



Review

# Systematic Review of Polygenic Risk Scores for Type 1 and Type 2 Diabetes

Felipe Padilla-Martínez <sup>1,2,\*</sup> , Francois Collin <sup>1</sup> , Miroslaw Kwasniewski <sup>1</sup> and Adam Kretowski <sup>2,3</sup>

<sup>1</sup> Centre for Bioinformatics and Data Analysis, Medical University of Bialystok, 15-276 Bialystok, Poland; francois.collin@umb.edu.pl (F.C.); miroslaw.kwasniewski@umb.edu.pl (M.K.)

<sup>2</sup> Clinical Research Centre, Medical University of Bialystok, 15-276 Bialystok, Poland; adamkretowski@wp.pl

<sup>3</sup> Department of Endocrinology, Diabetology and Internal Medicine, Medical University of Bialystok, 15-276 Bialystok, Poland

\* Correspondence: luis.padilla@umb.edu.pl; Tel.: +48-5182-93299

Received: 3 February 2020; Accepted: 28 February 2020; Published: 2 March 2020



**Abstract:** Recent studies have led to considerable advances in the identification of genetic variants associated with type 1 and type 2 diabetes. An approach for converting genetic data into a predictive measure of disease susceptibility is to add the risk effects of loci into a polygenic risk score. In order to summarize the recent findings, we conducted a systematic review of studies comparing the accuracy of polygenic risk scores developed during the last two decades. We selected 15 risk scores from three databases (Scopus, Web of Science and PubMed) enrolled in this systematic review. We identified three polygenic risk scores that discriminate between type 1 diabetes patients and healthy people, one that discriminate between type 1 and type 2 diabetes, two that discriminate between type 1 and monogenic diabetes and nine polygenic risk scores that discriminate between type 2 diabetes patients and healthy people. Prediction accuracy of polygenic risk scores was assessed by comparing the area under the curve. The actual benefits, potential obstacles and possible solutions for the implementation of polygenic risk scores in clinical practice were also discussed. Develop strategies to establish the clinical validity of polygenic risk scores by creating a framework for the interpretation of findings and their translation into actual evidence, are the way to demonstrate their utility in medical practice.

**Keywords:** polygenic risk score; type 1 diabetes; type 2 diabetes; diagnosis; genetics

## 1. Introduction

Diabetes mellitus is a complex and heterogeneous group of chronic metabolic diseases characterized by hyperglycemia, now recognized as one of the most important 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% of adults of 18 years and older developed diabetes. Diabetes can be treated and its consequences avoided or delayed with diet, physical activity, medication and regular screening and treatment for complications [2].

Diabetes is commonly divided into three subtypes. Type 1 diabetes (T1D) occurs predominantly in people < 30 years old and is generally thought to be precipitated by an immune-associated destruction of insulin-producing pancreatic beta cells, leading to insulin deficiency and requiring exogenous insulin supplement [3]. Type 2 diabetes (T2D) is a progressive metabolic disease characterized by insulin resistance [4] and eventual functional failure of pancreatic beta cells [5,6]. Maturity-onset diabetes of the young (MODY) is a monogenic form of diabetes showing autosomal dominant mode of inheritance. It accounts for 1%–5% of all the diabetic forms of the young and is characterized by anomalous pancreatic beta-cell activity [7–9].

Between 2002 and 2012, there was an incidence increase of 1.8% and 4.8% for T1D and T2D among American youths, respectively. Variations in the prevalence of obesity over time may contribute to variations in insulin resistance and to the increasing incidence of T2D [10]. Differences in incidence have been reported in populations from the same ethnic group living in different environments, thereby highlighting the importance of environmental risk factors [11]. Different approaches in the diagnosis area present potential for reducing the mortality and the incidence of cardiovascular complications among patients with T1D and T2D [12]. There is an international epidemic in diabetes with increased prevalence reported globally [13]; as the proportion of diagnosed cases of diabetes increases, a similar increase in the cases of diabetes errors occur [14]. Errors in the primary care of diabetes are misclassification, miscoding and misdiagnosis. A recent study conducted a cross-sectional study in UK, trying to identify cases with potential errors; a high rate of errors was found (57%) compared with previous studies, demonstrating that the prevalence of errors in people with diabetes in primary care is growing [15,16]. The implications of wrong diagnosis, coding or classification affect optimal treatment regimens and cause inappropriate financial and psychological impacts in such patients. Patients with correct diagnosis achieve a significant improvement in their glycemic control [16,17].

The key mission of genomics medicine aims to predict the genetic disease risk on the basis of an individual's genotype [18]. Identifying those in the population who are at greater risk of disease can result in an improvement in the healthcare sector and lower costs by reducing unnecessary disease concern and by introducing preemptive therapies of lifestyle changes for those at greater risk [19]. The generation of genome-wide variation data has become common for prediction of metabolic diseases [20,21]. Many of the metabolic diseases such as coronary heart disease, atrial fibrillation and T2D [19,22] have well-established risk loci and likely contain many genetic determinants with effects too small to be detected at genome-wide levels of statistical significance [23]. This demonstrates that all common variants across the genome actually explain a much higher proportion of heritability (50% or more) in many complex traits than could be seen based on a small subset of significant single nucleotide polymorphism (SNP) only [24].

The risk of developing T1D or T2D is influenced by the combination of genetic variation in multiple sites across the genome [25,26]. Over the past decades, large-scale genetic studies have described over 400 distinct genetic signals affecting T2D risk [27] and over 50 with influence on T1D predisposition [28]. Genetic testing for T1D risk is not part of routine clinical care. This may in part be due to very modest individual risk effects of non-human leukocyte antigen (HLA) SNPs, historic expense in genotyping HLA alleles and SNPs, lack of available working treatments and a lack of widespread understanding of the complex HLA nomenclature [29]. The increasing prevalence of T2D is one of the greatest challenges in public health [30]. Although obesity is the strongest predictor of T2D, it is also known that heritability of T2D is 26%–69% depending on age of onset, thus motivating the search for genetic predictors for T2D [31–33]. An approach to convert genetic data to a predictive measure of disease susceptibility is to add the risk effects of loci into a single genetic risk score (GRS)-polygenic risk score (PRS) [34–36].

Prediction accuracy of a PRS is often assessed by measuring the area under the receiver operating characteristic (ROC) curve (AUC). The AUC compares the rates of true positives (sensitivity) and false positives (1—specificity) accounting for the overall performance of predictive models [37]. The use of PRSs could become useful to identify a group of patients at risk; this will offer substantial clinical benefits while preventing growing morbidity and mortality associated with diabetes [38–41]. Several research groups have developed diabetes PRSs, fitting the scoring models to their study area [19,36,38,39,42–51]. All of them have used AUC as a predictive parameter to identify the sensitivity and specificity of the outcome of interest. The estimation of T1D and T2D PRSs can be used for diagnosis-support in scientific and clinical environment. Thus, this review aims to identify and compare the most recent studies where a PRS has been established. These may give a lead for researchers to develop an innovative PRS

for T1D and T2D and improve existing ones, taking into account variables that have not been used or exploring cutting-edge algorithms.

## 2. Methods

### 2.1. Search Strategies

The databases for the literature search were chosen based on a recommendation of the optimal database combinations [52] and database accessibility in our institution. The three databases chosen for the literature search were Scopus, Web of Science and PubMed. The databases were searched for studies of polygenic risk score for T1D, T2D and monogenic diabetes, published between 2000 and September 2019. The keywords of the queries were “diabetes”, “type 1”, “type 2”, “maturity onset diabetes of the young”, “genetic risk scores”, “polygenic risk scores” and their combinations: “diabetes type 1 polygenic risk score”, “diabetes type 2 polygenic risk score”, “maturity onset diabetes of the young polygenic risk score” and “diabetes genetic risk scores”.

### 2.2. Study Selection

During the screening stage, the exclusion criteria was based on the criteria “the strengthening the reporting of genetic association studies (STREGA): An extension of the strengthening the reporting of observational studies in epidemiology (STROBE)” [53]. From the 22 items mentioned in the article, we took into account the following 12 items that studies must have extensively explained in order to go forward for the eligibility stage: in the introduction section, objectives; in the methods section, study design, setting, participants, variables, quantitative variables and statistical methods; in the results section, participants, descriptive data and main results; and in the discussion section, limitations and interpretation of the results.

### 2.3. Data Collection Processing

The items collected from the full text and Supplementary Information were first author, year of publication, digital object identifier (DOI) when available, ethnicity of study panel, country, data set of study, validation set when available, number of patients and controls, method of sequencing/genotyping, panel of genes used, the number of SNPs used to obtain the PRS, clinical risk factors, the AUC for the clinical risk factors and the AUC for the PRS.

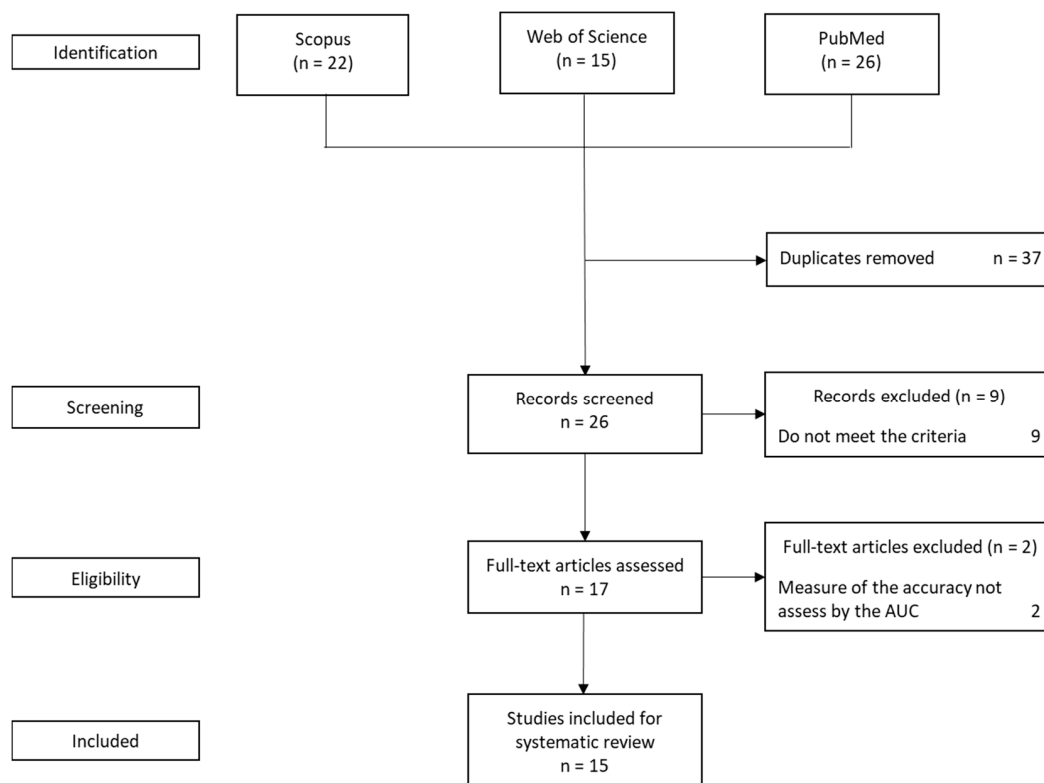
### 2.4. Synthesis of the Results

The AUC of the PRS was taken into account to assess the accuracy and for inter-PRS comparison purpose. The selected AUCs were grouped into three categories based on the diabetes subtypes to discriminate. The first group included T1D PRS comparison. The second group included T2D PRS comparison. The third group included T1D PRS comparison used to discriminate T1D vs. T2D and T1D vs. monogenic diabetes.

## 3. Results

### 3.1. Selected Studies for the Systematic Review

A total of 63 articles were retrieved from Pubmed, Scopus and Web of Science. After removing the duplicates, the total number of studies obtained was 26. The quality of selected studies for the next phase of the screening stage was evaluated using a modified criteria [53], and nine articles were excluded due to the lack of strong arguments on the 12 items of the criteria selected. In the final stage, two articles were excluded as a result of using a different measuring technique for the accuracy of the results. For the systematic review, 15 studies were selected (Figure 1). These studies have varying sources of data sets, panel of genes and genotyping strategies.



**Figure 1.** Flow diagram for study selection.

There were six studies eligible for the systematic review, which developed PRSs for T1D [38,39,42–45], and there were nine that studied PRSs for T2D [19,36,46–51,54] (Table 1). The majority of the studies were conducted in Caucasian populations, while some of them conducted the studies in Hispanic, African-American, Asian-American, South African and Iranian populations. Apart from the Iranian and South African cohort, all of the other studies had large sample sizes in their patients and their controls. The databases from where the articles were retrieved are shown in Table 1.

**Table 1.** The studies selected for the systematic review.

Study	Year	Country/Ethnicity	Patients	Controls	Database
<b>Studies focusing on type 1 diabetes</b>					
Winkler et al. [42]	2014	Caucasian	4574	1207	PubMed, Scopus
Oram et al. [38]	2015	Caucasian	<i>n</i> = 1938		PubMed
Patel et al. [39]	2016	Caucasian	1963	805	PubMed, Scopus
Perry et al. [43]	2018	Caucasian, Hispanic, African-American and Asian-American	627	423	PubMed, Scopus, Web of Science
Sharp et al. [44]	2019	Caucasian	6670	9416	PubMed
Yaghootkar et al. [45]	2019	Iranian	121	6	PubMed, Web of Science
<b>Studies focusing on type 2 diabetes</b>					
Weedon et al. [46]	2006	British	2409	3669	PubMed, Scopus
Lango et al. [47]	2008	Scotland	2309	2598	PubMed
Lyssenko et al. [48]	2008	Finland	2201	16,630	PubMed
Meigs et al. [49]	2008	European ancestry in USA	<i>n</i> = 2776		PubMed, Scopus, Web of Science
Chatterjee et al. [50]	2013	Caucasian	130	38,987	PubMed
Vassy et al. [51]	2014	European ancestry in USA	5941	5942	PubMed, Scopus
Läll et al. [36]	2016	Estonia	1181	9092	PubMed, Scopus
Chikowore et al. [54]	2016	South African	178	178	PubMed
Amit et al. [19]	2018	British	26,676	120,280	PubMed, Web of Science

The studies relied on datasets from different sources: T1DGC [55], WTCCC [56], UFDI, Iranian Hospitals [35], the PURE study [57], UK hospital [45], GoDARTS [58], MPP [59], BPS [60], Framingham

Offspring Study [61], Voight [62], CARDIA [63], the Estonian Biobank [64] and the UK Biobank [65] (Table 2).

**Table 2.** Data set source, panel of genes used and genotyping strategies.

Study	Year	Data Set	Panel of Genes	Platform
<b>Studies focusing on type 1 diabetes</b>				
Winkler et al. [42]	2014	T1DGC	T1DGC	TaqMan 5'nuclease assay
Oram et al. [38]	2015	WTCCC	1000 genomes and T1DGC	Affymetrix 500K SNP chip
Patel et al. [39]	2016	WTCCC	1000 genomes and T1DGC	Affymetrix 500K SNP chip
Perry et al. [43]	2018	University of Florida diabetes institute (UFDI)	Immunobase.org October 2017	Taqman SNP genotyping array
Sharp et al. [44]	2019	T1DGC	1000 genomes	Affymetrix Axiom Array
Yaghootkar et al. [45]	2019	Imam Reza Hospital and Children's Medical Centre in Iran	1000 genomes and T1DGC	Targeted next-generation sequencing (unspecified)
<b>Studies focusing on type 2 diabetes</b>				
Weedon et al. [46]	2006	UK	KCNK11, PPARG, TCF7L2.	Modified TaqMan
Lango et al. [47]	2008	GoDARTS	Frayling [66] and Zeggini et al. [67]	Modified TaqMan
Lyssenko et al. [48]	2008	Malmö Preventive Project (MPP) and Botnia Prospective Study (BPS).	Gloyn et al. [68], Grant et al. [69], Saxena et al. [70], Frayling [66], Scott et al. [71], Sladek et al. [72], Steinthorsdottir et al. [73], Zeggini et al. [74], Zeggini et al. [67], Lyssenko et al. [75].	Allelic discrimination assay-by-design method, Allele-specific (KASPar)
Meigs et al. [49]	2008	The Framingham Offspring Study	Saxena et al. [70], Zeggini et al. [67]	iPLEX technology
Chatterjee et al. [50]	2013	Voight [62]	Voight et al. [62]	Illumina Omni 2.5M Platform
Vassy et al. [51]	2014	The Framingham Offspring Study and CARDIA	DIAGRAMv3	Taqman, Illumina's OPA technology, Affymetrix 6.0, Illumina 370 and 550
Läll et al. [36]	2016	The Estonian Biobank	DIAGRAM Consortium	Illumina Human OmniExpress, Illumina Cardio-MetaboChip
Chikowore et al. [54]	2016	The PURE study	Chikowore et al. [57]	BeadXpress platform, Illumina
Amit et al. [19]	2018	The UK Biobank	1000 genome phase 3 version 5 (Linkage disequilibrium panel)	Affymetrix UK BiLEVE Axiom array, Affymetrix UK Biobank Axiom

The studies also differed in the panel of genes included to obtain the PRS (Table 2). For T1D PRS, the studies combined either the panel of genes from T1DGC ( $n = 4$ ) [55], 1000 genomes project ( $n = 4$ ) [56] or the Immunobase.org on October 2017 ( $n = 1$ ). For T2D PRS, the studies used either specific genes from previous studies ( $n = 6$ ), different versions of the DIAGRAM Consortium panel of genes ( $n = 2$ ) [76] or the 1000 genomes project ( $n = 1$ ) [56].

Lastly the studies differed in the platform used for genotyping or sequencing (Table 2). Most of the studies used modified TaqMan assays ( $n = 5$ ), different versions of Affymetrix microarrays ( $n = 5$ ) and Illumina technology ( $n = 4$ ). One study used KASPar genotyping, another the iPLEX technology and another failed reporting the sequencing platform that was operated.

### 3.2. Polygenic Risk Score for T1D prediction

Interacting factors such as background genetic risk, infant and adult diet, environmental exposure, beta-cell stress and immune phenotype increase the development of autoimmunity and beta-cell loss in clinical T1D [29]. Type 1 diabetes has a substantial heritable component, estimated to be between 65% and 88% [77,78]. Genes in the HLA region confer 50% of the genetic risk of T1D. The HLA gene family provides the genetic blueprint for a group of related proteins that help the immune

system to discriminate the endogenous proteins from bacterial proteins. The genes in this complex are categorized into two major classes: class I and class II. Class-I HLA presents antigen peptides found within the cell, to CD8 positive (cytotoxic T cells) while Class-II HLA presents antigen peptides found outside the cell, to CD4 positive (helper T cells) [79]. The dominant genetic drivers of this risk are Class II HLA DR and DQ genes on chromosome 6. The HLA haplotypes DR3 and DR4–DQ8, are the two most significant risk haplotypes, with highest genetic risk for T1D occurring in the compound heterozygote [80]. The HLA class 1 alleles have been associated with T1D; A24 is associated with both T1D risk and progression of beta-cell loss [81] and B3906 has been shown to modulate risk when present only with specific class 2 haplotypes [82] and B57 [83]. More than 60 common non-HLA T1D risk variants across the genome have been identified in linkage and genome wide association studies (GWAS) in genes including INS, PTPN22, CTLA-4 and IL2RA [84,85].

Genetic prediction for T1D has evolved from the use of HLA alleles alone [86] to the incorporation of non-HLA variants. In 2014, Winkler et al. developed a multivariate logistic regression model to estimate PRSs including 40 non-HLA genes SNPs, improving significantly the risk score with an AUC of 0.87 comparing to the control [42]. Oram and colleagues [38] adjusted a log-additive PRS model to discriminate patients versus controls for T1D and T2D. They applied a 69 SNP T2D-PRS and a 30 SNP T1D-PRS to a panel of T1D and T2D. They found the T1D-PRS was highly discriminant with an AUC of 0.88, while the T2D-PRS was less discriminant with an AUC of 0.64, and the combination of the two increased slightly the AUC to 0.89. 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 [44] (Table 3).

**Table 3.** Comparison of the accuracy of T1D PRS assessed by the AUC.

Year	Author	Polygenic Risk Scores	Single-Nucleotide Polymorphism	Area under the Curve for Polygenic Risk Scores	Ethnicity
2014	Winkler et al. [42]	T1D	41	0.87	Caucasian
2015	Oram et al. [38]	T1D	30	0.88	Caucasian
2015	Oram et al. [38]	T1D + T2D	99	0.89	Caucasian
2018	Perry et al. [43]	T1D	32	0.86	Caucasian
2018	Perry et al. [43]	T1D	32	0.90	Caucasian Hispanic
2018	Perry et al. [43]	T1D	32	0.75	African-American
2018	Perry et al. [43]	T1D	32	0.92	Asian-American
2019	Sharp et al. [44]	T1D	67	0.93	Caucasian

Currently, the majority of genetics studies on T1D are limited to Caucasian cohorts. However, Perry et al. investigated the hypothesis that race and/or ethnicity would be contextually important for evaluating genetic risk markers previously identified from Caucasian cohorts [43]. They applied the GRS used by Oram [39] to Hispanic Caucasian, African-American and Asian-American populations. The Hispanic Caucasian GRS was highly discriminant with an AUC of 0.90. The GRS for Asian-American was also highly discriminant with an AUC of 0.92, and the analysis indicated that this PRS could discriminate T1D subjects from controls in a small cohort for subjects of Asian-American, but larger studies are required to validate and extend these findings. The African-American obtained a less discriminant GRS with an AUC of 0.75; notable risk differences were observed for three SNPs: SH2B3, CTRB1/2, GAB3 in this population [43] (Table 3).

### 3.3. Polygenic Risk Scores for T2D prediction

The ability of drugs and lifestyle interventions to lead to reduction in progression of diabetes motivates efforts to determine those at the greatest future risk of developing T2D [87,88]. Dozens of genetic loci associated with T2D risk have been discovered using GWAS; currently, 243 have been associated [89–91]. There is still some disbelief regarding the practical value of identified SNPs in personalized risk prediction for the disease [92]. The main reason is that the effect of individual SNPs on complex common disease phenotypes is relatively weak; in other words that the environment is the main driver of T2D condition [93].

Before the first GWAS for T2D, a research describing three genetic variants (*KCNJ11*, *PPARG* and *TCF7L2*) that had been associated with T2D was published, assessing the combined risk of these variants and the predictive value of the genetic tests using the AUC. The AUC was 0.58, outpacing the 0.50 value that expressed no discriminative capacity, but short of the values seen for clinical tests [46]. Two years later, Lango and colleagues examined a 16 SNPs PGR; the score, adjusted for age, BMI and sex, predicted diabetes incidence with an AUC of 0.789. Adding the PRS to these clinical factors had only a modest impact on performance, pushing the AUC to 0.80 [47]. In a similar study, a research group assessed 16 SNPs PGR; the score, adjusted for age, sex, family history, BMI, blood pressure, triglycerides and fasting plasma glucose, predicted diabetes incidence with an AUC of 0.740; adding the PRS to clinical risk factors (CRF), the AUC of 0.750 had a small effect on the ability to predict T2D [48]. In the same year Meigs et al. estimated an 18 SNP PRS; the AUC for incident diabetes adjusted for age and sex was 0.534, and an enhanced clinical model incorporating age, sex, family history, BMI, glucose level, cholesterol level and triglyceride level reached 0.90. Adding genetic data to those two PRSs increased the AUC, respectively, to 0.58 and 0.910 [49]. The next analysis was made with a 22 SNP PRS; this group of researchers deduced that a ten-fold increase in effective GWAS sample size for T2D would result in a better performance. The result of the AUC adjusted for age, sex and family history was 0.570; after adding the PRS, the AUC increased to 0.740 [50]. An updated analysis of a 62 SNP PRS generated an improved AUC for T2D prediction. Combined with age and sex, the AUC was 0.72, but after the addition of other important clinical factors the score was 0.91 [51]. A recent study analyzed a South African population; a PRS using only four SNPs was created, taking into account sex, age, BMI and systolic blood pressure as clinical risk factors, generating an AUC of 0.665 [54].

Larger GWAS for T2D have been developed in the last years; these achievements have increased the number of significant loci identified to the hundreds. Although comparing variants that carry out genome-wide significance assures that the variants included in the score represent legitimate associations with disease, an inflexible threshold ignores many other variants, which even if they are truly associated with the phenotype, have escaped detection at genome-wide significance due to defined sample sizes [36]. Khera et al. [19] used a different approach, handling 7 million variants, and after adding sex, age and other important clinical factor, generated an AUC of 0.73 (Table 4).

**Table 4.** Comparison of the accuracy of T2D PRS assessed by the AUC.

Year	Author	Polygenic Risk Cores (PRS)	Single-Nucleotide Polymorphism	Area under the Curve (AUC) for Clinical Risk Factors	AUC PRS + Clinical Risk Factors	Difference	Clinical Risk Factors	Ethnicity
2006	Weedon et al. [46]	T2D	3	-	0.580	-	-	Caucasian
2008	Lango et al. [47]	T2D	18	0.780	0.800	0.020	Age, BMI, sex	Caucasian
2008	Lyssenko et al. [48]	T2D	16	0.740	0.750	0.010	Age, sex, family, BMI, blood pressure, triglycerides, glucose	Caucasian
2008	Meigs et al. [49]	T2D	18	0.534	0.581	0.047	Age, sex	Caucasian
2008	Meigs et al. [49]	T2D	18	0.595	0.615	0.020	Sex, age, family	Caucasian
2008	Meigs et al. [49]	T2D	18	0.900	0.910	0.010	Age, sex, family, BMI, glucose, cholesterol, triglycerides	Caucasian
2013	Chatterjee et al. [50]	T2D	22	0.570	0.740	0.170	Age, sex, family	Caucasian
2014	Vassy et al. [51]	T2D	62	0.698	0.726	0.028	Age, sex	Caucasian, USA population
2014	Vassy et al. [51]	T2D	62	0.903	0.906	0.003	Sex, family, BMI, blood pressure, HDL cholesterol, triglyceride levels, age	Caucasian, USA population
2016	Läll et al. [36]	T2D-double weighted	1000	0.699	0.74	0.042	Sex, age	Caucasian
2016	Läll et al. [36]	T2D-dw	1000	0.718	0.767	0.049	Sex, age, BMI	Caucasian
2016	Läll et al. [36]	T2D-dw	1000	0.777	0.79	0.012	Sex, age, BMI, hypertension, high blood glucose, physical activity, smoking, food consumption	Caucasian
2016	Chikowore et al. [54]	T2D	4	0.652	0.665	0.013	Sex, age, BMI and blood pressure	African
2018	Amit et al. [19]	T2D	7 million	0.66	0.73	0.070	Sex, age	Caucasian



### 3.4. Polygenic Risk Scores to discriminate different subtypes of diabetes

As it was mentioned before there are three types of errors in primary care of diabetes, in this section the misclassifications will be reviewed. Misclassification refers to giving the patient a type of diabetes classification that he/she does not have [16]. Diagnosis of diabetes into its subtypes is becoming an increasing dispute. There is a growing problem of obesity in young adults and teenagers, and consequently an increase in T2D; this takes us to a misclassification of T1D [94]. There is just a small overlap in the genetics of T1D and T2D, thus a PRS could be used as a diagnostic tool [29,74]. Oram and colleagues [38] were the first ones to accomplish the use of a PRS to discriminate T1D and T2D. They developed a PRS of 30 SNPs, which consists of HLA and non-HLA loci; the AUC was of 0.880, being highly discriminant. They also demonstrated that the use of the top nine SNPs had the same highly discriminant effect that as the use of the 30 SNPs; the AUC for the PRS using the top nine SNPs was 0.873 (Table 5).

**Table 5.** Comparison of accuracy of T1D PRS to discriminate diabetes subtypes, assessed by the AUC.

Year	Author	Polygenic Risk Scores	Single-Nucleotide Polymorphism	Area under the Curve for Polygenic Risk Scores	Ethnicity
2015	Oram et al. [38]	T1D vs. T2D	30	0.88	Caucasian
2015	Oram et al. [38]	T1D vs. T2D	9	0.87	Caucasian
2016	Patel et al. [39]	T1D vs. MODY	30	0.87	Caucasian
2019	Yaghootkar et al. [45]	T1D vs. Monogenic	9	0.90	Iranian

Differentiating patients with monogenic diabetes from T1D has important significance in scientific and clinical environments, the treatment in the patient being a crucial factor [95,96]. Patel and colleagues [39], generated a T1D-PRS from 30 SNPs to discriminate monogenic diabetes from T1D. They assessed the ability of the PRS to discriminate T1D and confirmed monogenic diabetes. The analysis of the PRS using the AUC showed that it was highly discriminant, the value being 0.87. The latest study of Yaghootkar and his team [45] provided the first evidence to suggest that the T1D PRS proposed by Oram et al. [38] using the top nine SNPs for a European cohort may help to distinguish monogenic diabetes from T1D in an Iranian population. The AUC analysis showed that the T1D PRS was highly discriminant between monogenic and T1D in the non-European cohort with a value of 0.898, which was similar to the ability of the same PRS in the European cohort (Table 5).

## 4. Discussion

PRSs are powerful tools to support diagnosis; they are consistent throughout life, and thus they could be an effective tool to determine whether a particular patient has T1D, T2D or one of the other forms of diabetes. Thanks to them, it is less difficult to predict the risk of pre-symptomatic diabetes [97]. As stated earlier in the analysis of T1D PRS, with the exception of the cohort of African-Americans [43], the AUC had values of more than 0.80, meaning that all of the PRSs had enough sensitivity and specificity to discriminate patients with T1D. Genetic factors are making an important contribution in the prevention of T1D by giving a reliable risk score. PRS for T1D can diagnose young adults with diabetes that will require insulin treatment in European cohorts, and this will be important to classify accurately patients, when clinical factors make incorrect diagnosis. Using PRS as a tool to discriminate between diabetes subtypes is another advantage provided; the latest studies have shown that T1D PRS is great at discriminating between the patients with T1D versus T2D [38] and monogenic diabetes [39]. The option of using T1D PRS validated for the Caucasian cohort in other ethnicities is under study and could become feasible [45]. The correct diagnosis, as a result of using PRS, could help to generate a lifestyle modification and a pharmaceutical intervention to reduce T1D progression.

Nevertheless, after an extensive review, it was found that there are potential obstacles in the construction of PRSs that could affect how they perform in real world population studies. As it was mentioned in the review, all the studies from the last decade have concluded that clinical risk factors

perform quite well in predicting T2D, and there is almost no improvement when adding the PRS; therefore, the PRSs do not have relevant value in the prediction, challenging their clinical relevancy. Further work is still needed to be done to achieve a complete understanding of how PRS is a functional tool for the diagnosis of T2D.

The first obstacle to overcome is the lack of innovation in the generation of PRSs for T2D. Right now, the central target of developing a PRS is to have a correct prediction to recognize individuals at risk [98,99]. The use of SNPs and logistic regression when making a PRS could be improved, since logistic regression is made to understand the process but is not optimized for prediction [100]. There are two approaches to build a PRS model, namely regression-based methods (e.g., logistic regression) and tree-based methods (e.g., random forest) [101,102]. Regression-based methods employ polynomial parametric or non-parametric regression methods to make a relation of the input to the output data. Tree-based methods use the binary split rule to have correlation between the input and output data [103–105]. Tree-based methods using machine learning approaches have been extensively used in risk prediction for diseases such as cancer, Alzheimer's disease, and cardiovascular disease [106–109]. The use of machine learning techniques, combined with data from GWAS will improve the prediction of polygenic traits [110].

Secondly, the underestimation of population heterogeneity in the prediction of T2D could be the cause of another problem, namely overfitting, which is a common concern in PRS studies. It is required to calibrate, validate and optimize the PRS to every cohort of study to prove that it does not overfit the training data, producing inflated results [18]. The point of reference to avoid the generation of overfit prediction models is to implement a prediction using out of sample data [92,111]. The majority of PRS using Caucasian GWAS are biased by the allelic drift when compared to other populations, even when picking the same variants [25,112]. Diverse ethnic groups have different frequencies of key risk and probably different SNP associations [36]. Therefore, there is a need for developing generalized risk prediction methods and the inclusion of more diverse individuals in risk score studies [25]. To avoid the possibility of obtaining false positive results due to overfitting, the adaptation of existing T2D PRSs validated for Caucasian cohorts to other ethnicities could be the answer.

Thirdly, the environmental effect in genetic studies could be a bias in the development of a T2D PRS, an important aspect to take into consideration. Gene–environment (GxE) interactions can be defined as “genetic effects on a disease that differ in magnitude across environmental contexts” [113]. In most GWAS studies, it is assumed that no GxE interactions exist. If this assumption is incorrect, then the clinical effects of genetic risk factors may be missed [114]. Genetic and environmental factors may jointly contribute to clarify the importance of analyzing GxE interactions [113,114] and the benefits that these interactions have, such as the detection of new disease susceptibility loci [115–117]. The high power of PRS approaches to identify GxE interactions has been confirmed [118,119].

A genetic background seems to be fundamental for the development of diabetes, but it is only absolutely enough in Mendelian forms of diabetes, such as monogenic diabetes [120]. The identification of genetic variants using GWAS explain only about 10% of T2D heritability. Studies have assessed the importance of heredity and environment on the etiology of T2D; therefore, the missing heritability may be attributed to GxE interaction [34,121]. Documentation that environmental factors adjust phenotypic expression in genetic risk cohorts has been demonstrated in individuals with glucose intolerance using a multiethnic cohort [87,120], among others studies [122]. Because GxE interactions have proposed as a way to improve genetic risk studies, these environmental factors are relevant to the diagnosis of T2D [123]. The interactions of GxE contribute to the total genetic variance of a given trait [124], demonstrating the importance of GxE interactions in explaining the variance of diabetes-related studies [125]. To discard a bias of the environment in a genetic study, it will be necessary to obtain a similar PRS for T2D prediction when fitting the model on a single homogeneous population but exposed to different environments. In cases where the PRS will be different for the same case exposed before, the neglect of the environment effect will be the reason.

## 5. Conclusions

We have identified 15 studies that developed PRS, 12 to discriminate between patients and controls, and three to discriminate between T1D and diabetes subtypes. We consistently assessed the accuracy of PRSs using the AUC, regardless of the source of data, panel of genes used and genotyping strategies. However, these findings should be taken with caution. These PRSs were identified from 15 studies with variable study designs. In order to have a better prediction of diabetes, the use of PRSs that combine clinical, environmental and genetic interactions must be used. It is necessary to develop strategies to establish the clinical validity of PRSs by creating a pipeline for the interpretation of findings and their translation into actual evidence. Taking into account all the factors for implementation is the way to demonstrate the utility of PRSs in medical practice.

**Author Contributions:** Conceptualization, F.P.-M., F.C., M.K. and A.K.; methodology, F.P.-M.; formal analysis, F.P.-M.; writing—review and editing, F.P.-M., F.C., M.K. and A.K.; visualization, F.P.-M.; supervision, F.C., M.K. and A.K.; project administration, F.P.-M.; funding acquisition, A.K. All authors have read and agree to the published version of the manuscript.

**Funding:** This research was conducted within the project that 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.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

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. [CrossRef]
2. WHO. Available online: <https://www.who.int/news-room/fact-sheets/detail/diabetes> (accessed on 30 August 2019).
3. Bluestone, J.A.; Herold, K.; Eisenbarth, G. Genetics, Pathogenesis and Clinical Interventions in Type 1 Diabetes. *Nature* **2010**, *464*, 1293–1300.
4. Kahn, S.E.; Mark, E.C.; Stefano, D.P. Pathophysiology and Treatment of Type 2 Diabetes: Perspectives on the Past, Present, and Future. *Lancet* **2014**, *383*, 1068–1083. [CrossRef]
5. 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*, 29. [CrossRef]
6. 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. [CrossRef]
7. 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. [CrossRef]
8. Gloyn, A.L.; Ellard, S.; 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. [CrossRef]
9. Stride, A.H.A. Different Genes, Different Diabetes: Lessons from Maturity Onset Diabetes of the Young. *Ann. Med.* **2002**, *34*, 207–216. [CrossRef]
10. Mayer-Davis, E.J.; Lawrence, J.M.; Dabelea, D.; Divers, J.; Isom, S.; Dolan, L.; Imperatore, G.; Linder, B.; Marcovina, S.; Pettitt, D.J.; et al. Incidence Trends of Type 1 and Type 2 Diabetes among Youths, 2002–2012. *N. Engl. J. Med.* **2017**, *376*, 1419–1429. [CrossRef]
11. Karvonen, M.; Viik-Kajander, M.; Moltchanova, E.; Libman, I.; LaPorte, R.; Tuomilehto, J. Incidence of Childhood Type 1 Diabetes Worldwide. Diabetes Mondiale (DiaMond) Project Group. *Diabetes Care* **2000**, *23*, 1516–1526. [CrossRef]

12. Rawshani, A.; Rawshani, A.; Franzén, S.; Eliasson, B.; Svensson, A.-M.; Miftaraj, M.; McGuire, D.K.; Sattar, N.; Rosengren, A.; Gudbjörnsdóttir, S. Mortality and Cardiovascular Disease in Type 1 and Type 2 Diabetes. *N. Engl. J. Med.* **2017**, *376*, 1407–1418. [[CrossRef](#)] [[PubMed](#)]
13. Pan, C.; Shang, S.; Kirch, W.; Thoenes, M. Burden of Diabetes in the Adult Chinese Population: A Systematic Literature Review and Future Projections. *Int. J. Gen. Med.* **2010**, *3*, 173–179. [[PubMed](#)]
14. Abubakari, A.R.; Lauder, W.; Jones, M.C.; Kirk, A.; Agyemang, C.; Bhopal, R.S. Prevalence and Time Trends in Diabetes and Physical Inactivity among Adult West African Populations: The Epidemic Has Arrived. *Public Health* **2009**, *123*, 602–614. [[CrossRef](#)] [[PubMed](#)]
15. De Lusignan, S.; Khunti, K.; Belsey, J.; Hattersley, A.; van Vlymen, J.; Gallagher, H.; Millett, C.; Hague, N.J.; Tomson, C.; Harris, K.; et al. A Method of Identifying and Correcting Miscoding, Misclassification and Misdiagnosis in Diabetes: A Pilot and Validation Study of Routinely Collected Data. *Diabet. Med.* **2010**, *27*, 203–209. [[CrossRef](#)]
16. Seidu, S.; Davies, M.J.; Mostafa, S.; de Lusignan, S.; Khunti, K. Prevalence and Characteristics in Coding, Classification and Diagnosis of Diabetes in Primary Care. *Postgrad. Med. J.* **2013**, *90*, 13–17. [[CrossRef](#)]
17. 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. [[CrossRef](#)]
18. Vega, F.M.D.L.; Bustamante, C.D. Polygenic Risk Scores: A Biased Prediction? *Genome Med.* **2018**, *10*, 1–3.
19. 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. [[CrossRef](#)]
20. 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. [[CrossRef](#)]
21. Evangelou, E.; Warren, 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. [[CrossRef](#)]
22. Whitfield, J.B. Genetic Insights and into Cardiometabolic and Risk Factors. *Clin. Biochem. Rev.* **2014**, *35*, 15. [[PubMed](#)]
23. Zeng, P. Statistical Analysis for Genome-Wide Association Study. *J. Biomed. Res.* **2015**, *29*, 285. [[PubMed](#)]
24. Golan, D.; Lander, E.S.; Rosset, S. Measuring Missing Heritability: Inferring the Contribution of Common Variants. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E5272–E5281. [[CrossRef](#)] [[PubMed](#)]
25. Kolb, H.; Martin, S. Environmental/Lifestyle Factors in the Pathogenesis and Prevention of Type 2 Diabetes. *BMC Med.* **2017**, *15*, 131. [[CrossRef](#)]
26. Sanna, S.; van Zuydam, N.R.; Mahajan, A.; Kurilshikov, A.; Vila, A.V.; Vösa, 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. [[CrossRef](#)]
27. Mahajan, A.; Taliun, D.; Thurner, M.; Robertson, N.R.; Torres, J.M.; Rayner, N.W.; Payne, A.J.; Steinthorsdóttir, 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. [[CrossRef](#)]
28. Onengut-Gumuscu, S.; Chen, 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. [[CrossRef](#)]
29. Sharp, S.A.; 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. [[CrossRef](#)]
30. 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. [[CrossRef](#)]
31. 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. [[CrossRef](#)]

32. 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. [[CrossRef](#)] [[PubMed](#)]
33. Almgren, P. Heritability and Familiality of Type 2 Diabetes and Related Quantitative Traits in the Botnia Study. *Diabetologia* **2011**, *54*, 2811–2819. [[CrossRef](#)] [[PubMed](#)]
34. 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. [[CrossRef](#)] [[PubMed](#)]
35. Bailey, J.N.C.; Igo, R.P. Genetic Risk Scores. *Curr. Protoc. Hum. Genet.* **2016**, *91*, 1291–1299. [[CrossRef](#)]
36. 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. [[CrossRef](#)]
37. Wray, N.; Yang, J.; Goddard, M.E.; Visscher, P.M. The Genetic Interpretation of Area under the ROC Curve in Genomic Profiling. *PLoS Genet.* **2010**, *6*, e1000864. [[CrossRef](#)]
38. 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. [[CrossRef](#)]
39. 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. [[CrossRef](#)]
40. Kim, M.S.; Patel, K.P.; Teng, A.K.; Berens, A.J.; Lachance, J. Genetic Disease Risks Can Be Misestimated across Global Populations. *Genome Biol.* **2018**, *19*, 1–14. [[CrossRef](#)]
41. Thomas, N.J.; Jones, S.E.; Weedon, M.N.; Shields, B.M.; Oram, R.A.; Hattersley, A.T. Frequency and Phenotype of Type 1 Diabetes in the First Six Decades of Life: A Cross-Sectional, Genetically Stratified Survival Analysis from UK Biobank. *Lancet Diabetes Endocrinol.* **2018**, *6*, 122–129. [[CrossRef](#)]
42. 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. [[CrossRef](#)] [[PubMed](#)]
43. 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*, 1–13. [[CrossRef](#)] [[PubMed](#)]
44. Sharp, S.A.; 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.
45. 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**, *36*, 1694–1702. [[CrossRef](#)] [[PubMed](#)]
46. 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. [[CrossRef](#)]
47. Lango, H.; Palmer, C.N.A.; Morris, A.D.; Zeggini, E.; Hattersley, A.T.; McCarthy, M.I.; Frayling, T.M.; Weedon, M.N. Assessing the Combined Impact of 18 Common Genetic Variants of Modest Effect Sizes on Type 2 Diabetes Risk. *Diabetes* **2008**, *57*, 3129–3135. [[CrossRef](#)]
48. 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. [[CrossRef](#)]
49. 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’Agostino, 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. [[CrossRef](#)]

50. 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. [[CrossRef](#)]
51. 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. [[CrossRef](#)]
52. Bramer, W.M.; Rethlefsen, M.L.; Kleijnen, J.; Franco, O.H. Optimal Database Combinations for Literature Searches in Systematic Reviews: A Prospective Exploratory Study. *Syst. Rev.* **2017**, *6*, 245. [[CrossRef](#)] [[PubMed](#)]
53. Little, J.; Higgins, J.P.T.; Ioannidis, J.P.A.; Moher, D.; Gagnon, F.; von Elm, E.; Khoury, M.J.; Cohen, B.; Davey-Smith, G.; Grimshaw, J.; et al. Strengthening the REporting of Genetic Association Studies (STREGA)-an Extension of the STROBE Statement. *Genet. Epidemiol.* **2009**, *33*, 581–598. [[CrossRef](#)] [[PubMed](#)]
54. Chikowore, T.; van Zyl, T.; Feskens, E.J.M.; Conradie, K.R. Predictive Utility of a Genetic Risk Score of Common Variants Associated with Type 2 Diabetes in a Black South African Population. *Diabetes Res. Clin. Pract.* **2016**, *122*, 1–8. [[CrossRef](#)] [[PubMed](#)]
55. 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. [[CrossRef](#)] [[PubMed](#)]
56. Devuyst, O. The 1000 Genomes Project: Welcome to a New World. *Perit. Dial. Int.* **2015**, *35*, 676–677. [[CrossRef](#)] [[PubMed](#)]
57. Chikowore, T.; Conradie, K.R.; Towers, G.W.; van Zyl, T. Common Variants Associated with Type 2 Diabetes in a Black South African Population of Setswana Descent: African Populations Diverge. *OMICS: A J. Integr. Biol.* **2015**, *19*, 617–626. [[CrossRef](#)]
58. 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. [[CrossRef](#)]
59. 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. [[CrossRef](#)]
60. 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. [[CrossRef](#)]
61. 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. [[CrossRef](#)]
62. Voight, B.F.; Scott, 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. [[CrossRef](#)] [[PubMed](#)]
63. 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. [[CrossRef](#)]
64. 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. [[CrossRef](#)] [[PubMed](#)]
65. 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. [[CrossRef](#)]
66. Frayling, T.M. Genome-Wide Association Studies Provide New Insights into Type 2 Diabetes Aetiology. *Nat. Rev. Genet.* **2007**, *8*, 657–662. [[CrossRef](#)]
67. 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. [[CrossRef](#)]
68. 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. [[CrossRef](#)]
69. 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. [[CrossRef](#)]
70. 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.
71. 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. [[CrossRef](#)]
72. 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. [[CrossRef](#)] [[PubMed](#)]
73. Steinthorsdottir, V.; Thorleifsson, G.; Reynisdottir, I.; Benediktsson, R.; Jonsdottir, T.; Walters, G.B.; Styrkarsdottir, U.; Gretarsdottir, S.; Emilsson, V.; Ghosh, S.; et al. A Variant in CDKAL1 Influences Insulin Response and Risk of Type 2 Diabetes. *Nat. Genet.* **2007**, *39*, 770–775. [[CrossRef](#)] [[PubMed](#)]
74. Zeggini, E.; Weedon, M.N.; Lindgren, C.M.; Frayling, T.M.; Elliott, K.S.; Lango, H.; Timpson, N.J.; Perry, J.R.B.; Rayner, N.W.; Freathy, R.M.; et al. Replication of Genome-Wide Association Signals in UK Samples Reveals Risk Loci for Type 2 Diabetes. *Science* **2007**, *316*, 1336–1341. [[CrossRef](#)] [[PubMed](#)]
75. Lyssenko, V.; Almgren, P.; Anevski, D.; Orho-Melander, M.; Sjögren, M.; Saloranta, C.; Tuomi, T.; Groop, L.; the Botnia Study Group. Genetic Prediction of Future Type 2 Diabetes. *PLoS Med.* **2005**, *2*, e345. [[CrossRef](#)]
76. 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. [[CrossRef](#)]
77. 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. [[CrossRef](#)]
78. 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**, *10*, 1447–1455. [[CrossRef](#)]
79. Gale, E.A.M. The Rise of Childhood Type 1 Diabetes in the 20th Century. *Diabetes* **2014**, *51*, 3353–3361. [[CrossRef](#)]
80. Noble, J.A. Immunogenetics of Type 1 Diabetes: A Comprehensive Review. *J. Autoimmun.* **2015**, *64*, 101–112. [[CrossRef](#)]
81. Roshandel, D.; Gubitosi-Klug, R.; Bull, S.B.; Canty, A.J.; Pezzolesi, M.G.; King, G.L.; Keenan, H.A.; Snell-Bergeon, J.K.; Maahs, D.M.; Klein, R.; et al. Meta-Genome-Wide Association Studies Identify a Locus on Chromosome 1 and Multiple Variants in the MHC Region for Serum C-Peptide in Type 1 Diabetes. *Diabetologia* **2018**, *61*, 1098–1111. [[CrossRef](#)]
82. Baschal, E.E.; Baker, P.R.; Eyring, K.R.; Siebert, J.C.; Jasinski, J.M.; Eisenbarth, G.S. The HLA-B\*3906 Allele Imparts a High Risk of Diabetes Only on Specific HLA-DR/DQ Haplotypes. *Diabetologia* **2011**, *54*, 1702–1709. [[CrossRef](#)] [[PubMed](#)]
83. Noble, J.A.; Valdes, A.M.; Varney, M.D.; Carlson, J.A.; Moonsamy, P.; Fear, A.L.; Lane, J.A.; Lavant, E.; Rappner, R.; Louey, A.; et al. HLA Class I and Genetic Susceptibility to Type 1 Diabetes: Results From the Type 1 Diabetes Genetics Consortium. *Diabetes* **2010**, *59*, 2972–2979. [[CrossRef](#)] [[PubMed](#)]
84. Barrett, J.C.; Clayton, 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. [[CrossRef](#)] [[PubMed](#)]
85. Pociot, F. Type 1 Diabetes Genome-Wide Association Studies: Not to Be Lost in Translation. *Clin. Transl. Immunol.* **2017**, *6*, e162. [[CrossRef](#)]

86. Aly, T.A.; Ide, A.; Jahromi, M.M.; Barker, J.M.; Fernando, M.S.; Babu, S.R.; Yu, L.; Miao, D.; Erlich, H.A.; Fain, P.R.; et al. Extreme Genetic Risk for Type 1A Diabetes. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 14074–14079. [[CrossRef](#)]
87. Knowler, W. Reduction in the Incidence of Type 2 Diabetes with Lifestyle Intervention or Metformin. *N. Engl. J. Med.* **2002**, *346*, 393–403.
88. Lindstrom, J.; Louheranta, A.; Mannelin, M.; Rastas, M.; Salminen, V.; Eriksson, J.; Uusitupa, M.; Tuomilehto, J. The Finnish Diabetes Prevention Study (DPS): Lifestyle Intervention and 3-Year Results on Diet and Physical Activity. *Diabetes Care* **2003**, *26*, 3230–3236. [[CrossRef](#)]
89. Morris, A. Large-Scale Association Analysis Provides Insights into the Genetic Architecture and Pathophysiology of Type 2 Diabetes. *Nat. Genet.* **2012**, *44*, 981–990.
90. 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. [[CrossRef](#)]
91. 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. [[CrossRef](#)]
92. 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. [[CrossRef](#)] [[PubMed](#)]
93. Lyssenko, V.; Laakso, M. Genetic Screening for the Risk of Type 2 Diabetes: Worthless or Valuable? *Diabetes Care* **2013**, *36*, S120–S126. [[CrossRef](#)] [[PubMed](#)]
94. Hope, S.V.; Wienand-Barnett, S.; Shepherd, M.; King, S.M.; Fox, C.; Khunti, K.; Oram, R.A.; Knight, B.A.; Hattersley, A.T.; Jones, A.G.; et al. Practical Classification Guidelines for Diabetes in Patients Treated with Insulin: A Cross-Sectional Study of the Accuracy of Diabetes Diagnosis. *Br. J. Gen. Pract.* **2016**, *66*, e315–e322. [[CrossRef](#)] [[PubMed](#)]
95. Pearson, E.R.; Starkey, B.J.; Powell, R.J.; Gribble, F.M.; Clark, P.M.; Hattersley, A.T. Genetic Cause of Hyperglycaemia and Response to Treatment in Diabetes. *Lancet* **2003**, *362*, 1275–1281. [[CrossRef](#)]
96. Pearson, E.R.; Flechtner, I.; Njølstad, P.R.; Malecki, M.T.; Flanagan, S.E.; Larkin, B.; Ashcroft, F.M.; Klimes, I.; Codner, E.; Iotova, V.; et al. Switching from Insulin to Oral Sulfonylureas in Patients with Diabetes Due to Kir6.2 Mutations. *N. Engl. J. Med.* **2006**, *355*, 467–477. [[CrossRef](#)]
97. 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. [[CrossRef](#)]
98. 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. *Lancet* **2010**, *375*, 1525–1535. [[CrossRef](#)]
99. Manolio, T.A. Bringing Genome-Wide Association Findings into Clinical Use. *Nat. Rev. Genet.* **2013**, *14*, 549–558. [[CrossRef](#)]
100. Shmueli, G. To Explain or To Predict? *SSRN Electron. J.* **2010**, *25*, 289–310.
101. 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. [[CrossRef](#)]
102. Abraham, G.; Inouye, M. Genomic Risk Prediction of Complex Human Disease and Its Clinical Application. *Curr. Opin. Genet. Dev.* **2015**, *33*, 10–16. [[CrossRef](#)] [[PubMed](#)]
103. 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. [[CrossRef](#)] [[PubMed](#)]
104. 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. [[CrossRef](#)] [[PubMed](#)]
105. 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. [[CrossRef](#)]
106. Cruz, J.A.; Wishart, D.S. Applications of Machine Learning in Cancer Prediction and Prognosis. *Cancer Inform.* **2006**, *2*, 117693510600200. [[CrossRef](#)]



107. 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, Doha, Qatar, 31 March–4 April 2008.
108. 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*, 16. [[CrossRef](#)]
109. Zhang, D.; Shen, D. Multi-Modal Multi-Task Learning for Joint Prediction of Multiple Regression and Classification Variables in Alzheimer's Disease. *NeuroImage* **2012**, *59*, 895–907. [[CrossRef](#)]
110. Paré, G.; Mao, S.; Deng, W.Q. A Machine-Learning Heuristic to Improve Gene Score Prediction of Polygenic Traits. *Sci. Rep.* **2017**, *7*, 12665. [[CrossRef](#)]
111. Choi, S.W.; Mak, T.S.H.; O'Reilly, P.F. A Guide to Performing Polygenic Risk Score Analyses. *BioRxiv* **2018**. [[CrossRef](#)]
112. 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. [[CrossRef](#)]
113. Dick, D.M. Gene-Environment Interaction in Psychological Traits and Disorders. *Annu. Rev. Clin. Psychol.* **2011**, *7*, 383–409. [[CrossRef](#)] [[PubMed](#)]
114. Franks, P.W. Gene × Environment Interactions in Type 2 Diabetes. *Curr. Diabetes Rep.* **2011**, *11*, 552–561. [[CrossRef](#)] [[PubMed](#)]
115. Thomas, D. Gene-Environment-Wide Association Studies: Emerging Approaches. *Nat. Rev. Genet.* **2010**, *11*, 259–272. [[CrossRef](#)] [[PubMed](#)]
116. Boffetta, P.; Winn, D.M.; Ioannidis, J.P.; Thomas, D.C.; Little, J.; Smith, G.D.; Coglianò, 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. [[CrossRef](#)]
117. 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. [[CrossRef](#)]
118. 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. [[CrossRef](#)]
119. Aschard, H. A Perspective on Interaction Effects in Genetic Association Studies. *Genet. Epidemiol.* **2016**, *40*, 678–688. [[CrossRef](#)]
120. 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. [[CrossRef](#)]
121. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorf, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A.; et al. Finding the Missing Heritability of Complex Diseases. *Nature* **2009**, *461*, 747–753. [[CrossRef](#)]
122. Eze, I.C.; Imboden, M.; Kumar, A.; von Eckardstein, A.; Stolz, D.; Gerbase, M.W.; Künzli, N.; Pons, M.; Kronenberg, F.; Schindler, C.; et al. Air Pollution and Diabetes Association: Modification by Type 2 Diabetes Genetic Risk Score. *Environ. Int.* **2016**, *94*, 263–271. [[CrossRef](#)]
123. Lee, Y.-C.; Lai, C.-Q.; Ordovas, J.M.; Parnell, L.D. A Database of Gene-Environment Interactions Pertaining to Blood Lipid Traits, Cardiovascular Disease and Type 2 Diabetes. *J. Data Min. Genom. Proteom.* **2011**, *2*. [[CrossRef](#)] [[PubMed](#)]
124. Visscher, P.M.; Hill, W.G.; Wray, N.R. Heritability in the Genomics Era—Concepts and Misconceptions. *Nat. Rev. Genet.* **2008**, *9*, 255–266. [[CrossRef](#)] [[PubMed](#)]
125. Zheng, J.-S.; Arnett, D.K.; Lee, Y.-C.; Shen, J.; Parnell, L.D.; Smith, C.E.; Richardson, K.; Li, D.; Borecki, I.B.; Ordovas, J.M.; et al. Genome-Wide Contribution of Genotype by Environment Interaction to Variation of Diabetes-Related Traits. *PLoS ONE* **2013**, *8*, e77442. [[CrossRef](#)] [[PubMed](#)]

