction Based on Polygenic Analyses of Genome-Wide Association Studies. *Nat. Genet.* **2013**, *45*, 400–405, doi:10.1038/ng.2579.
86. Vassy, J.L.; Hivert, M.-F.; Porneala, B.; Dauriz, M.; Florez, J.C.; Dupuis, J.; Siscovick, D.S.; Fornage, M.; Rasmussen-Torvik, L.J.; Bouchard, C.; et al.

Polygenic Type 2 Diabetes Prediction at the Limit of Common Variant Detection. *Diabetes* **2014**, *63*, 2172–2182, doi:10.2337/db13-1663.

87. Szczerbinski, L. Polish Registry of Diabetes (PolReD). *Identifier NCT04657367* **2020**, doi:https://clinicaltrials.gov/ct2/show/study/NCT04657367.
88. Sidorkiewicz, I.; Niemira, M.; Maliszewska, K.; Erol, A.; Bielska, A.; Szalkowska, A.; Adamska-Patruno, E.; Szczerbinski, L.; Gorska, M.; Kretowski, A. Circulating MiRNAs as a Predictive Biomarker of the Progression from Prediabetes to Diabetes: Outcomes of a 5-Year Prospective Observational Study. *J. Clin. Med.* **2020**, *9*, 2184, doi:10.3390/jcm9072184.
89. American Diabetes Association. *Access 20 August* **2021**, doi:https://www.diabetes.org/a1c/diagnosis.
90. Type 2 Diabetes Knowledge Portal - Home Available online: https://t2d.hugeamp.org/ (accessed on 21 June 2022).
91. R Core Team R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput.* **2021**, Vienna, Austria, doi:https://www.R-project.org/.
92. Padilla-Martínez, F.; Collin, F.; Kwasniewski, M.; Kretowski, A. Systematic Review of Polygenic Risk Scores for Type 1 and Type 2 Diabetes. *Int. J. Mol. Sci.* **2020**, *21*, 1703, doi:10.3390/ijms21051703.
93. Oram, R.A.; Patel, K.; Hill, A.; Shields, B.; McDonald, T.J.; Jones, A.; Hattersley, A.T.; Weedon, M.N. A Type 1 Diabetes Genetic Risk Score Can Aid Discrimination Between Type 1 and Type 2 Diabetes in Young Adults. *Diabetes Care* **2015**, *39*, 337–344, doi:10.2337/dc15-1111.
94. Patel, K.A.; Oram, R.A.; Flanagan, S.E.; Franco, E.D.; Colclough, K.; Shepherd, M.; Ellard, S.; Weedon, M.N.; Hattersley, A.T. Type 1 Diabetes Genetic Risk Score: A Novel Tool to Discriminate Monogenic and Type 1 Diabetes. *Diabetes* **2016**, *65*, 2094–2099, doi:10.2337/db15-1690.
95. Winkler, C.; Krumsiek, J.; Buettner, F.; Angermüller, C.; Giannopoulou, E.Z.; Theis, F.J.; Ziegler, A.-G.; Bonifacio, E. Feature Ranking of Type 1 Diabetes Susceptibility Genes Improves Prediction of Type 1 Diabetes. *Diabetologia* **2014**, *57*, 2521–2529, doi:10.1007/s00125-014-3362-1.

96. Perry, D.J.; Wasserfall, C.H.; Oram, R.A.; Williams, M.D.; Posgai, A.; Muir, A.B.; Haller, M.J.; Schatz, D.A.; Wallet, M.A.; Mathews, C.E.; et al. Application of a Genetic Risk Score to Racially Diverse Type 1 Diabetes Populations Demonstrates the Need for Diversity in Risk-Modeling. *Sci. Rep.* **2018**, *8*, doi:10.1038/s41598-018-22574-5.

97. Sharp; Rich, S.S.; Wood, A.R.; Jones, S.E.; Beaumont, R.N.; Harrison, J.W.; Schneider, D.A.; Locke, J.M.; Tyrrell, J.; Weedon, M.N.; et al. Development and Standardization of an Improved Type 1 Diabetes Genetic Risk Score for Use in Newborn Screening and Incident Diagnosis. *Diabetes Care* **2019**, *42*, 200–207, doi:10.2337/dc18-1785.

98. Yaghootkar, H.; Abbasi, F.; Ghaemi, N.; Rabbani, A.; Wakeling, M.N.; Eshraghi, P.; Enayati, S.; Vakili, S.; Heidari, S.; Patel, K.; et al. Type 1 Diabetes Genetic Risk Score Discriminates between Monogenic and Type 1 Diabetes in Children Diagnosed at the Age of Less of 5 Years in the Iranian Population. *Diabet. Med.* **2019**, doi:10.1111/dme.14071.

99. Khera, A.V.; Chaffin, M.; Aragam, K.G.; Haas, M.E.; Roselli, C.; Choi, S.H.; Natarajan, P.; Lander, E.S.; Lubitz, S.A.; Ellinor, P.T.; et al. Genome-Wide Polygenic Scores for Common Diseases Identify Individuals with Risk Equivalent to Monogenic Mutations. *Nat. Genet.* **2018**, *50*, 1219–1224, doi:10.1038/s41588-018-0183-z.

100. Weedon, M.N.; McCarthy, M.I.; Hitman, G.; Walker, M.; Groves, C.J.; Zeggini, E.; Rayner, N.W.; Shields, B.; Owen, K.R.; Hattersley, A.T.; et al. Combining Information from Common Type 2 Diabetes Risk Polymorphisms Improves Disease Prediction. *PLoS Med.* **2006**, *3*, e374, doi:10.1371/journal.pmed.0030374.

101. Lango, H.; Palmer, C.N.A.; Morris, A.D.; Zeggini, E.; Hattersley, A.T.; McCarthy, M.I.; Frayling, T.M.; and, M.N.W. Assessing the Combined Impact of 18 Common Genetic Variants of Modest Effect Sizes on Type 2 Diabetes Risk. *Diabetes* **2008**, *57*, 3129–3135, doi:10.2337/db08-0504.

102. Lyssenko, V.; Jonsson, A.; Almgren, P.; Pulizzi, N.; Isomaa, B.; Tuomi, T.; Berglund, G.; Altshuler, D.; Nilsson, P.; Groop, L. Clinical Risk Factors, DNA

Variants, and the Development of Type 2 Diabetes. *N. Engl. J. Med.* **2008**, *359*, 2220–2232, doi:10.1056/nejmoa0801869.

103. Meigs, J.B.; Shrader, P.; Sullivan, L.M.; McAteer, J.B.; Fox, C.S.; Dupuis, J.; Manning, A.K.; Florez, J.C.; Wilson, P.W.F.; D\textquotesingleAgostino, R.B.; et al. Genotype Score in Addition to Common Risk Factors for Prediction of Type 2 Diabetes. *N. Engl. J. Med.* **2008**, *359*, 2208–2219, doi:10.1056/nejmoa0804742.
104. Chatterjee, N.; Wheeler, B.; Sampson, J.; Hartge, P.; Chanock, S.J.; Park, J.-H. Projecting the Performance of Risk Prediction Based on Polygenic Analyses of Genome-Wide Association Studies. *Nat. Genet.* **2013**, *45*, 400–405, doi:10.1038/ng.2579.
105. Vassy, J.L.; Hivert, M.-F.; Porneala, B.; Dauriz, M.; Florez, J.C.; Dupuis, J.; Siscovick, D.S.; Fornage, M.; Rasmussen-Torvik, L.J.; Bouchard, C.; et al. Polygenic Type 2 Diabetes Prediction at the Limit of Common Variant Detection. *Diabetes* **2014**, *63*, 2172–2182, doi:10.2337/db13-1663.
106. Rich, S.S.; Akolkar, B.; Concannon, P.; Erlich, H.; Hilner, J.E.; Julier, C.; Morahan, G.; Nerup, J.; Nierras, C.; Pociot, F.; et al. Overview of the Type I Diabetes Genetics Consortium. *Genes Immun.* **2009**, *10*, S1–S4, doi:10.1038/gene.2009.84.
107. Consortium, W.T.C.C. Genome-Wide Association Study of 14,000 Cases of Seven Common Diseases and 3,000 Shared Controls. *Nature* **2007**, *447*, 661–678, doi:10.1038/nature05911.
108. Morris, A.D.; Boyle, D.I.; MacAlpine, R.; Emslie-Smith, A.; Jung, R.T.; Newton, R.W.; MacDonald, T.M. The Diabetes Audit and Research in Tayside Scotland (Darts) Study: Electronic Record Linkage to Create a Diabetes Register. *BMJ* **1997**, *315*, 524–528, doi:10.1136/bmj.315.7107.524.
109. Eriksson, K.-F.; Lindgarde, F. Prevention of Type 2 (Non-Insulin-Dependent) Diabetes Mellitus by Diet and Physical Exercise The 6-Year Malmö Feasibility Study. *Diabetologia* **1991**, *34*, 891–898, doi:10.1007/BF00400196.
110. Groop, L.; Forsblom, C.; Lehtovirta, M.; Tuomi, T.; Karanko, S.; Nissen, M.; Ehrnstrom, B.-O.; Forsen, B.; Isomaa, B.; Snickars, B.; et al. Metabolic Consequences of a Family History of NIDDM (The Botnia Study): Evidence

for Sex-Specific Parental Effects. *Diabetes* **1996**, *45*, 1585–1593, doi:10.2337/diab.45.11.1585.

111. Feinleib, M.; Kannel, W.B.; Garrison, R.J.; McNamara, P.M.; Castelli, W.P. The Framingham Offspring Study. Design and Preliminary Data. *Prev. Med.* **1975**, *4*, 518–525, doi:10.1016/0091-7435(75)90037-7.
112. Voight, B.F.; Scott, and L.J.; Steinthorsdottir, V.; Morris, A.P.; Dina, C.; Welch, R.P.; Zeggini, E.; Huth, C.; Aulchenko, Y.S.; Thorleifsson, G.; et al. Twelve Type 2 Diabetes Susceptibility Loci Identified through Large-Scale Association Analysis. *Nat. Genet.* **2010**, *42*, 579–589, doi:10.1038/ng.609.
113. Friedman, G.D.; Cutter, G.R.; Donahue, R.P.; Hughes, G.H.; Hulley, S.B.; Jacobs, D.R.; Liu, K.; Savage, P.J. Cardia: Study Design, Recruitment, and Some Characteristics of the Examined Subjects. *J. Clin. Epidemiol.* **1988**, *41*, 1105–1116, doi:10.1016/0895-4356(88)90080-7.
114. Leitsalu, L.; Haller, T.; Esko, T.; Tammesoo, M.-L.; Alavere, H.; Snieder, H.; Perola, M.; Ng, P.C.; Mägi, R.; Milani, L.; et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *Int. J. Epidemiol.* **2015**, *44*, 1137–1147, doi:10.1093/ije/dyt268.
115. Sudlow, C.; Gallacher, J.; Allen, N.; Beral, V.; Burton, P.; Danesh, J.; Downey, P.; Elliott, P.; Green, J.; Landray, M.; et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Med.* **2015**, *12*, e1001779, doi:10.1371/journal.pmed.1001779.
116. Devuyst, O. The 1000 Genomes Project: Welcome to a New World. *Perit. Dial. Int.* **2015**, *35*, 676–677, doi:10.3747/pdi.2015.00261.
117. DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium; Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; Mexican American Type 2 Diabetes (MAT2D) Consortium; Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) Consortium; Mahajan, A.; Go, M.J.; Zhang, W.; Below, J.E.; Gaulton, K.J.; et al. Genome-Wide Trans-Ancestry Meta-Analysis Provides

Insight into the Genetic Architecture of Type 2 Diabetes Susceptibility. *Nat. Genet.* **2014**, *46*, 234–244, doi:10.1038/ng.2897.

118. Vassy, J.L.; Hivert, M.-F.; Porneala, B.; Dauriz, M.; Florez, J.C.; Dupuis, J.; Siscovick, D.S.; Fornage, M.; Rasmussen-Torvik, L.J.; Bouchard, C.; et al. Polygenic Type 2 Diabetes Prediction at the Limit of Common Variant Detection. *Diabetes* **2014**, *63*, 2172–2182, doi:10.2337/db13-1663.
119. Frayling, T.M. Genome–Wide Association Studies Provide New Insights into Type 2 Diabetes Aetiology. *Nat. Rev. Genet.* **2007**, *8*, 657–662, doi:10.1038/nrg2178.
120. Zeggini, E.; Scott, L.J.; Saxena, R.; Voight, B.F.; Marchini, J.L.; Hu, T.; de Bakker, P.I.; Abecasis, G.R.; Almgren, P.; Andersen, G.; et al. Meta-Analysis of Genome-Wide Association Data and Large-Scale Replication Identifies Additional Susceptibility Loci for Type 2 Diabetes. *Nat. Genet.* **2008**, *40*, 638–645, doi:10.1038/ng.120.
121. Gloyn, A.L.; Weedon, M.N.; Owen, K.R.; Turner, M.J.; Knight, B.A.; Hitman, G.; Walker, M.; Levy, J.C.; Sampson, M.; Halford, S.; et al. Large-Scale Association Studies of Variants in Genes Encoding the Pancreatic -Cell KATP Channel Subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) Confirm That the KCNJ11 E23K Variant Is Associated With Type 2 Diabetes. *Diabetes* **2003**, *52*, 568–572, doi:10.2337/diabetes.52.2.568.
122. Grant, S.F.A.; Thorleifsson, G.; Reynisdottir, I.; Benediktsson, R.; Manolescu, A.; Sainz, J.; Helgason, A.; Stefansson, H.; Emilsson, V.; Helgadottir, A.; et al. Variant of Transcription Factor 7-like 2 (TCF7L2) Gene Confers Risk of Type 2 Diabetes. *Nat. Genet.* **2006**, *38*, 320–323, doi:10.1038/ng1732.
123. Saxena, R.; Voight, B.F.; Lyssenko, V.; Burtt, N.P.; de Bakker, P.I.W.; Chen, H.; Roix, J.J.; Kathiresan, S.; Hirschhorn, J.N.; Daly, M.J.; et al. Genome-Wide Association Analysis Identifies Loci for Type 2 Diabetes and Triglyceride Levels. *Science* **2007**, *316*, 1331–1336, doi:10.1126/science.1142358.
124. Scott, L.J.; Mohlke, K.L.; Bonnycastle, L.L.; Willer, C.J.; Li, Y.; Duren, W.L.; Erdos, M.R.; Stringham, H.M.; Chines, P.S.; Jackson, A.U.; et al. A Genome-Wide Association Study of Type 2 Diabetes in Finns Detects Multiple

Susceptibility Variants. *Science* **2007**, *316*, 1341–1345, doi:10.1126/science.1142382.

125. Sladek, R.; Rocheleau, G.; Rung, J.; Dina, C.; Shen, L.; Serre, D.; Boutin, P.; Vincent, D.; Belisle, A.; Hadjadj, S.; et al. A Genome-Wide Association Study Identifies Novel Risk Loci for Type 2 Diabetes. *Nature* **2007**, *445*, 881–885, doi:10.1038/nature05616.
126. FinnGen 2018 GWAS: European Ancestry | Knowledge Portal Network Available online: https://www.kp4cd.org/node/352 (accessed on 10 August 2022).
127. Bonifacio, E.; Beyerlein, A.; Hippich, M.; Winkler, C.; Vehik, K.; Weedon, M.N.; Laimighofer, M.; Hattersley, A.T.; Krumsiek, J.; Frohnert, B.I.; et al. Genetic Scores to Stratify Risk of Developing Multiple Islet Autoantibodies and Type 1 Diabetes: A Prospective Study in Children. *PLOS Med.* **2018**, *15*, e1002548, doi:10.1371/journal.pmed.1002548.
128. Ashley, E.A.; Butte, A.J.; Wheeler, M.T.; Chen, R.; Klein, T.E.; Dewey, F.E.; Dudley, J.T.; Ormond, K.E.; Pavlovic, A.; Morgan, A.A.; et al. Clinical Assessment Incorporating a Personal Genome. *The Lancet* **2010**, *375*, 1525–1535, doi:10.1016/s0140-6736(10)60452-7.
129. Manolio, T.A. Bringing Genome-Wide Association Findings into Clinical Use. *Nat. Rev. Genet.* **2013**, *14*, 549–558, doi:10.1038/nrg3523.
130. Shmueli, G. To Explain or To Predict? *SSRN Electron. J.* **2010**, doi:10.2139/ssrn.1351252.
131. Wei, Z.; Wang, K.; Qu, H.-Q.; Zhang, H.; Bradfield, J.; Kim, C.; Frackleton, E.; Hou, C.; Glessner, J.T.; Chiavacci, R.; et al. From Disease Association to Risk Assessment: An Optimistic View from Genome-Wide Association Studies on Type 1 Diabetes. *PLoS Genet.* **2009**, *5*, e1000678, doi:10.1371/journal.pgen.1000678.
132. Abraham, G.; Inouye, M. Genomic Risk Prediction of Complex Human Disease and Its Clinical Application. *Curr. Opin. Genet. Dev.* **2015**, *33*, 10–16, doi:10.1016/j.gde.2015.06.005.
133. Dasgupta, A.; Sun, Y.V.; König, I.R.; Bailey-Wilson, J.E.; Malley, J.D. Brief Review of Regression-Based and Machine Learning Methods in Genetic

Epidemiology: The Genetic Analysis Workshop 17 Experience. *Genet. Epidemiol.* **2011**, *35*, S5–S11, doi:10.1002/gepi.20642.

134. Okser, S.; Pahikkala, T.; Airola, A.; Salakoski, T.; Ripatti, S.; Aittokallio, T. Regularized Machine Learning in the Genetic Prediction of Complex Traits. *PLoS Genet.* **2014**, *10*, e1004754, doi:10.1371/journal.pgen.1004754.
135. Mehta, P.; Bukov, M.; Wang, C.-H.; Day, A.G.R.; Richardson, C.; Fisher, C.K.; Schwab, D.J. A High-Bias, Low-Variance Introduction to Machine Learning for Physicists. *Phys. Rep.* **2019**, *810*, 1–124, doi:10.1016/j.physrep.2019.03.001.
136. Cruz, J.A.; Wishart, D.S. Applications of Machine Learning in Cancer Prediction and Prognosis. *Cancer Inform.* **2006**, *2*, 117693510600200, doi:10.1177/117693510600200030.
137. Palaniappan, S.; Awang, R. Intelligent Heart Disease Prediction System Using Data Mining Techniques. In Proceedings of the 2008 IEEE/ACS International Conference on Computer Systems and Applications; IEEE, March 2008.
138. Yu, W.; Liu, T.; Valdez, R.; Gwinn, M.; Khoury, M.J. Application of Support Vector Machine Modeling for Prediction of Common Diseases: The Case of Diabetes and Pre-Diabetes. *BMC Med. Inform. Decis. Mak.* **2010**, *10*, doi:10.1186/1472-6947-10-16.
139. Zhang, D.; Shen, D. Multi-Modal Multi-Task Learning for Joint Prediction of Multiple Regression and Classification Variables in Alzheimer\textquotesingles Disease. *NeuroImage* **2012**, *59*, 895–907, doi:10.1016/j.neuroimage.2011.09.069.
140. Paré, G.; Mao, S.; Deng, W.Q. A Machine-Learning Heuristic to Improve Gene Score Prediction of Polygenic Traits. *Sci. Rep.* **2017**, *7*, 12665, doi:10.1038/s41598-017-13056-1.
141. Vega, F.M.D.L.; Bustamante, C.D. Polygenic Risk Scores: A Biased Prediction? *Genome Med.* **2018**, *10*, doi:10.1186/s13073-018-0610-x.
142. Choi, S.W.; Mak, T.S.H.; O'Reilly, P.F. A Guide to Performing Polygenic Risk Score Analyses. **2018**, doi:10.1101/416545.
143. Reisberg, S.; Iljasenko, T.; Läll, K.; Fischer, K.; Vilo, J. Comparing Distributions of Polygenic Risk Scores of Type 2 Diabetes and Coronary Heart

Disease within Different Populations. *PLOS ONE* **2017**, *12*, e0179238, doi:10.1371/journal.pone.0179238.

144. Dick, D.M. Gene-Environment Interaction in Psychological Traits and Disorders. *Annu. Rev. Clin. Psychol.* **2011**, *7*, 383–409, doi:10.1146/annurev-clinpsy-032210-104518.
145. Franks, P.W. Gene × Environment Interactions in Type 2 Diabetes. *Curr. Diab. Rep.* **2011**, *11*, 552–561, doi:10.1007/s11892-011-0224-9.
146. Thomas, D. Gene–Environment-Wide Association Studies: Emerging Approaches. *Nat. Rev. Genet.* **2010**, *11*, 259–272, doi:10.1038/nrg2764.
147. Boffetta, P.; Winn, D.M.; Ioannidis, J.P.; Thomas, D.C.; Little, J.; Smith, G.D.; Cogliano, V.J.; Hecht, S.S.; Seminara, D.; Vineis, P.; et al. Recommendations and Proposed Guidelines for Assessing the Cumulative Evidence on Joint Effects of Genes and Environments on Cancer Occurrence in Humans. *Int. J. Epidemiol.* **2012**, *41*, 686–704, doi:10.1093/ije/dys010.
148. McAllister, K.; Mechanic, L.E.; Amos, C.; Aschard, H.; Blair, I.A.; Chatterjee, N.; Conti, D.; Gauderman, W.J.; Hsu, L.; Hutter, C.M.; et al. Current Challenges and New Opportunities for Gene-Environment Interaction Studies of Complex Diseases. *Am. J. Epidemiol.* **2017**, *186*, 753–761, doi:10.1093/aje/kwx227.
149. Kraft, P.; Yen, Y.-C.; Stram, D.O.; Morrison, J.; Gauderman, W.J. Exploiting Gene-Environment Interaction to Detect Genetic Associations. *Hum. Hered.* **2007**, *63*, 111–119, doi:10.1159/000099183.
150. Aschard, H. A Perspective on Interaction Effects in Genetic Association Studies. *Genet. Epidemiol.* **2016**, *40*, 678–688, doi:10.1002/gepi.21989.
151. Muhammed, A.; Zaki, M.T.; Elserafy, A.S.; Amin, S.A. Correlation between Prediabetes and Coronary Artery Disease Severity in Patients Undergoing Elective Coronary Angiography. *Egypt. Heart J.* **2019**, *71*, 34, doi:10.1186/s43044-019-0034-y.
152. Jackson, A.M.; Rørth, R.; Liu, J.; Kristensen, S.L.; Anand, I.S.; Claggett, B.L.; Cleland, J.G.F.; Chopra, V.K.; Desai, A.S.; Ge, J.; et al. Diabetes and Pre-Diabetes in Patients with Heart Failure and Preserved Ejection Fraction. *Eur. J. Heart Fail.* **2022**, *24*, 497–509, doi:10.1002/ejhf.2403.

153. Tuso, P. Prediabetes and Lifestyle Modification: Time to Prevent a Preventable Disease. *Perm. J.* **2014**, *18*, 88–93, doi:10.7812/TPP/14-002.
154. Aldossari, K.K.; Aldiab, A.; Al-Zahrani, J.M.; Al-Ghamdi, S.H.; Abdelrazik, M.; Batais, M.A.; Javad, S.; Nooruddin, S.; Razzak, H.A.; El-Metwally, A. Prevalence of Prediabetes, Diabetes, and Its Associated Risk Factors among Males in Saudi Arabia: A Population-Based Survey. *J. Diabetes Res.* **2018**, *2018*, 2194604, doi:10.1155/2018/2194604.
155. Nishida, C.; Ko, G.T.; Kumanyika, S. Body Fat Distribution and Noncommunicable Diseases in Populations: Overview of the 2008 WHO Expert Consultation on Waist Circumference and Waist-Hip Ratio. *Eur. J. Clin. Nutr.* **2010**, *64*, 2–5, doi:10.1038/ejcn.2009.139.
156. Why Some Patients Don't Take Prediabetes Seriously Available online: https://www.ama-assn.org/delivering-care/diabetes/why-some-patients-dont-take-prediabetes-seriously (accessed on 14 September 2022).
157. Lee, Y.; Siddiqui, W.J. Cholesterol Levels. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2022.
158. Pappan, N.; Rehman, A. Dyslipidemia. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2022.
159. Kansal, S.; Kamble, T.K. Lipid Profile in Prediabetes. *J. Assoc. Physicians India* **2016**, *64*, 18–21.
160. Al Amri, T.; Bahijri, S.; Al-Raddadi, R.; Ajabnoor, G.; Al Ahmadi, J.; Jambi, H.; Borai, A.; Tuomilehto, J. The Association Between Prediabetes and Dyslipidemia Among Attendants of Primary Care Health Centers in Jeddah, Saudi Arabia. *Diabetes Metab. Syndr. Obes. Targets Ther.* **2019**, *12*, 2735–2743, doi:10.2147/DMSO.S233717.
161. Seral-Cortes, M.; Sabroso-Lasa, S.; De Miguel-Etayo, P.; Gonzalez-Gross, M.; Gesteiro, E.; Molina-Hidalgo, C.; De Henauw, S.; Gottrand, F.; Mavrogianni, C.; Manios, Y.; et al. Development of a Genetic Risk Score to Predict the Risk of Overweight and Obesity in European Adolescents from the HELENA Study. *Sci. Rep.* **2021**, *11*, 3067, doi:10.1038/s41598-021-82712-4.
162. Hüls, A.; Krämer, U.; Carlsten, C.; Schikowski, T.; Ickstadt, K.; Schwender, H. Comparison of Weighting Approaches for Genetic Risk Scores in Gene-

Environment Interaction Studies. *BMC Genet.* **2017**, *18*, doi:10.1186/s12863-017-0586-3.

163. Sheikh, M.A.; Lund, E.; Braaten, T. The Predictive Effect of Body Mass Index on Type 2 Diabetes in the Norwegian Women and Cancer Study. *Lipids Health Dis.* **2014**, *13*, 164, doi:10.1186/1476-511X-13-164.
164. Boye, K.S.; Lage, M.J.; Thieu, V.; Shinde, S.; Dhamija, S.; Bae, J.P. Obesity and Glycemic Control among People with Type 2 Diabetes in the United States: A Retrospective Cohort Study Using Insurance Claims Data. *J. Diabetes Complications* **2021**, *35*, 107975, doi:10.1016/j.jdiacomp.2021.107975.
165. Nguyen, N.T.K.; Vo, N.-P.; Huang, S.-Y.; Wang, W. Fat-Free Mass and Skeletal Muscle Mass Gain Are Associated with Diabetes Remission after Laparoscopic Sleeve Gastrectomy in Males but Not in Females. *Int. J. Environ. Res. Public. Health* **2022**, *19*, 978, doi:10.3390/ijerph19020978.
166. Hudda, M.T.; Aarestrup, J.; Owen, C.G.; Cook, D.G.; Sørensen, T.I.A.; Rudnicka, A.R.; Baker, J.L.; Whincup, P.H.; Nightingale, C.M. Association of Childhood Fat Mass and Weight With Adult-Onset Type 2 Diabetes in Denmark. *JAMA Netw. Open* **2021**, *4*, e218524, doi:10.1001/jamanetworkopen.2021.8524.
167. Sun, J.; Liu, Z.; Zhang, Z.; Zeng, Z.; Kang, W. The Correlation of Prediabetes and Type 2 Diabetes With Adiposity in Adults. *Front. Nutr.* **2022**, *9*.
168. CDC Defining Adult Overweight and Obesity Available online: https://www.cdc.gov/obesity/basics/adult-defining.html (accessed on 15 September 2022).
169. Baker, C.F.; Overvad, K.; Dahm, C.C. Lean Body Mass and Risk of Type 2 Diabetes - a Danish Cohort Study. *J. Diabetes Metab. Disord.* **2019**, *18*, 445–451, doi:10.1007/s40200-019-00438-7.
170. Bellou, V.; Belbasis, L.; Tzoulaki, I.; Evangelou, E. Risk Factors for Type 2 Diabetes Mellitus: An Exposure-Wide Umbrella Review of Meta-Analyses. *PloS One* **2018**, *13*, e0194127, doi:10.1371/journal.pone.0194127.
171. Al Hommos, N.A.; Ebenibo, S.; Edeoga, C.; Dagogo-Jack, S. Trajectories of Body Weight and Fat Mass in Relation to Incident Prediabetes in a Biracial

Cohort of Free-Living Adults. *J. Endocr. Soc.* **2020**, *5*, bvaa164, doi:10.1210/jendso/bvaa164.

172. LeCroy, M.N.; Hua, S.; Kaplan, R.C.; Sotres-Alvarez, D.; Qi, Q.; Thyagarajan, B.; Gallo, L.C.; Pirzada, A.; Daviglus, M.L.; Schneiderman, N.; et al. Associations of Changes in Fat Free Mass with Risk for Type 2 Diabetes: Hispanic Community Health Study/Study of Latinos. *Diabetes Res. Clin. Pract.* **2021**, *171*, 108557, doi:10.1016/j.diabres.2020.108557.
173. Xu, M.; Hu, H.; Deng, D.; Chen, M.; Xu, Z.; Wang, Y. Prediabetes Is Associated with Genetic Variations in the Gene Encoding the Kir6.2 Subunit of the Pancreatic ATP-Sensitive Potassium Channel (KCNJ11): A Case-Control Study in a Han Chinese Youth Population. *J. Diabetes* **2018**, *10*, 121–129, doi:10.1111/1753-0407.12565.
174. Scuteri, A.; Sanna, S.; Chen, W.-M.; Uda, M.; Albai, G.; Strait, J.; Najjar, S.; Nagaraja, R.; Orrú, M.; Usala, G.; et al. Genome-Wide Association Scan Shows Genetic Variants in the FTO Gene Are Associated with Obesity-Related Traits. *PLoS Genet.* **2007**, *3*, e115, doi:10.1371/journal.pgen.0030115.
175. Frayling, T.M.; Timpson, N.J.; Weedon, M.N.; Zeggini, E.; Freathy, R.M.; Lindgren, C.M.; Perry, J.R.B.; Elliott, K.S.; Lango, H.; Rayner, N.W.; 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*, 889–894, doi:10.1126/science.1141634.
176. Loos, R.J.F.; Lindgren, C.M.; Li, S.; Wheeler, E.; Zhao, J.H.; Prokopenko, I.; Inouye, M.; Freathy, R.M.; Attwood, A.P.; Beckmann, J.S.; et al. Common Variants near MC4R Are Associated with Fat Mass, Weight and Risk of Obesity. *Nat. Genet.* **2008**, *40*, 768–775, doi:10.1038/ng.140.
177. Adamska-Patruno, E.; Goscik, J.; Czajkowski, P.; Maliszewska, K.; Ciborowski, M.; Golonko, A.; Wawrusiewicz-Kurylonek, N.; Citko, A.; Waszczeniuk, M.; Kretowski, A.; et al. The MC4R Genetic Variants Are Associated with Lower Visceral Fat Accumulation and Higher Postprandial Relative Increase in Carbohydrate Utilization in Humans. *Eur. J. Nutr.* **2019**, *58*, 2929–2941, doi:10.1007/s00394-019-01955-0.

178. Blaak, E.E.; Antoine, J.-M.; Benton, D.; Björck, I.; Bozzetto, L.; Brouns, F.; Diamant, M.; Dye, L.; Hulshof, T.; Holst, J.J.; et al. Impact of Postprandial Glycaemia on Health and Prevention of Disease. *Obes. Rev.* **2012**, *13*, 923–984, doi:10.1111/j.1467-789X.2012.01011.x.
179. Kumar, A.A.; Satheesh, G.; Vijayakumar, G.; Chandran, M.; Prabhu, P.R.; Simon, L.; Kutty, V.R.; Kartha, C.C.; Jaleel, A. Postprandial Metabolism Is Impaired in Overweight Normoglycemic Young Adults without Family History of Diabetes. *Sci. Rep.* **2020**, *10*, 353, doi:10.1038/s41598-019-57257-2.
180. Haupt, A.; Thamer, C.; Heni, M.; Machicao, F.; Machann, J.; Schick, F.; Stefan, N.; Fritsche, A.; Häring, H.-U.; Staiger, H. Novel Obesity Risk Loci Do Not Determine Distribution of Body Fat Depots: A Whole-Body MRI/MRS Study. *Obesity* **2010**, *18*, 1212–1217, doi:10.1038/oby.2009.413.
181. Speliotes, E.K.; Willer, C.J.; Berndt, S.I.; Monda, K.L.; Thorleifsson, G.; Jackson, A.U.; Lango Allen, H.; Lindgren, C.M.; Luan, J.; Mägi, R.; et al. Association Analyses of 249,796 Individuals Reveal 18 New Loci Associated with Body Mass Index. *Nat. Genet.* **2010**, *42*, 937–948, doi:10.1038/ng.686.
182. Kulyté, A.; Rydén, M.; Mejhert, N.; Dungner, E.; Sjölin, E.; Arner, P.; Dahlman, I. MTCH2 in Human White Adipose Tissue and Obesity. *J. Clin. Endocrinol. Metab.* **2011**, *96*, E1661–E1665, doi:10.1210/jc.2010-3050.
183. Mariman, E.C.M.; Bouwman, F.G.; Aller, E.E.J.G.; van Baak, M.A.; Wang, P. Extreme Obesity Is Associated with Variation in Genes Related to the Circadian Rhythm of Food Intake and Hypothalamic Signaling. *Physiol. Genomics* **2015**, *47*, 225–231, doi:10.1152/physiolgenomics.00006.2015.
184. Chen, Y.; Shen, M.; Ji, C.; Huang, Y.; Shi, Y.; Ji, L.; Qin, Y.; Gu, Y.; Fu, Q.; Chen, H.; et al. Genome-Wide Identification of N6-Methyladenosine Associated SNPs as Potential Functional Variants for Type 1 Diabetes. *Front. Endocrinol.* **2022**, *13*.
185. McAllan, L.; Baranasic, D.; Villicaña, S.; Zhang, W.; Lehne, B.; Adamo, M.; Jenkinson, A.; Elkalaawy, M.; Mohammadi, B.; Hashemi, M.; et al. *Integrative Genomic Analyses in Adipocytes Implicate DNA Methylation in Human Obesity and Diabetes*; Genetic and Genomic Medicine, 2021;

186. Tait, S.; Baldassarre, A.; Masotti, A.; Calura, E.; Martini, P.; Varì, R.; Scazzocchio, B.; Gessani, S.; Del Cornò, M. Integrated Transcriptome Analysis of Human Visceral Adipocytes Unravels Dysregulated MicroRNA-Long Non-Coding RNA-MRNA Networks in Obesity and Colorectal Cancer. *Front. Oncol.* **2020**, *10*, 1089, doi:10.3389/fonc.2020.01089.
187. Lv, D.; Zhou, D.; Zhang, Y.; Zhang, S.; Zhu, Y.-M. Two Obesity Susceptibility Loci in LYPLAL1 and ETV5 Independently Associated with Childhood Hypertension in Chinese Population. *Gene* **2017**, *627*, 284–289, doi:10.1016/j.gene.2017.06.030.
188. Ofori, J.K.; Karagiannopoulos, A.; Nagao, M.; Westholm, E.; Ramadan, S.; Wendt, A.; Esguerra, J.L.S.; Eliasson, L. Human Islet MicroRNA-200c Is Elevated in Type 2 Diabetes and Targets the Transcription Factor ETV5 to Reduce Insulin Secretion. *Diabetes* **2022**, *71*, 275–284, doi:10.2337/db21-0077.
189. Gutierrez-Aguilar, R.; Kim, D.-H.; Casimir, M.; Dai, X.-Q.; Pfluger, P.T.; Park, J.; Haller, A.; Donelan, E.; Park, J.; D'Alessio, D.; et al. The Role of the Transcription Factor ETV5 in Insulin Exocytosis. *Diabetologia* **2014**, *57*, 383–391, doi:10.1007/s00125-013-3096-5.
190. Plaza-Florido, A.; Pérez-Prieto, I.; Molina-Garcia, P.; Radom-Aizik, S.; Ortega, F.B.; Altmäe, S. Transcriptional and Epigenetic Response to Sedentary Behavior and Physical Activity in Children and Adolescents: A Systematic Review. *Front. Pediatr.* **2022**, *10*.
191. Cheung, C.Y.Y.; Tso, A.W.K.; Cheung, B.M.Y.; Xu, A.; Ong, K.L.; Fong, C.H.Y.; Wat, N.M.S.; Janus, E.D.; Sham, P.C.; Lam, K.S.L. Obesity Susceptibility Genetic Variants Identified from Recent Genome-Wide Association Studies: Implications in a Chinese Population. *J. Clin. Endocrinol. Metab.* **2010**, *95*, 1395–1403, doi:10.1210/jc.2009-1465.
192. Kotnik, P.; Knapič, E.; Kokošar, J.; Kovač, J.; Jerala, R.; Battelino, T.; Horvat, S. Identification of Novel Alleles Associated with Insulin Resistance in Childhood Obesity Using Pooled-DNA Genome-Wide Association Study Approach. *Int. J. Obes.* **2018**, *42*, 686–695, doi:10.1038/ijo.2017.293.

193. Ng, M.C.Y.; Tam, C.H.T.; So, W.Y.; Ho, J.S.K.; Chan, A.W.; Lee, H.M.; Wang, Y.; Lam, V.K.L.; Chan, J.C.N.; Ma, R.C.W. Implication of Genetic Variants Near NEGR1, SEC16B, TMEM18, ETV5/DGKG, GNPDA2, LIN7C/BDNF, MTCH2, BCDIN3D/FAIM2, SH2B1, FTO, MC4R, and KCTD15 with Obesity and Type 2 Diabetes in 7705 Chinese. *J. Clin. Endocrinol. Metab.* **2010**, *95*, 2418–2425, doi:10.1210/jc.2009-2077.

194. Virtanen, K.A.; Iozzo, P.; Hällsten, K.; Huupponen, R.; Parkkola, R.; Janatuinen, T.; Lönnqvist, F.; Viljanen, T.; Rönnemaa, T.; Lönnroth, P.; et al. Increased Fat Mass Compensates for Insulin Resistance in Abdominal Obesity and Type 2 Diabetes : A Positron-Emitting Tomography Study. *Diabetes* **2005**, *54*, 2720–2726, doi:10.2337/diabetes.54.9.2720.

195. Wondmkun, Y.T. Obesity, Insulin Resistance, and Type 2 Diabetes: Associations and Therapeutic Implications. *Diabetes Metab. Syndr. Obes. Targets Ther.* **2020**, *13*, 3611–3616, doi:10.2147/DMSO.S275898.