The effect of type 2 diabetes genetic predisposition on non-cardiovascular comorbidities

Ana Luiza Arruda[1,2,*,×], Ozvan Bocher[1,3,*], Henry J. Taylor[4,5,6,*], Davis Cammann[7,*], Satoshi Yoshiji[8,9,10], Xianyong Yin[11,12], Chi Zhao[13], Jingchun Chen[7], Alexis C. Wood[14], Ken Suzuki[15], Josep M. Mercader[8,16,17], Cassandra N. Spracklen[13], James B. Meigs[8,18,19], Marijana Vujkovic[20,21,22], George Davey Smith[23], Jerome I. Rotter[24], Benjamin F. Voight[20,21,25,26,‡], Andrew P. Morris[27,‡], Eleftheria Zeggini[1,28,‡,×]

[1]*Institute of Translational Genomics, Helmholtz Munich, Neuherberg, 85764, Germany*
[2]*Technical University of Munich (TUM), School of Medicine and Health, Graduate School of Experimental Medicine, Munich, 81675, Germany*
[3]*Univ Brest, Inserm, EFS, UMR 1078, GGB, F-29200 Brest, France*
[4]*Center for Precision Health Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA*
[5]*British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK*
[6]*Heart and Lung Research Institute, University of Cambridge, Cambridge, UK*
[7]*Nevada Institute of Personalized Medicine, University of Nevada, Las Vegas, 4505 S. Maryland Parkway, Las Vegas, NV 89154, USA*
[8]*Programs in Metabolism and Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA*
[9]*Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan*
[10]*McGill Genome Centre, McGill University, Montreal, QC, Canada*
[11]*Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing, China*
[12]*Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA*
[13]*Department of Biostatistics and Epidemiology, University of Massachusetts Amherst, Amherst, MA, USA*
[14]*USDA/ARS Children's Nutrition Center, Baylor College of Medicine, Houston, TX, USA*
[15]*Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, University of Tokyo, Tokyo, Japan*
[16]*Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA*
[17]*Harvard Medical School, Boston, MA, USA*
[18]*Department of Medicine, Harvard Medical School, Boston, MA, USA*
[19]*Division of General Internal Medicine, Massachusetts General Hospital, Boston, MA, USA*
[20]*Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, USA*
[21]*Department of Genetics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA*
[22]*Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA*
[23]*MRC Integrative Epidemiology Unit, University of Bristol, BS8 2BN Bristol, UK*
[24]*Institute for Translational Genomics and Population Sciences, Department of Pediatrics, Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA, USA*
[25]*Department of Systems Pharmacology and Translational Therapeutics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA*
[26]*Institute for Translational Medicine and Therapeutics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA*
[27]*Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, The University of Manchester, Manchester, United Kingdom*

[28]*TUM School of Medicine and Health, Technical University of Munich and Klinikum Rechts der Isar, Munich, 81675, Germany*
*equal contribution
‡shared last authors
×corresponding authors

Abstract

Type 2 diabetes (T2D) is epidemiologically associated with a wide range of non-cardiovascular comorbidities, yet their shared etiology has not been fully elucidated. Leveraging eight non-overlapping mechanistic clusters of T2D genetic profiles, each representing distinct biological pathways, we investigate putative causal links between cluster-stratified T2D genetic predisposition and 21 non-cardiovascular comorbidities. Most of the identified putative causal effects are driven by distinct T2D genetic clusters. For example, the risk-increasing effects of T2D genetic predisposition on cataracts and erectile dysfunction are primarily attributed to obesity and glucose regulation mechanisms, respectively. When surveyed in populations across the globe, we observe opposing effect directions for depression, asthma and chronic obstructive pulmonary disease between populations. We identify a putative causal link between T2D genetic predisposition and osteoarthritis. To underscore the translational potential of our findings, we intersect high-confidence effector genes for osteoarthritis with targets of T2D-approved drugs and identify metformin as a potential candidate for drug repurposing in osteoarthritis.

Introduction

Worldwide, over 500 million individuals have type 2 diabetes (T2D), a number expected to exceed 1.27 billion by 2050[1]. In 2022, diabetes-related healthcare expenditures were estimated to be over $412 billion in the United States of America alone[2]. The cumulative economic strain on healthcare resources is much greater when considering the comorbidities associated with T2D. For instance, a study looking at multimorbid T2D patients found that 73% also suffered from hypertension, followed by 69% with a diagnosis of back pain, 67% of depression, 45% of asthma, and 36% of osteoarthritis[3]. While advances in treatment, such as improved blood sugar monitoring[4] and glucagon-like peptide-1 (GLP-1) receptor agonists[5], can mitigate some of the challenges associated with T2D, its rising prevalence and diverse comorbidities call for a paradigm shift toward prevention strategies. Despite their prevalence and significant impact on quality of life[6], non-cardiovascular T2D comorbidities remain understudied compared to cardiovascular conditions. Building on the understanding of shared etiology, we can identify common biological mechanisms to develop targeted interventions tailored to an individual's genetic predisposition to multimorbidity profiles[7, 8].

Genetic and environmental factors are involved in the pathophysiology of T2D, a biologically heterogeneous disease characterized by the interplay of different cell-type specific mechanisms[9]. Efforts to identify different T2D genetic profiles have yielded promising results[10, 11]. The T2D Global Genomics Initiative (T2DGGI) has identified 1,289 genetic risk variants in a multi-ancestry genome-wide association study (GWAS) meta-analysis [11] and grouped these variants into eight non-overlapping mechanistic clusters representing distinct biological pathways. Clustering was performed based on the association profile of the T2D genetic risk variants with cardiometabolic traits, yielding the following clusters: body fat (n=233 variants), obesity (n=233), beta cell associated with increased proinsulin (PI) levels (n=91), beta cell associated with decreased PI levels (n=89), metabolic syndrome (n=166), lipodystrophy (n=45), liver/lipid metabolism (n=3) and residual glycaemic (n=389).

Associations between the T2D partitioned genetic risk scores (GRS) and cardiovascular outcomes, including coronary artery disease and diabetic retinopathy, have now been described[10, 11]. However, previous studies investigating the causal links between genetic predisposition to T2D and its non-cardiovascular comorbidities have not considered genetic subtypes of disease heterogeneity[12-20]. Here, we apply a cluster-stratified causal inference analysis framework using Mendelian randomization (MR) to 21 diseases epidemiologically associated with T2D, spanning five classes: musculoskeletal, respiratory, reproductive, neuropsychiatric, and ophthalmic conditions. We identify causal links between the genetic burden of T2D mechanistic subtypes and 15 commonly co-occurring diseases and highlight repurposing opportunities for T2D-approved drugs. Our findings demonstrate the potential of translational genomics to identify targeted preventive strategies.

Results

Overview of the conceptual framework of the study

We investigate potential causal effects between T2D genetic predisposition and 21 non-cardiovascular comorbidities (SupTab1) using MR. The T2D comorbidities were selected after a comprehensive literature review of epidemiological studies and the availability of GWAS data. In addition to performing pairwise bi-directional two-sample MR analyses (Figure 1), we investigated cluster-stratified effects of T2D genetic predisposition on the comorbidities by restricting the genetic instrumental variables (IVs) to T2D genetic risk variants assigned to each mechanistic cluster described in the latest T2DGGI GWAS meta-analysis (Figure 1). We performed all analyses within each genetic ancestry group and meta-analyzed the single ancestry estimates where possible. Although genetic ancestry is on a continuous scale, here we use the continental genetic ancestry groups defined by the 1000 Genomes Project phase 3[21]. We corrected for multiple testing using false-discovery-rate (FDR) across all MR analyses and defined relationships showing an FDR-adjusted p-value (q-value) < 0.05 as putative causal. We applied multiple sensitivity analyses to assess the validity of the MR assumptions, including adjusting the estimates for potential confounders or mediators of the identified potentially causal relationships through multivariable MR (Methods, Supplemental Note, Figure 1). Causal estimates are expressed as the odds ratio (OR) for each comorbidity per 2-fold increase in genetically determined dichotomous T2D risk. Finally, to validate the MR findings, we performed a phenome-wide association study (PheWAS) for each T2D genetic cluster using data from the All of Us program, an ancestrally and culturally diverse cohort that was not included in the T2DGGI GWAS meta-analysis[22]. Similarly to the MR analysis, we meta-analyzed the PheWAS results across genetic ancestry groups (Methods).

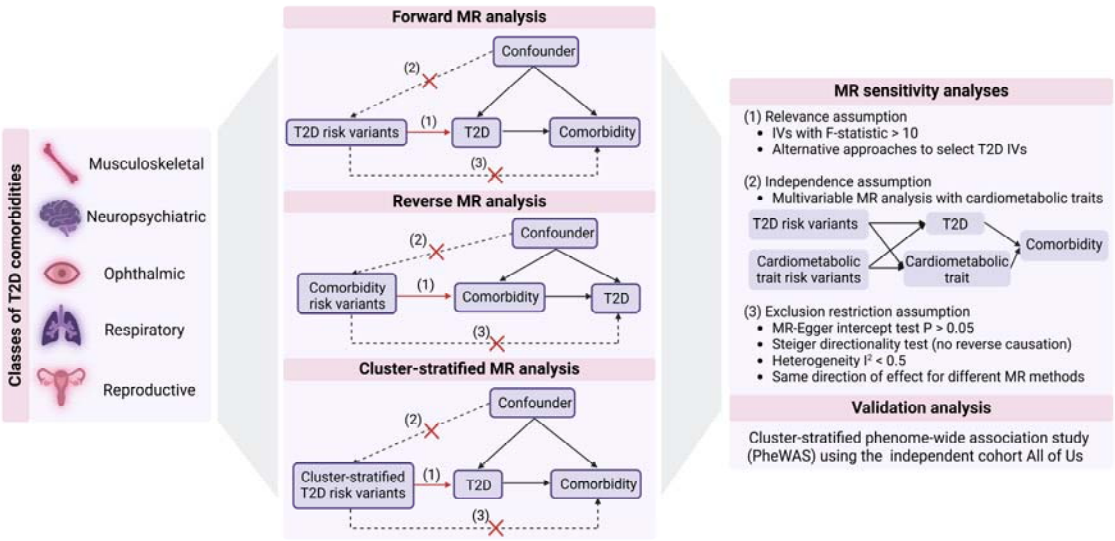


Figure 1: Overview of study design (MR = Mendelian randomization; T2D = type 2 diabetes; IVs = genetic instrumental variables; P = p-value).

Genetic predisposition to T2D is a potential driver of comorbidities

We find statistical evidence (q-value < 0.05) of a putative causal link between overall T2D genetic predisposition and 10 comorbidities (Figure 2, SupTab2). Our findings show that genetic predisposition to T2D has risk-increasing effects on carpal tunnel syndrome (CTS) (OR=1.10, q-value=3.12×10^{-28}), osteoarthritis (OR=1.01, q-value=4.43×10^{-2}), attention-deficit/hyperactivity disorder (ADHD) (OR=1.09, q-value=5.26×10^{-14}), primary open-angle glaucoma (glaucoma) (OR=1.07, q-value=6.23×10^{-8}), cataracts (OR=1.04, q-value=5.82×10^{-9}), polycystic ovary syndrome (PCOS) (OR=1.12, q-value=1.43×10^{-5}), and erectile dysfunction (OR=1.09, q-value=7.95×10^{-8}). For anorexia nervosa (OR=0.96, q-value=1.72×10^{-2}), osteoporosis (OR=0.95, q-value=1.15×10^{-3}), and obsessive-compulsive disorder (OCD) (OR=0.90, q-value=2.57×10^{-3}), we find T2D genetic predisposition to have a protective effect. In the reverse direction, we find statistical evidence of a protective effect of genetic predisposition to bipolar disorder (OR=0.96, q-value=7.95×10^{-3}) and clinically diagnosed Alzheimer's disease (OR=0.98, q-value=4.33×10^{-2}) on risk for T2D (Supplemental Figure 1, SupTab3).

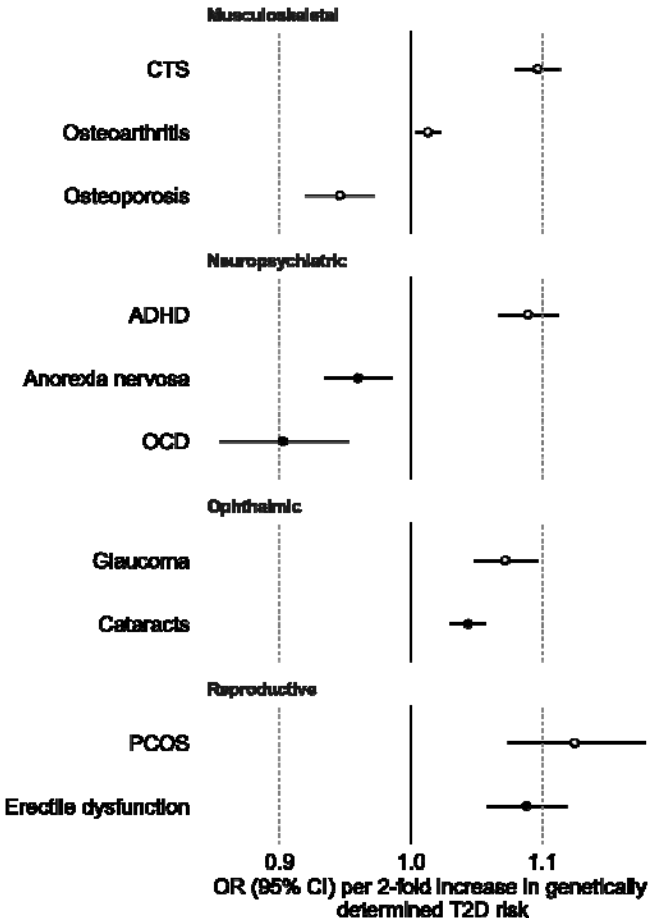


Figure 2: Results of two-sample Mendelian randomization (MR) analysis of genetic predisposition to type 2 diabetes (T2D) on non-cardiovascular comorbidity risk for the putative causal relationships (q-value < 0.05). Causal estimates are expressed as the odds ratio (OR) for each comorbidity per doubling (2-fold increase) in genetically determined dichotomous T2D risk. Filled circles mark estimates that passed the sensitivity analyses to assess the validity of the MR assumptions. (CI =

confidence interval; CTS = Carpal tunnel syndrome; ADHD = attention-deficit/hyperactivity disorder; OCD = obsessive-compulsive disorder; PCOS = polycystic ovary syndrome)

Distinct biological mechanisms contribute to the effect of T2D genetic predisposition on comorbidities

In addition to the effects of overall genetic predisposition to T2D on comorbidity risk, we investigated the effects of genetic burden for T2D mechanistic subtypes using the T2DGGI genetic clusters (Extended Figure 1). All identified putative causal relationships between overall genetic predisposition to T2D and its comorbidities, except for anorexia nervosa, are accompanied by at least one cluster-stratified effect, suggesting specific underlying biological pathways (Figure 3).

Within the power constraints of our study, we also provide evidence of putative causal effects of certain T2D genetic clusters without the presence of a causal effect of overall genetic predisposition to T2D on five further comorbidities: chronic back pain, rheumatoid arthritis, Alzheimer's disease, asthma, and chronic obstructive pulmonary disease (COPD) (Figure 3, SupTab2). To investigate whether the observed effects are specific to the respective clusters, we performed leave-one-cluster-out MR analyses using all T2D genetic risk variants except those assigned to the putative causal cluster (Methods, Supplemental Note, SupTab4). In the leave-one-cluster-out MR analyses, we find no evidence of a causal effect of T2D genetic predisposition on any of the five diseases, indicating specific biological mechanisms, represented by the putative causal clusters, that may underlie these relationships.

The impact of T2D obesity-related pathways on cataracts, asthma, and back pain

We find potential causal effects of the obesity cluster on 9 out of the 15 comorbidities (60%) causally affected by any genetic predisposition to T2D. The effects of the obesity cluster are consistently the largest across all clusters (Figure 3, SupTab2. For chronic back pain (OR=1.09, q-value=6.35×10^{-5}), cataracts (OR=1.08, q-value=3.84×10^{-6}), and asthma (OR=1.12, q-value=5.95×10^{-3}), we find a risk-increasing effect of the obesity cluster but not of any of the other T2D genetic clusters. Concordantly, we observe a positive association between the obesity cluster genetic risk variants and all three diseases in the independent All of Us cohort dataset (back pain: OR=1.03, q-value=1.52×10^{-5}; asthma: OR=1.04, q-value=7.9×10^{-7}; cataracts: OR=1.08, q-value=2.1×10^{-2}). Within the power bracket of our study, our findings indicate that obesity is the main biological mechanism driving the observed effect of T2D genetic predisposition on chronic back pain, cataracts, and asthma.

Body mass index (BMI) can act as a confounder of the identified relationships as it is a shared risk factor of T2D and the 9 comorbidities putatively causally affected by the obesity cluster (Supplemental Figure 2, Methods, SupTab5). When we adjust the obesity cluster estimates for BMI through multivariable MR, we no longer observe an effect of T2D genetic predisposition on any disease except cataracts, CTS, and osteoporosis (Supplemental Figure 28-33, SupTab6). Our sensitivity analyses show statistical evidence of a robust risk-increasing causal effect of T2D genetic predisposition on cataracts driven by the obesity genetic cluster.

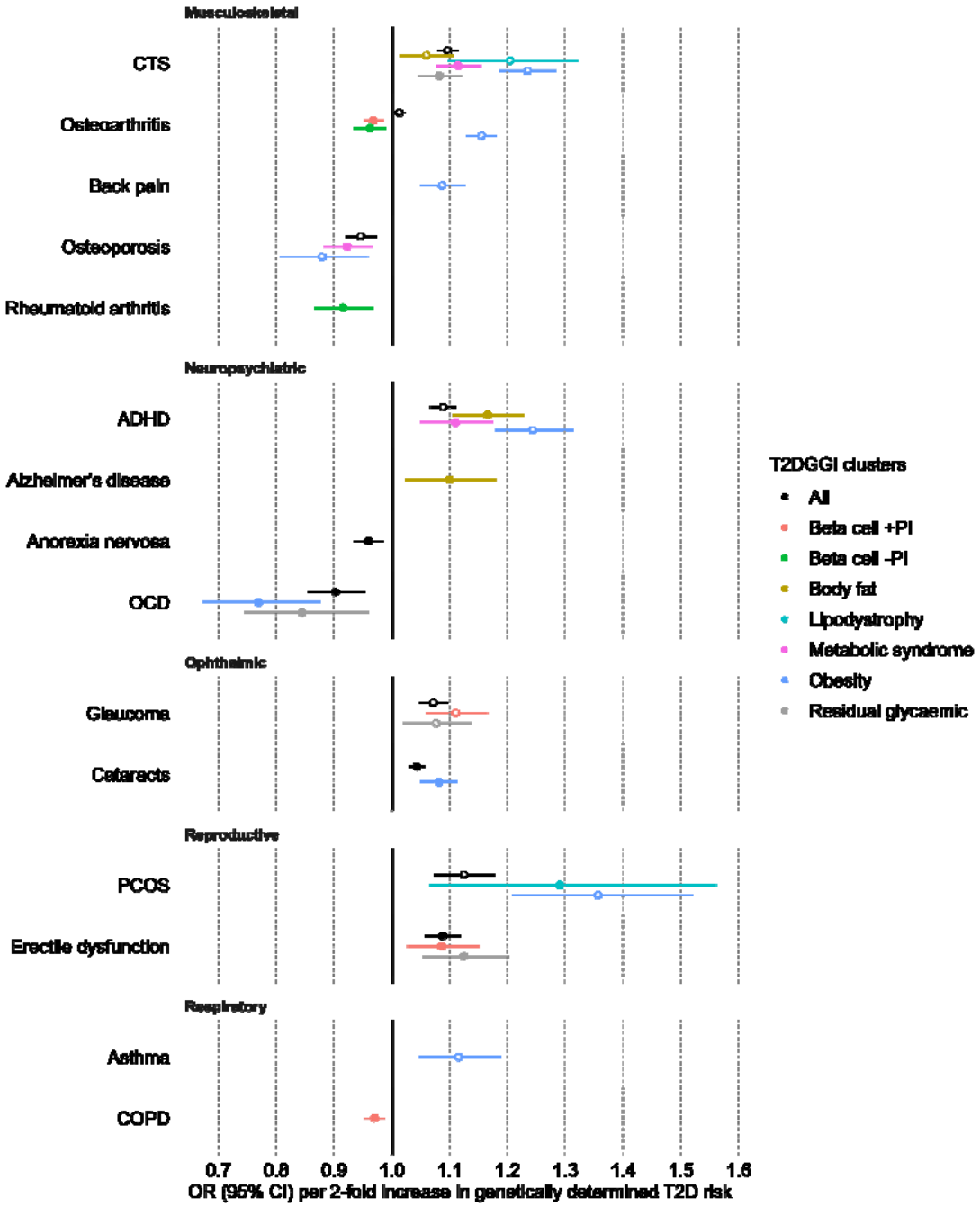


Figure 3: Results of cluster-stratified two-sample Mendelian randomization (MR) analysis of genetic predisposition to type 2 diabetes (T2D) on non-cardiovascular comorbidities risk for the putative causal relationships (q-value < 0.05). Causal estimates are expressed as the odds ratio (OR) for each comorbidity per doubling (2-fold increase) in genetically determined dichotomous T2D risk. Filled circles mark estimates that passed all sensitivity analyses to assess the validity of the MR assumptions. (T2DGGI = Type 2 Diabetes Global Genomics Initiative; CI = confidence interval; PI = proinsulin; CTS = Carpal tunnel syndrome; ADHD = attention-deficit/hyperactivity disorder; OCD = obsessive-compulsive disorder; PCOS = polycystic ovary syndrome; COPD = chronic obstructive pulmonary disease)

The impact of T2D glucose regulation mechanisms on glaucoma, COPD, rheumatoid arthritis, and erectile dysfunction

Beta cell-related clusters drive the putative protective effect of T2D genetic predisposition on rheumatoid arthritis (-PI: OR=0.92, q-value=1.35×10^{-2}) and COPD (+PI: OR=0.97, q-value=7.42×10^{-3}), adding genetic evidence to the potential role of beta cell dysfunction in these diseases[23-26]. In addition to the beta cell clusters, variants in the residual glycaemic cluster, which are most strongly associated with fasting glucose, also drive the risk-increasing effects of genetic predisposition to T2D on erectile dysfunction (residual glycaemic: OR=1.12, q-value=4.54×10^{-3}; beta cell +PI: OR=1.09, q-value=3.06×10^{-2}) and glaucoma (residual glycaemic: OR=1.08, q-value=4.78×10^{-2}; beta cell +PI: OR=1.11, q-value=2.61×10^{-4}). For erectile dysfunction, sensitivity analyses show statistical evidence of a robust risk-increasing causal effect of T2D genetic predisposition driven by glucose regulation clusters (Supplemental Note).

The impact of T2D non-glucose-related pathways on ADHD, PCOS, Alzheimer's disease, and osteoporosis

The protective effect of genetic predisposition to T2D on osteoporosis is driven by the metabolic syndrome (OR=0.89, q-value=3.62×10^{-3}) and obesity clusters (OR=0.83, q-value=2.44×10^{-2}). Similarly, we find negative associations in the All of Us cohort between T2D genetic risk variants assigned to the obesity cluster and osteoporosis (obesity: OR=0.94, q-value=3.99×10^{-7}). For PCOS, the risk-increasing effect of genetic predisposition to T2D is driven by the obesity (OR=1.55, q-value=3.84×10^{-6}) and lipodystrophy clusters (OR=1.44, q-value=4.93×10^{-2}).

For Alzheimer's disease, our findings indicate a risk-increasing effect of T2D genetic predisposition only through the body fat cluster (OR=1.10, q-value=4.83×10^{2}). Adjusting for high-density lipoprotein (HDL) levels fully attenuates this effect (Supplemental Figure 47, SupTab6). For ADHD, a combination of the obesity (OR=1.24, q-value=7.57×10^{-14}), body fat (OR=1.17, q-value=2.83×10^{-7}), and metabolic syndrome clusters (OR=1.11, q-value=2.77×10^{-3}) drive the risk-increasing effect of T2D genetic predisposition. These results suggest that non-glucose-related metabolic and adiposity pathways may contribute to the link between T2D genetic predisposition and increased ADHD and Alzheimer's disease risks.

The multifactorial impact of T2D on CTS, OCD and osteoarthritis

We find evidence of a putative causal link between all T2D genetic clusters except the beta cell clusters and CTS (Figure 3, SupTab2), indicating a multifactorial association mediated by diverse biological pathways. The strongest effects are observed for the obesity (OR=1.23, q-value=6.77×10^{-24}) and lipodystrophy clusters (OR=1.20, q-value=1.15×10^{-3}). In line with these findings, T2D genetic risk variants assigned to these two clusters are positively associated with CTS in the All of Us data (obesity: OR=1.07, q-value=5.02×10^{-8}, lipodystrophy: OR=1.04, q-value=4.08×10^{-2}) (SupTab7). The obesity (OR=0.77, q-value=1.15×10^{-3}) and residual glycaemic clusters (OR=0.84, q-value=4.85×10^{-2}) drive the protective effect of T2D genetic predisposition on OCD.

Osteoarthritis is the only disease for which we find divergent patterns of potentially causal effects across the T2D genetic clusters. The identified risk-increasing effect of genetic predisposition to T2D on osteoarthritis is driven by the obesity cluster (OR=1.15, q-value=3.55×10^{-35}). Supporting our findings, the T2D risk variants assigned to the obesity cluster are positively associated with osteoarthritis risk in the All of Us cohort (OR=1.06, q-value=3.4×10^{-13}). In contrast, genetic predisposition to T2D restricted to the beta cell clusters is associated with reduced osteoarthritis risk (+PI: OR=0.97, q-value=2.17×10^{-3}, -PI: OR=0.96, q-value=4.43×10^{-2}). All estimates of T2D genetic predisposition on osteoarthritis risk are attenuated after adjusting for the effects of BMI, waist-to-hip ratio (WHR), and subcutaneous adipose tissue (SAT) volume (Supplemental Figure 32,34,37, SupTab6).

Potential causal effects of genetic predisposition to T2D vary across genetic ancestry groups

We examined the MR findings within each genetic ancestry group prior to meta-analysis. (Extended Figure 2). The groups comprise individuals genetically similar to Europeans (EUR), East Asians (EAS), South Asians (SAS), and Admixed Americans (AMR) as defined by the 1000 Genomes Project phase 3[21]. We observe divergent patterns of association between T2D genetic predisposition and depression, asthma, and COPD risk among different genetic ancestry groups (Figure 4, SupTab8). Although no statistical evidence of a causal effect of T2D genetic predisposition on depression was found in the meta-analysis across global populations, a protective effect driven by the obesity cluster was found in EAS. On the contrary, in EUR we find evidence of a risk-increasing effect of T2D genetic predisposition on depression linked to the obesity and body fat clusters (Figure 4, SupTab8). All T2D genetic risk variants and those assigned to the obesity cluster are positively associated with depression in the All of Us cohort (all: OR=1.03, q-value=5.35×10^{-4}, obesity: OR=1.05, q-value=4.02×10^{-4}).

We find a protective effect of T2D genetic predisposition on COPD in AMR, contrasting a risk-increasing effect in EUR. For EUR, we find additional risk-increasing effects from the body fat and obesity clusters and a protective effect from the beta cell +PI cluster (Figure 4, SupTab8). In addition to COPD, divergent effects across T2D genetic clusters are observed in osteoarthritis. For both diseases, the risk-increasing effects of the obesity cluster oppose the protective effects of beta cell-related clusters. Results in EAS and SAS show a putative protective effect of T2D genetic predisposition on asthma. In EAS, the effect is driven by the beta cell and the residual glycaemic clusters. Conversely, in EUR, we identify a risk-increasing effect driven by the obesity cluster. These opposing effect directions across different genetic ancestry groups may reflect differences in environmental factors, such as distinct effects of adiposity on T2D pathophysiology across global populations[27, 28]. A further explanation may be the different strengths of association between the T2D genetic clusters and the genetic ancestry groups. For example, in the latest T2DGGI work it was shown that variants assigned to the obesity cluster have greater allelic effects on T2D in EUR than EAS, whereas those assigned to the beta cell clusters have a greater allelic effect on T2D in EAS compared to EUR[11].

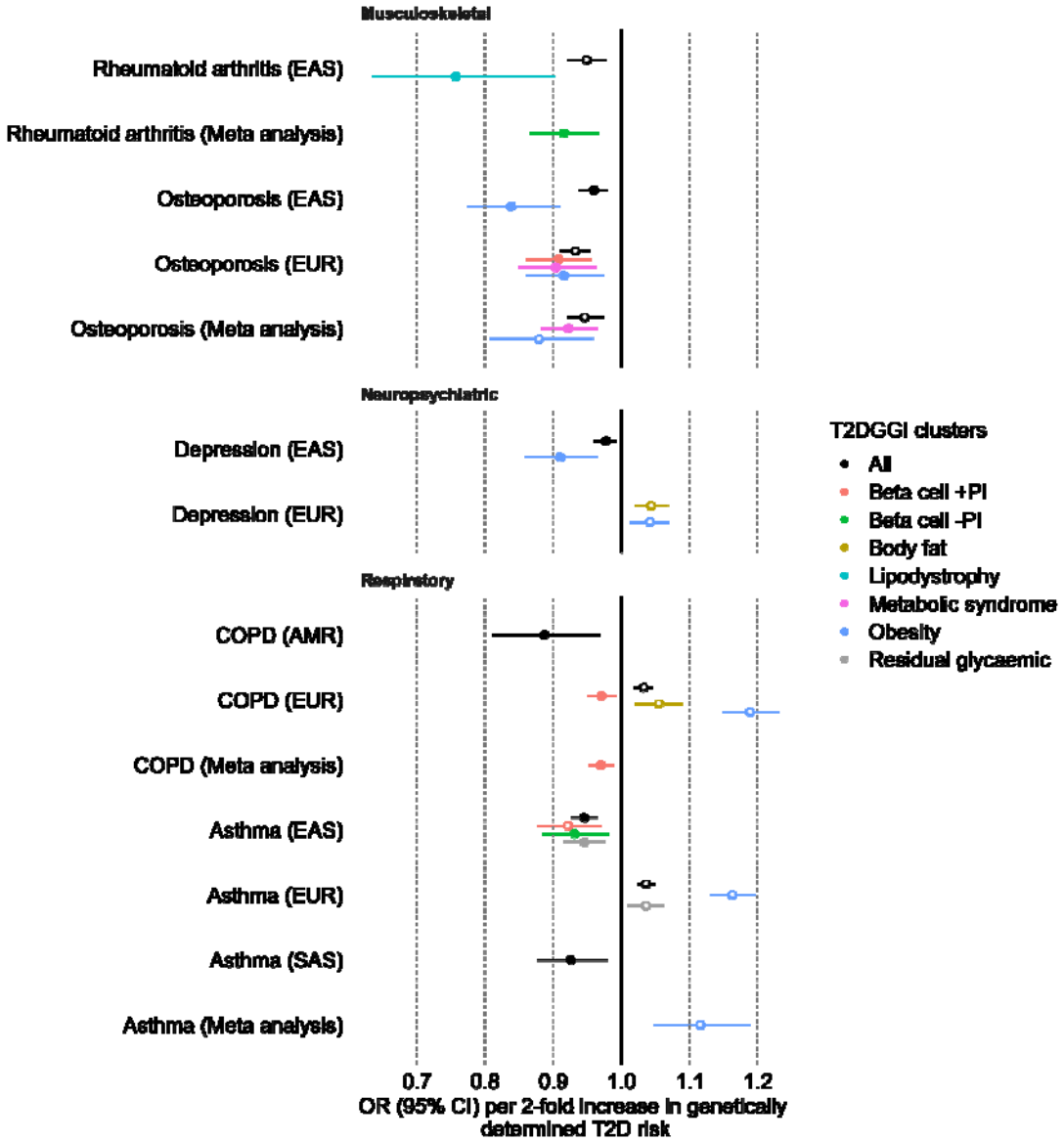


Figure 4: Mendelian randomization (MR) results for different global genetic ancestry groups using cluster-stratified genetic predisposition to T2D and risk for T2D non-cardiovascular comorbidities (EUR=individuals genetically similar to Europeans, EAS=individuals genetically similar to East Asians) for the putative causal relationships (q-value < 0.05). Causal estimates are expressed as the odds ratio (OR) of each comorbidity per doubling (2-fold increase) in genetically determined dichotomous T2D risk. Filled circles mark robust estimates that passed all sensitivity analyses (T2DGGI = Type 2 Diabetes Global Genomics Initiative; CI = confidence interval; PI = proinsulin).

Genetic evidence pointing to metformin as a drug repurposing opportunity for osteoarthritis

Drugs approved for the treatment of T2D may be relevant repurposing candidates for non-cardiovascular comorbidities. Given the availability of a curated list of high-confidence effector genes based on orthogonal functional genomics evidence for osteoarthritis[29], we have applied this hypothesis to the T2D-osteoarthritis comorbidity as proof of principle. By overlaying the osteoarthritis prioritized effector genes with genes targeted by T2D-approved

drugs, we find that metformin, an antihyperglycemic drug, targets the osteoarthritis effector gene *NDUFS3* (GRCh38 chr11:47,565,336-47,584,562). In addition to its effect on lowering blood glucose levels, metformin also has anti-inflammatory and body weight-reducing effects[30]. The T2D index genetic risk variant closest to *NDUSF3*, rs2125838 (GRCh38 chr11:47,790,381), is assigned to the obesity cluster, for which we find a risk-increasing effect on osteoarthritis. In GTEx v10, rs2125838 is a *cis*-eQTL for *NDUFS3* in fibroblasts and cerebellar hemisphere. Previous studies, including Phase III clinical trials, have shown evidence of the beneficial effects of metformin treatment on a subset of osteoarthritis symptoms[31] and in delaying the onset of this disease in T2D patients[32]. Here we provide genetic evidence supporting the weight-reducing effect of metformin as a potential mechanism of action in alleviating the burden of osteoarthritis.

Discussion

T2D is a leading global health concern that impacts individuals and healthcare systems. Effective prevention strategies for T2D and its comorbidities require a deeper understanding of the shared etiology underlying these relationships. Here, we sought to investigate the biological mechanisms underlying putative causal effects of cluster-stratified T2D genetic predisposition on 21 non-cardiovascular, non-oncogenic comorbidities. Our results show that T2D genetic predisposition is a potential driver of comorbidities, rather than the risk of T2D being causally affected by genetic predisposition to its comorbidities. Moreover, T2D genetic predisposition is primarily linked to risk-increasing effects on its comorbidities, reinforcing the role of genetic burden to T2D as a risk factor. The obesity cluster drives most of the observed associations and shows the strongest effects, supporting the well-known role of obesity as a common risk factor for multiple chronic diseases[33]. Finally, we identify evidence of potential causal effects of T2D genetic predisposition on asthma, depression, and COPD with opposing directions across global populations. This observed variation may be largely due to differences in environmental factors across populations [28].

Our genetic findings align with epidemiological observations for many of the identified associations, including cataracts[34], glaucoma[35, 36], PCOS[37, 38], and erectile dysfunction[39-41]. For instance, cohort studies have shown a decreased risk of T2D in anorexia nervosa patients[42], and a longitudinal observational study using the Polish National Health Fund data showed a decreasing trend of anorexia nervosa prevalence in T2D patients[43]. The risk of fractures, a proxy for osteoporosis, is lower in T2D patients than in healthy controls[44], in line with the identified protective effect of T2D genetic predisposition on osteoporosis. We validate previously described evidence of putative causal effects of T2D genetic predisposition on CTS[12], osteoporosis[13], glaucoma[14, 15], cataracts[16], PCOS[17], erectile dysfunction[18], and ADHD risk[19, 20] and prioritize the biological mechanisms underlying these associations. For instance, we show that glucose regulation mechanisms contribute to the associations between T2D genetic predisposition and erectile dysfunction and glaucoma, which aligns with previous genetics-based studies using glycaemic traits data[15, 18, 45, 46]. Obesity has been positively associated with cataracts in a meta-analysis of over 1.6 million individuals[47]. Concordantly, our findings highlight obesity as one of the mechanisms underlying the putative causal effect of T2D genetic predisposition on cataract risk. We also find evidence that genetic variants associated with increased HDL levels have a putative protective effect on clinically diagnosed Alzheimer's disease risk, opposing the previously established risk-increasing

effect[48, 49]. Possible explanations for the observed opposing effect directions include Alzheimer's disease case definitions of previous reports relying on proxy cases and non-age-matched controls[50]. Further studies are needed to disentangle the relationship between HDL levels, T2D, and Alzheimer's disease. For osteoarthritis, we find evidence of diverging cluster-stratified effects of genetic predisposition to T2D, validating the previously shown opposite association between both diseases in addition to shared risk-increasing obesity mechanisms[51].

In addition, our work highlights the T2D-approved drug metformin as a drug repurposing opportunity for osteoarthritis. Metformin is an inhibitor of the mitochondrial complex I, leading to reduced ATP production and activation of the AMPK pathway, a central regulator of cellular energy homeostasis[30]. It targets the osteoarthritis effector gene *NDUFS3* that encodes a core subunit of the mitochondrial complex 1, essential for its catalytic activity and assembly. Homozygous or compound heterozygous mutations in *NDUFS3* lead to mitochondrial complex 1 deficiency nuclear type 8 in humans[52]. Patients suffering from this condition show a wide range of phenotypes, including lactic acidosis and kyphoscoliosis. NDUFS3 is more highly expressed in degraded compared to intact osteoarthritis cartilage[53], pointing to a potential role of this protein in advancing cartilage degradation.

Cluster-stratified MR analysis has been used previously to identify specific biological mechanisms driving putatively causal relationships and potential pleiotropic pathways[54-56]. Although derived from biological associations, one limitation of the eight T2D mechanistic clusters is that they reflect cardiometabolic processes, excluding other potential biological mechanisms involved in T2D. Moreover, the T2D genetic risk variants were clustered based on a hard-clustering approach, which assigns all risk variants to exactly one cluster, with no overlap. The lack of outliers in this approach might decrease the robustness of each cluster. Given the broad nature of the cardiometabolic traits used to cluster the T2D genetic risk variants, they are posited to influence the risk for several diseases, including the comorbidities investigated here. We addressed the potential of confounding or mediation relationships by adjusting the MR estimates for the T2D-related cardiometabolic traits in a multivariable MR approach. If the adjusted effect of T2D genetic predisposition on comorbidity risk is attenuated to zero, care should be taken when interpreting the results given that the unadjusted estimates might be biased.

We recognize that our findings are constrained by the discovery power of the GWAS used. Larger sample sizes may reveal associations with additional T2D genetic clusters. Moreover, the sample size of the T2DGGI GWAS meta-analysis used to identify T2D genetic risk variants and create the mechanistic clusters is much larger than some of the non-cardiovascular comorbidity GWAS used in this work. This may influence the power to detect causal effects in the reverse direction MR analysis. We acknowledge that validation of the findings using sex-specific T2D GWAS data is needed for sex-specific diseases. A well-known limiting factor in statistical genetics analyses is the relative paucity of GWAS data from diverse global populations[57]. In our work, this bias might lead to T2D genetic IVs having a different effect in under-represented global populations and, hence, being less powerful IVs for these populations. Despite the limitations, we identify evidence of a potential causal effect of genetic predisposition to T2D on rheumatoid arthritis only when meta-analyzing the causal estimates across global populations, highlighting the discovery power

gain of integrating diverse GWAS data. Our results underscore the need for the community to continue to pursue efforts to further increase diversity in genetic studies.

In conclusion, we provide evidence of potential causal links between different mechanistic subtypes of T2D genetic predisposition and its non-cardiovascular comorbidities. These findings can inform preventive strategies to mitigate the onset of long-term co-occurring conditions by stratifying and monitoring patients based on their genetic burden to T2D mechanistic subtypes. Moreover, we explore opportunities for drug repurposing, with T2D-approved drugs holding promise for targeting comorbidities. Our work paves the way for more stratified treatment approaches aligned with the genetic and multimorbidity profiles of patients.

Methods

Datasets

The T2DGGI consortium considered six genetic ancestry groups, which refer to the 1000 Genomes Project phase 3[21]: European, East Asian, African American, admixed American, South Asian, and South African (Table_T2DGGI_samplesize). In the meta-analysis across all global populations (2,535,601 individuals including 428,452 cases), 1,289 index (r^2 < 0.05) T2D genetic risk variants were identified at genome-wide significance (p-value < 5×10^{-8}) (SupTab9).

Previous efforts from the T2DGGI consortium have clustered the 1,289 T2D genetic risk variants based on their association profile with 37 cardiometabolic traits. The traits used for clustering included glycaemic traits, anthropometric measures, body fat and adipose tissue volume, blood pressure, circulating plasma lipids levels, and liver function and lipid metabolism biomarkers. A hard clustering approach was performed in an unsupervised manner. This resulted in eight non-overlapping mechanistic clusters of T2D risk variants that represent distinct biological pathways: obesity (n=233), beta-cell associated with positive proinsulin (n=91), beta-cell associated with negative proinsulin (n=89), lipodystrophy (n=45), liver and lipid metabolism (n=3), residual glycaemic (n=389), body fat (n=273) and metabolic syndrome (n=166). We used the T2DGGI GWAS summary statistics to perform MR analyses within genetic ancestry groups[11]. For all other diseases and quantitative traits employed in this work, an overview of the sample sizes and underlying populations is found in SupTable10 and SupTab1.

Approaches to select instrumental variables for T2D

To best maintain the robustness of the T2D genetic risk clusters, we used all 1,289 T2D GWAS index risk variants identified in the T2DGGI global meta-analysis to derive the main results. Index variants were defined as genome-wide significant (p-value < $5x10^{-8}$) variants with r^2 < 0.05 over a 5Mb window[11]. If any genetic instrumental variable (IV) was not present in the outcome trait, we replaced it with an LD-based proxy (r^2 >0.8) if available using the *LDlinkR::Ldproxy()* R function (v1.3.0)[58].

Using all T2D index risk variants, our definition of independent IVs is not as strict as the one commonly employed in MR analyses, defined as LD-based clumped genome-wide

significant (p-value < $5x10^{-8}$) variants with r^2 < 0.001 over a 10Mb window. To address this, we performed sensitivity analyses to select IVs for T2D (Supplemental Figures 3-27). Firstly, we have removed all IVs with evidence of weak instrumental bias, defined as an F-statistic = ($beta^2/se^2$) < 10, where *beta* is the effect size estimate, and *se* is its standard error from the T2D GWAS. Secondly, we compared the effect magnitude of our results with alternative approaches to select T2D IVs. We have employed three additional approaches to define T2D IVs: selecting one variant per locus, LD-based clumping all the T2D index risk variants, and performing cluster-wise LD-based clumping (SupTab11, SupTab12).

Selection of instrumental variables for other traits

For all other traits employed here (non-cardiovascular comorbidities and cardiometabolic traits), we defined IVs as LD-based clumped genome-wide significant (p-value < $5x10^{-8}$) variants with r^2 < 0.001 over a 10Mb window. Clumping was performed with *plink* (v1.9), and LD was calculated based on the 1000 Genomes Project phase 3 release[21] with matching genetic ancestry groups. If any IV was absent in the T2D matching genetic ancestry group GWAS, we replaced it with a proxy. We used the output of plink, which assigns all variants to a clumped result, to search for proxies using the *ieugwasr::ld_matrix()* R function (v1.0.2). We removed variants with an F-statistic<10.

Two-sample Mendelian randomization analysis

Following the STROBE-MR guidelines[59], we performed bi-directional two-sample MR analyses[60] between genetic predisposition to T2D and 21 non-cardiovascular comorbidities. We used the *TwoSampleMR* R package (v0.5.7), which is curated by MR-Base[61]. It has been shown that the estimates of two-sample MR remain unbiased in the presence of sample overlap between exposure and outcome when using large sample sizes[62]. All analyses were performed within genetic ancestry groups using the LD panel from the corresponding genetic ancestry from the 1000 Genomes Project phase 3 release[21] to conduct clumping when necessary. Using matching genetic ancestry group data between exposure and outcome increases the accuracy of the causal estimate by reducing bias due to heterogeneous underlying data distributions.

For the main results, we applied the inverse variance weighted (IVW) method, which performs a random-effects meta-analysis of the Wald ratio estimate of each IV. If only one IV was available, the Wald ratio estimate was used. We applied additional MR methods to ensure consistency of causal effect direction under different methodological assumptions, namely the weighted median, MR-Egger[63], MR-PRESSO[64], and Steiger-filtered IVW methods [65]. We ran the MR-PRESSO distortion test with 1000 iterations to compute the null distribution for the outlier test and increased it to 1500 if it failed. If the distortion test was significant (p-value < 0.05), we considered the effect estimate of the MR-PRESSO outlier-corrected method. Otherwise, we used the estimate of the raw MR-PRESSO result. To facilitate interpretability when using binary traits as exposure, we report the MR results as OR for the outcome per doubling (2-fold change) in genetically determined dichotomous exposure risk. This unit is calculated by multiplying the MR effect sizes by ln(2) and converting them to OR via exponentiation[66].

Assessment of the Mendelian randomization assumptions

MR relies on three main assumptions: relevance, independence, and the exclusion restriction criteria (Supplemental Note). The relevance assumption states that the IVs need to be strongly associated with the exposure and can be directly tested by selecting genome-wide significant and strong (F-statistic > 10) IVs. The independence assumption ensures that the IVs are not associated with any confounder of the exposure-outcome association. The exclusion restriction criterion assumes no horizontal pleiotropy by ensuring that the IVs affect the outcome only through the exposure and not through any alternative pathways. Both assumptions cannot be directly tested but can be assessed via sensitivity analyses. We have conducted multivariable MR analyses to address the independence assumption by adjusting the univariable MR effect for known biological confounders or mediators of the exposure and outcome relationship. To test for horizontal pleiotropy, we applied the MR-Egger intercept test (*TwoSampleMR::mr_pleiotropy_test()*) and the MR-PRESSO outlier and distortion test. Both methods provide causal estimates corrected for the detected pleiotropy, and we looked for a concordant effect direction across different MR methods and the IVW estimate. Moreover, we assessed heterogeneity using I^2, a measure based on Cochran's Q-statistic (*TwoSampleMR::mr_heterogeneity()*) that is more interpretable and independent of the number of studies[67]. Evidence for heterogeneity, defined as $I^2 > 50$, implies that some IVs may influence the outcome through pathways other than the exposure, possibly violating the exclusion restriction assumption. Finally, we have tested for reverse causation by applying a bi-directional MR framework and a directionality test based on the Steiger filter (*TwoSampleMR::directionality_test()*) (SupTab13). Finally, we compared our results with the MR-Clust method, which groups together IVs with similar causal effect estimates on the outcome trait into distinct cluster[55] (Supplemental Note, SupTab14).

Cluster-stratified two-sample Mendelian randomization analysis

To infer the putative causal effects of the eight mechanistic clusters of T2D genetic risk on the analyzed non-cardiovascular comorbidities, we conducted additional MR analyses restricting the T2D IVs to risk variants assigned to each cluster. For these analyses, we followed the same sensitivity analysis procedures described above. In addition, we used the correlated IVW method implemented in the *MendelianRandomization* R package (v0.10)[68], which allows for correlated IVs. This method is designed for less than 500 IVs and, therefore, could not be employed in the MR analyses using all 1,289 T2D genetic risk variants. As a sensitivity analysis to assess the specificity of the clusters, we performed a leave-one-out cluster MR analysis using as IVs the T2D genetic risk variants excluding all variants assigned to one cluster at a time.

Meta-analysis across genetic ancestry groups

In cases of multiple publicly available GWAS summary statistics for a comorbidity from different genetic ancestry groups, we conducted a meta-analysis of the MR estimates across these groups. We used the *rma.uni()* function of the *metafor* R package (v4.6)[69] to conduct a random-effects meta-analysis using the IVW estimate. We tested for heterogeneity using a restricted maximum likelihood estimator (REML). We considered evidence for heterogeneity if $I^2 > 50$.

Definition of significance

We account for multiple testing burden for both tested directions (T2D genetic predisposition on comorbidity risk and genetic predisposition to comorbidity on T2D risk) together by correcting the p-values of the IVW estimate and, if applicable, the correlated IVW estimate across all analyses (including the meta-analyses) using the FDR method (referred to as q-values). We defined statistical significance as IVW estimates and, if applicable, correlated IVW estimates with a q-value<0.05.

Multivariable Mendelian randomization analysis with cardiometabolic traits

For the cluster-stratified genetic predisposition to T2D MR results with FDR<5%, we subsequently performed multivariable MR (MVMR) analyses using the mechanistic clusters of T2D genetic predisposition and several cardiometabolic traits as exposure and the non-cardiovascular comorbidities as outcome. By adjusting cluster-stratified MR estimates for cardiometabolic traits causally associated with the comorbidities, we can estimate the direct effect of the mechanistic clusters of T2D genetic predisposition on its comorbidities. To compare the univariable and multivariable MR results, we first performed univariable MR analyses using the cardiometabolic traits as exposures and the investigated T2D comorbidities as outcomes (Supplemental Figure 2). The results were adjusted for multiple testing using FDR correction. IVs were selected using the same approach as for the comorbidities. We then conducted an MVMR analysis only for the cardiometabolic traits that showed a potential causal effect on comorbidity at FDR of 5%. We applied the *TwoSampleMR::mv_multiple()* function, which uses the IVW method to perform the MVMR analysis. We performed sensitivity analyses using the *MVMR* R package (v0.4)[70], including testing for heterogeneity (*qhet_mvmr()*), pleiotropy (*pleiotropy_mvmr()*), and instrumental strength of each exposure (*strength_mvmr()*).

Phenome-wide association study of cluster-stratified T2D risk variants

To identify non-cardiovascular phenotypes associated with the T2DGGI clusters, we performed PheWAS in each cluster of genetic predisposition to T2D using the National Institutes of Health (NIH) All of Us cohort[22], which was not part of the T2DGGI meta-analysis. All research was carried out on the All of Us Researcher Workbench. Quality control (QC) and genetic ancestry group assignment were performed on 206,000 participants with EHR and WGS data. We pruned the dataset to a maximal set of unrelated individuals using a kinship coefficient < 0.1 and removed samples with ambiguous sex (SupTab15). To reduce the computational complexity, we used the All of Us ACAF variant set, which retains those variants with an allele frequency (AF) > 1% or allele count (AC) >100 in any computed genetic ancestry group.

For each genetic ancestry group, we calculated the first 10 principal components (PCs) using smartpca (v7.2.1)[71] with LD-pruned genotypes and the *fastmode* option enabled. To derive the LD-pruned genotypes, we restricted variants to autosomal variants that were present in 1000 Genomes Project phase 3 release[21], not in the major histocompatibility complex (MHC) region, minor allele frequency (MAF) ≥ 1%, and Hardy-Weinberg equilibrium (HWE) p-value ≥ 1x10-6. LD pruning was conducted using plink2[72] with r^2= 0.05 in a 1000

KB window with an 80b step size. For each genetic ancestry group and T2D genetic cluster, we calculated a GRS summing the cluster-stratified risk alleles, weighted by the effect sizes from each genetic ancestry group. The GRS was adjusted for the 10 first PCs and scaled within each genetic ancestry group.

For each genetic ancestry group and T2D genetic cluster, we performed a PheWAS using the PheTK package (v0.2.1rc5)[73]. Phenotype code (phecode) counts were calculated using the count_phecode() function of PheTK, with the following options: `phecode_version="X"` and `icd_version="US"`. Participants without any recorded phecodes were excluded to avoid incomplete EHRs. PheWAS were conducted with the PheWAS function of PheTK using logistic regression and the model `phecode ~ sex + age at last EHR + PCs 1-10 + cluster GRS`. Only phecodes with a minimum case count of 50 were included, and cases were defined as having at least two instances of the respective phecode. We considered phecodes from the following categories: mental, musculoskeletal, neurological, respiratory, and sense organs. Multiple testing correction was applied for each genetic ancestry group separately using a q-value < 0.05. We meta-analyzed the cluster-stratified PheWAS across genetic ancestry groups using the random-effects IVW method and the *rma.uni()* function of the metafor R package (v4.6)[69]. The meta-analysis results were FDR corrected for multiple testing separately.

Drug repurposing opportunity exemplified by osteoarthritis

By querying the OpenTargets database[74], we identified genes targeted by T2D-approved drugs. Among these genes, we searched for prioritized effector genes for osteoarthritis[29]. The prioritization of osteoarthritis effector genes was based on functional genomics evidence. To gather additional information on the drugs representing repurposing opportunities, we queried the DrugsBank database[75].

Author contributions

A.L.A, O.B, H.T and D.C. contributed equally to this project. A.L.A, O.B, B.F.V., A.P.M. and E.Z. conceived the project and analysis plan. All authors reviewed the analysis plan. A.L.A, O.B, H.T and D.C. conducted all the analyses. The interpretation of the results was conducted by all authors. A.L.A and O.B. wrote the manuscript, which was reviewed by all authors.

Acknowledgments

O.B has received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement No 101017802 (OPTOMICS). Supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute (NHLBI) grant R01HL105756. J.M.M. is supported by American Diabetes Association grant #11-22-ICTSPM-16 and by NHGRI U01HG011723, by the National Institute Of Diabetes And Digestive And Kidney Diseases of the National Institutes of Health under Award Number R01DK137993 and U01 DK140757, AMP CMD award from RFP 6

from the Foundation for the National Institutes of Health, and a Medical University of Bialystok (MUB) grant from the Ministry of Science and Higher Education (Poland).

Data availability

The GWAS summary statistics used in this work are publicly available and referenced in Supplementary Tables 1 and 10. Researchers can apply to access the individual-level data of the All of Us program (https://researchallofus.org/) used to perform the PheWAS. The list of prioritized effector genes for osteoarthritis is publicly available in the referenced publication. The GTEx v10 data used in this manuscript were obtained from: the GTEx Portal on 02/21/25 (dbGaP Accession phs000424.v10.p2).

Code availability

The code used to perform all the MR-related analysis can be found on GitHub under GNU General Public License (GPL-3) at https://github.com/nalu357/T2DGGI_comorbidity and on Zenodo at https://doi.org/10.5281/zenodo.15168490. The PheTK package used to perform the All Of Us PheWAS is freely available on the Python Package Index, on GitHub under GNU General Public License (GPL-3) at https://github.com/nhgritctran/PheTK and on Zenodo at https://doi.org/10.5281/zenodo.14217954.

Extended figures

Extended Figure 1: Results of cluster-stratified two-sample Mendelian randomization (MR) analysis of genetic predisposition to type 2 diabetes (T2D) on non-cardiovascular comorbidities risk. Causal estimates are expressed as the odds ratio (OR) of each comorbidity per doubling (2-fold increase) in genetically determined dichotomous T2D risk. Filled circles mark estimates with a q-value < 0.05 that passed all sensitivity analyses to assess the validity of the MR assumptions. (T2DGGI = Type 2 Diabetes Global Genomics Initiative; CI = confidence interval; PI = proinsulin; CTS = Carpal tunnel syndrome; ADHD = attention-deficit/hyperactivity disorder; OCD = obsessive-compulsive disorder; PCOS = polycystic ovary syndrome; COPD = chronic obstructive pulmonary disease)

Extended Figure 2: Results of the single-ancestry cluster-stratified two-sample Mendelian randomization (MR) analysis of genetic predisposition to type 2 diabetes (T2D) on non-cardiovascular comorbidities risk. Causal estimates are expressed as the odds ratio (OR) of each comorbidity per doubling (2-fold increase) in genetically determined dichotomous T2D risk. Filled circles mark estimates with a q-value < 0.05 that passed all sensitivity analyses to assess the validity of the MR assumptions. (T2DGGI = Type 2 Diabetes Global Genomics Initiative; CI = confidence interval; PI = proinsulin; COPD = chronic obstructive pulmonary disease)

Supplemental tables

Supplemental Table 1: Overview of non-cardiovascular type 2 diabetes comorbidities GWAS data including sample size, genetic ancestry group information and reference of the publication.

Supplemental Table 2: Results of the Mendelian randomization analysis using cluster-stratified genetic predisposition to type 2 diabetes as the exposure and non-cardiovascular comorbidities as the outcome, along with sensitivity analysis results.

Supplemental Table 3: Results of the reverse Mendelian randomization analysis using genetic predisposition to non-cardiovascular comorbidities as the exposure and type 2 diabetes as the outcome, along with sensitivity analysis results.

Supplemental Table 4: Results of the leave-one-cluster-out Mendelian randomization analysis using genetic predisposition to type 2 diabetes as the exposure and non-cardiovascular comorbidities as the outcome.

Supplemental Table 5: Results of the univariate cardiometabolic Mendelian randomization analysis using genetic predisposition to cardiometabolic traits as the exposure and non-cardiovascular comorbidities as the outcome, along with sensitivity analysis results.

Supplemental Table 6: Results of the multivariate cardiometabolic Mendelian randomization analysis using genetic predisposition to cardiometabolic traits and type 2 diabetes as the exposures and non-cardiovascular comorbidities as the outcome, along with sensitivity analysis results.

Supplemental Table 7: PheWAS results of each type 2 diabetes genetic cluster in the All of Us cohort performed in each genetic ancestry group and meta-analyzed across these.

Supplemental Table 8: Results of the Mendelian randomization analysis for each genetic ancestry group using cluster-stratified genetic predisposition to type 2 diabetes as the exposure and non-cardiovascular comorbidities as the outcome, along with sensitivity analysis results. These results were meta-analyzed across genetic ancestry groups (Supplemental Table 2).

Supplemental Table 9: Sample size of each genetic ancestry group type 2 diabetes GWAS meta-analysis from T2DGGI consortium.

Supplemental Table 10: Overview of cardiometabolic traits that are potential confounders or mediators of the effects of genetic predisposition to type 2 diabetes, including sample size, genetic ancestry group information and reference of the publication.

Supplemental Table 11: Number of type 2 diabetes genetic instrumental variables (IVs) for each mechanistic cluster using different approaches to select IVs employed as sensitivity analysis for the Mendelian randomization analysis.

Supplemental Table 12A: Results of the Mendelian randomization analysis and sensitivity analysis results using type 2 diabetes genetic instrument variables selected using cluster-stratified LD-based clumping.

Supplemental Table 12B: Results of the Mendelian randomization analysis and sensitivity analysis results using type 2 diabetes (T2D) genetic instrument variables selected using the strongest association per T2D genetic risk region.

Supplemental Table 13: Results of the directionality test for Mendelian randomization results using cluster-stratified genetic predisposition to type 2 diabetes as the exposure and non-cardiovascular comorbidities as the outcome.

Supplemental Table 14: Results (FDR < 5%) of the MR-Clust method.

Supplemental Table 15: Sample size of each genetic ancestry group from the All of Us genetic data used to perform the PheWAS.

Supplementary figures

Supplemental Figure 1: Results of the reverse Mendelian randomization (MR) analysis using genetic predisposition to type 2 diabetes (T2D) comorbidities as exposure and T2D as the outcome. Causal estimates are expressed as the odds ratio of comorbidity risk per doubling (2-fold increase) in genetically determined dichotomous T2D risk. Filled circles mark estimates with a false-discovery-rate correction of 5%. No estimate passed the MR sensitivity analyses. The genetic ancestry groups represent individuals genetically similar to Africans (AFR), East Asians (EAS), Europeans (EUR), admixed Americans (AMR) and South Asians (SAS) as defined by the 1000 Genomes Project phase 3. (CI = confidence interval).

Supplemental Figure 2: Univariable Mendelian randomization (MR) results with genetic predisposition to cardiometabolic traits as exposures and type 2 diabetes (T2D) comorbidities as outcomes. Causal estimates are expressed as the odds ratio of comorbidity risk per doubling (2-fold increase) in genetically determined dichotomous T2D risk. Filled circles denote robust causal estimates that passed sensitivity analysis and false-discovery rate correction at 5%. (CI = confidence interval; PI = proinsulin; CTS = Carpal tunnel syndrome; ADHD = attention-deficit/hyperactivity disorder; OCD = obsessive-compulsive disorder; PCOS = polycystic ovary syndrome; COPD = chronic obstructive pulmonary disease; HDL=high-density lipoprotein cholesterol; WHR = waist-to-hip ratio; BMI=body mass index).

Supplemental Figures 3-27: Comparison between the results of Mendelian randomization (MR) results using the inverse variance weighted method and different approaches to select genetic instrumental variables (IVs). The forest plots depict all cluster-stratified estimates of genetic predisposition to type 2 diabetes (T2D) on comorbidity risk. Causal estimates are expressed as the odds ratio of comorbidity risk per doubling (2-fold increase) in genetically determined dichotomous T2D risk. Filled circles mark estimates with FDR < 5%. The genetic ancestry groups represent individuals genetically similar to Africans (AFR), East Asians (EAS), Europeans (EUR), admixed Americans (AMR) and South Asians (SAS) as defined by the 1000 Genomes Project. (CI = confidence interval).

Supplemental Figures 28-48: Forest plots comparing the results of the univariable and multivariable Mendelian randomization (MR) analyses using data from individuals genetically similar to Europeans. Causal estimates are expressed as the odds ratio of comorbidity risk per doubling (2-fold increase) in genetically determined dichotomous T2D risk. (HDL=high-density lipoprotein cholesterol; WHR = waist-to-hip ratio; BMI=body mass index; CI=confidence interval; FDR=false-discovery-rate at 5%).

Supplemental Figures 49-63: MR-Clust scatterplots of putative T2D causal clusters on non-cardiometabolic comorbidities - MR-Clust (V.0.1.0) was used to identify putatively causal clusters of genetic predisposition to T2D as an exposure on non-cardiometabolic comorbidities as an outcome. We restricted putative causal clusters to instrumental variables (IVs) with a >80% probability of inclusion in the cluster, and removed any IVs placed in "Null" or "Junk" clusters by the MR-Clust algorithm. The genetic ancestry groups represent individuals genetically similar to Africans (AFR), East Asians (EAS), Europeans (EUR), admixed Americans (AMR) and South Asians (SAS) as defined by the 1000 Genomes Project. All scatterplots were generated using the "two_stage_plot" function implemented in MR-Clust. When possible, T2D IVs were annotated in the scatterplot by their cardiometabolic cluster from Suzuki et al (2024)(1).

References

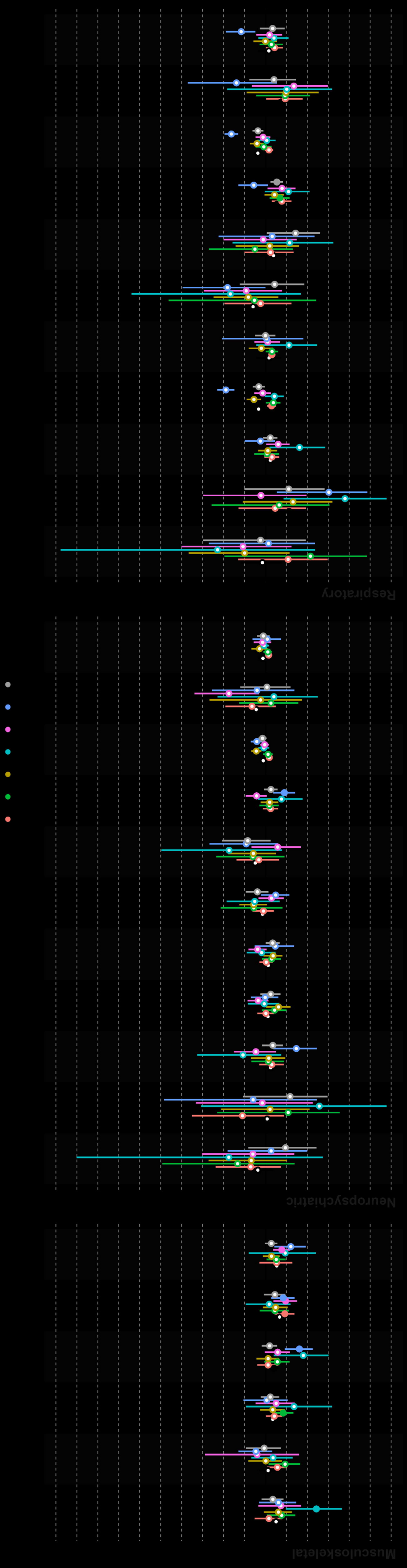
1. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet, 2023. **402**(10397): p. 203-234.
2. Parker, E.D., et al., Economic Costs of Diabetes in the U.S. in 2022. Diabetes Care, 2024. **47**(1): p. 26-43.
3. Pearson-Stuttard, J., et al., Variations in comorbidity burden in people with type 2 diabetes over disease duration: A population-based analysis of real world evidence. eClinicalMedicine, 2022. **52**: p. 101584-101584.
4. Martens, T., et al., Effect of Continuous Glucose Monitoring on Glycemic Control in Patients With Type 2 Diabetes Treated With Basal Insulin: A Randomized Clinical Trial. Jama, 2021. **325**(22): p. 2262-2272.
5. Drucker, D.J., Efficacy and Safety of GLP-1 Medicines for Type 2 Diabetes and Obesity. Diabetes Care, 2024. **47**(11): p. 1873-1888.
6. Gregg, E.W., et al., The burden of diabetes-associated multiple long-term conditions on years of life spent and lost. Nat Med, 2024. **30**(10): p. 2830-2837.
7. Tobias, D.K., et al., Second international consensus report on gaps and opportunities for the clinical translation of precision diabetes medicine. Nature Medicine, 2023. **29**(10): p. 2438-2457.
8. Bonnefond, A., et al., Dissection of type 2 diabetes: a genetic perspective. The Lancet Diabetes & Endocrinology, 2025.
9. DeFronzo, R.A., et al., Type 2 diabetes mellitus. Nature Reviews Disease Primers, 2015. **1**(1): p. 15019.
10. Smith, K., et al., Multi-ancestry polygenic mechanisms of type 2 diabetes. Nature Medicine, 2024. **30**(4): p. 1065-1074.
11. Suzuki, K., et al., Genetic drivers of heterogeneity in type 2 diabetes pathophysiology. Nature, 2024. **627**(8003): p. 347-357.
12. Mi, J. and Z. Liu, Obesity, Type 2 Diabetes, and the Risk of Carpal Tunnel Syndrome: A Two-Sample Mendelian Randomization Study. Front Genet, 2021. **12**: p. 688849.
13. Cheng, L., S. Wang, and H. Tang, Type 2 diabetes mellitus plays a protective role against osteoporosis --mendelian randomization analysis. BMC Musculoskelet Disord, 2023. **24**(1): p. 444.

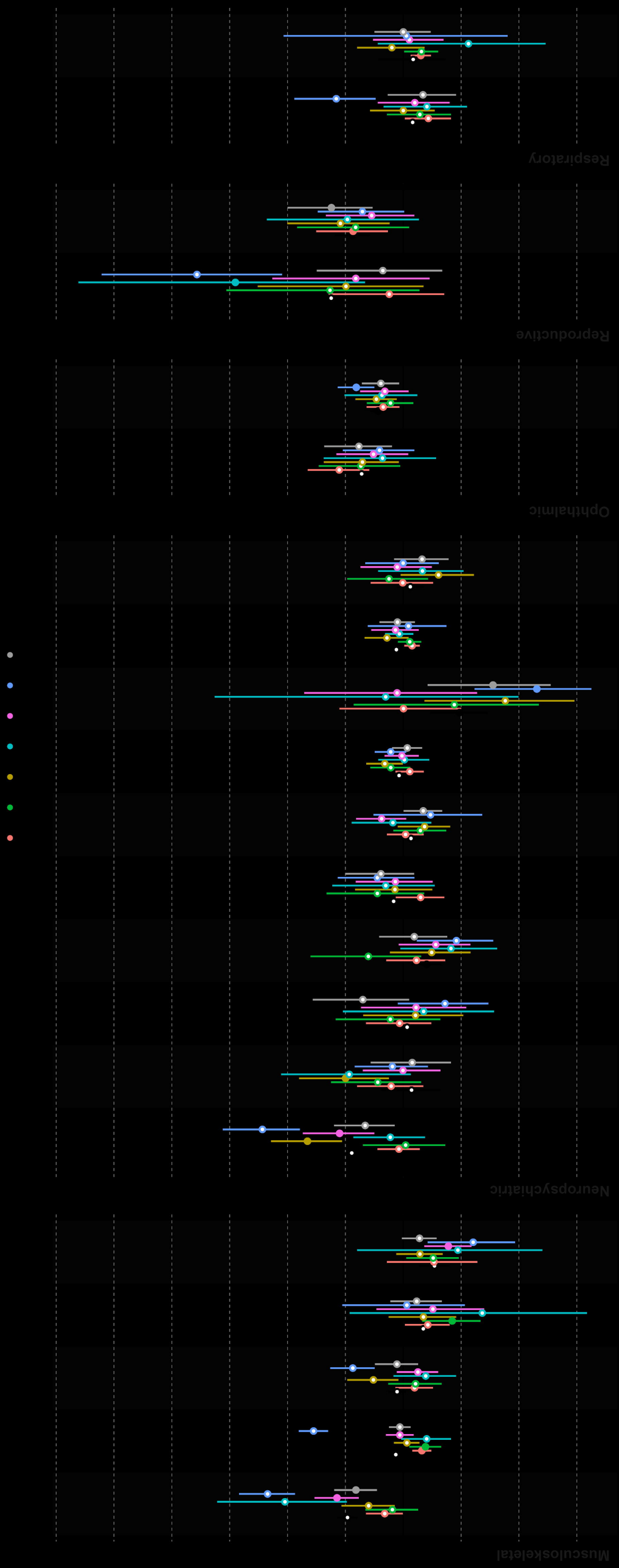
14. Julian, T.H., et al., Causal factors in primary open angle glaucoma: a phenome-wide Mendelian randomisation study. Sci Rep, 2023. **13**(1): p. 9984.
15. Chen, R., et al., Dissecting causal associations of type 2 diabetes with 111 types of ocular conditions: a Mendelian randomization study. Front Endocrinol (Lausanne), 2023. **14**: p. 1307468.
16. Yuan, S., A. Wolk, and S.C. Larsson, Metabolic and lifestyle factors in relation to senile cataract: a Mendelian randomization study. Sci Rep, 2022. **12**(1): p. 409.
17. Liu, Q., et al., A genome-wide cross-trait analysis identifies shared loci and causal relationships of type 2 diabetes and glycaemic traits with polycystic ovary syndrome. Diabetologia, 2022. **65**(9): p. 1483-1494.
18. Yuan, C., et al., Type 2 diabetes mellitus increases risk of erectile dysfunction independent of obesity and dyslipidemia: A Mendelian randomization study. Andrology, 2022. **10**(3): p. 518-524.
19. Baranova, A., et al., Shared genetics and bidirectional causal relationships between type 2 diabetes and attention-deficit/hyperactivity disorder
Development of Osteoarthritis in Adults With Type 2 Diabetes Treated With Metformin vs a Sulfonylurea. Gen Psychiatr, 2023. **36**(2): p. e100996.
20. Tao, H., et al., Psychiatric disorders and Type 2 diabetes mellitus: A bidirectional Mendelian randomization. Eur J Clin Invest, 2023. **53**(3): p. e13893.
21. Auton, A., et al., A global reference for human genetic variation. Nature 2015 526:7571, 2015. **526**(7571): p. 68-74.
22. Denny, J.C., et al., The "All of Us" Research Program. N Engl J Med, 2019. **381**(7): p. 668-676.
23. Ferraz-Amaro, I., et al., Impaired beta cell function is present in nondiabetic rheumatoid arthritis patients. Arthritis Res Ther, 2013. **15**(1): p. R17.
24. Nicolau, J., et al., Rheumatoid arthritis, insulin resistance, and diabetes. Joint Bone Spine, 2017. **84**(4): p. 411-416.
25. Machado, F.V.C., et al., Physiopathological relationship between chronic obstructive pulmonary disease and insulin resistance. Endocrine, 2018. **61**(1): p. 17-22.
26. Bolton, C.E., et al., Insulin resistance and inflammation - A further systemic complication of COPD. Copd, 2007. **4**(2): p. 121-6.
27. Ma, R.C. and J.C. Chan, Type 2 diabetes in East Asians: similarities and differences with populations in Europe and the United States. Ann N Y Acad Sci, 2013. **1281**(1): p. 64-91.
28. Eastwood, S.V., et al., Ancestry, ethnicity, and race: explaining inequalities in cardiometabolic disease. Trends in Molecular Medicine, 2024. **30**(6): p. 541-551.
29. Hatzikotoulas, K., et al., Translational genomics of osteoarthritis in 1,962,069 individuals. Nature, 2025.
30. Foretz, M., B. Guigas, and B. Viollet, Metformin: update on mechanisms of action and repurposing potential. Nat Rev Endocrinol, 2023. **19**(8): p. 460-476.
31. Alimoradi, N., et al., Metformin attenuates symptoms of osteoarthritis: role of genetic diversity of Bcl2 and CXCL16 in OA. Arthritis Res Ther, 2023. **25**(1): p. 35.
32. Baker, M.C., et al., Development of Osteoarthritis in Adults With Type 2 Diabetes Treated With Metformin vs a Sulfonylurea. JAMA Netw Open, 2023. **6**(3): p. e233646.
33. Mounier, N., et al., Genetics identifies obesity as a shared risk factor for co-occurring multiple long-term conditions. medRxiv, 2024: p. 2024.07.10.24309772.
34. Becker, C., et al., Cataract in patients with diabetes mellitus-incidence rates in the UK and risk factors. Eye (Lond), 2018. **32**(6): p. 1028-1035.

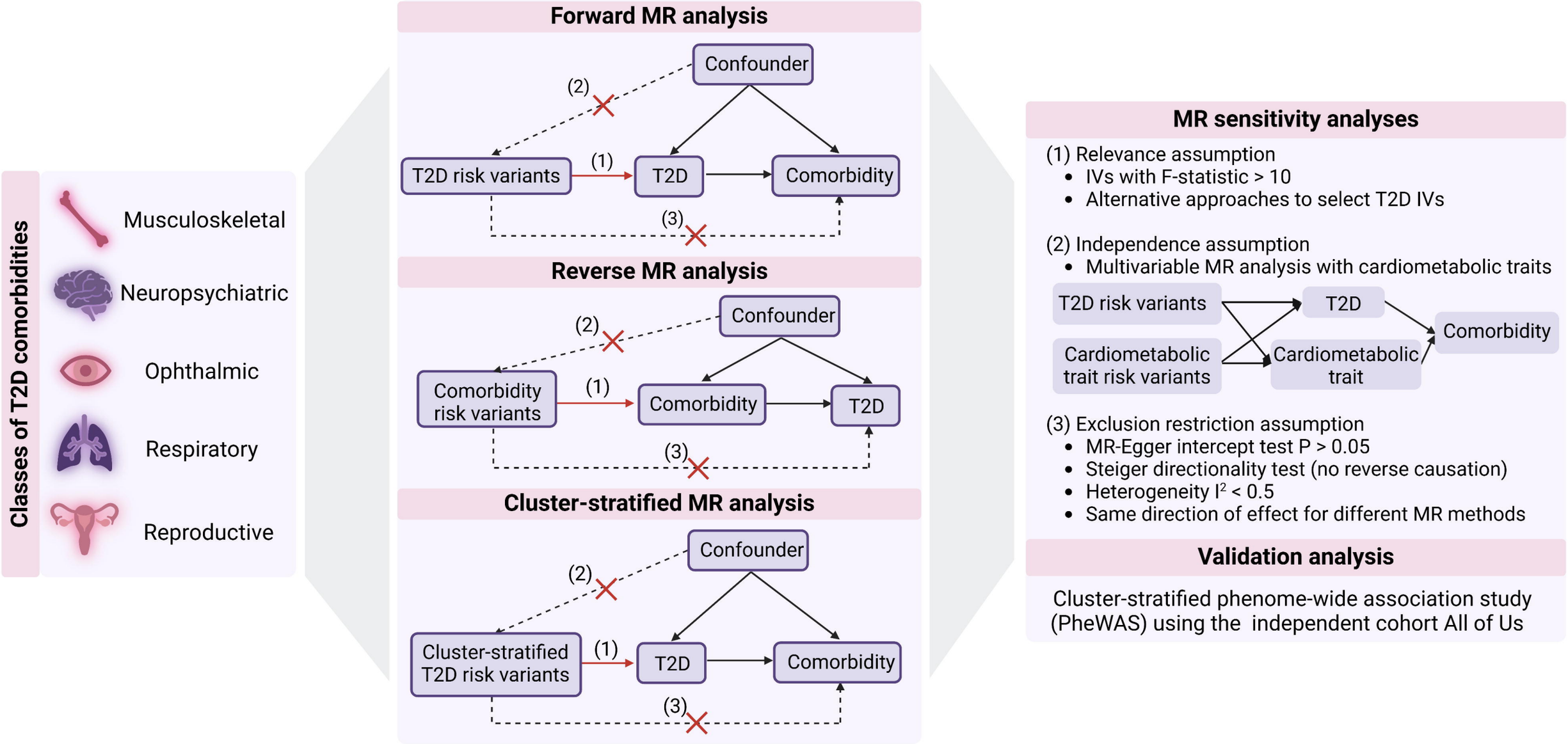
35. Chopra, V., et al., Type 2 diabetes mellitus and the risk of open-angle glaucoma the Los Angeles Latino Eye Study. Ophthalmology, 2008. **115**(2): p. 227-232.e1.
36. Singh, W., et al., Prevalence and Associated Risk Factors of Primary Open Angle Glaucoma Among Patients With Type 2 Diabetes Mellitus: A Cross-Sectional Study From North India. Cureus, 2022. **14**(9): p. e28908.
37. Cioana, M., et al., Prevalence of Polycystic Ovary Syndrome in Patients With Pediatric Type 2 Diabetes: A Systematic Review and Meta-analysis. JAMA Network Open, 2022. **5**(2): p. e2147454-e2147454.
38. Long, C., et al., Prevalence of polycystic ovary syndrome in patients with type 2 diabetes: A systematic review and meta-analysis. Front Endocrinol (Lausanne), 2022. **13**: p. 980405.
39. Parmar, R.S., et al., Prevalence of erectile dysfunction in Type 2 diabetes mellitus (T2DM) and its predictors among diabetic men. J Family Med Prim Care, 2022. **11**(7): p. 3875-3879.
40. Kouidrat, Y., et al., High prevalence of erectile dysfunction in diabetes: a systematic review and meta-analysis of 145 studies. Diabet Med, 2017. **34**(9): p. 1185-1192.
41. kitaw, T.A., et al., The global burden of erectile dysfunction and its associated risk factors in diabetic patients: an umbrella reviews. BMC Public Health, 2024. **24**(1): p. 2816.
42. Ji, J., J. Sundquist, and K. Sundquist, Association between anorexia nervosa and type 2 diabetes in Sweden: Etiological clue for the primary prevention of type 2 diabetes. Endocrine Research, 2016. **41**(4): p. 310-316.
43. Jaworski, M., et al., A Ten-Year Longitudinal Study of Prevalence of Eating Disorders in the General Polish Type 2 Diabetes Population. Med Sci Monit, 2018. **24**: p. 9204-9212.
44. Davie, G.S., et al., Incident Type 2 Diabetes and Risk of Fracture: A Comparative Cohort Analysis Using U.K. Primary Care Records. Diabetes Care, 2021. **44**(1): p. 58-66.
45. Jiang, C., et al., New evidence for the effect of type 2 diabetes and glycemic traits on testosterone levels: a two-sample Mendelian randomization study. Front Endocrinol (Lausanne), 2023. **14**: p. 1238090.
46. Seo, J.H. and Y. Lee, Possible Causal Association between Type 2 Diabetes and Glycaemic Traits in Primary Open-Angle Glaucoma: A Two-Sample Mendelian Randomisation Study. Biomedicines, 2024. **12**(4).
47. Niazi, S., et al., Association between obesity and age-related cataract: an updated systematic review and dose-response meta-analysis of prospective cohort studies. Front Nutr, 2023. **10**: p. 1215212.
48. Hussain, S.M., et al., Association of plasma high-density lipoprotein cholesterol level with risk of incident dementia: a cohort study of healthy older adults. Lancet Reg Health West Pac, 2024. **43**: p. 100963.
49. Luo, J., et al., Genetic Associations Between Modifiable Risk Factors and Alzheimer Disease. JAMA Netw Open, 2023. **6**(5): p. e2313734.
50. Wu, Y., et al., Pervasive biases in proxy genome-wide association studies based on parental history of Alzheimer's disease. Nat Genet, 2024.
51. Arruda, A.L., et al., Genetic underpinning of the comorbidity between type 2 diabetes and osteoarthritis. Am J Hum Genet, 2023. **110**(8): p. 1304-1318.
52. Haack, T.B., et al., Molecular diagnosis in mitochondrial complex I deficiency using exome sequencing. Journal of Medical Genetics, 2012. **49**(4): p. 277.
53. Steinberg, J., et al., A molecular quantitative trait locus map for osteoarthritis. Nature Communications 2021 12:1, 2021. **12**(1): p. 1-11.

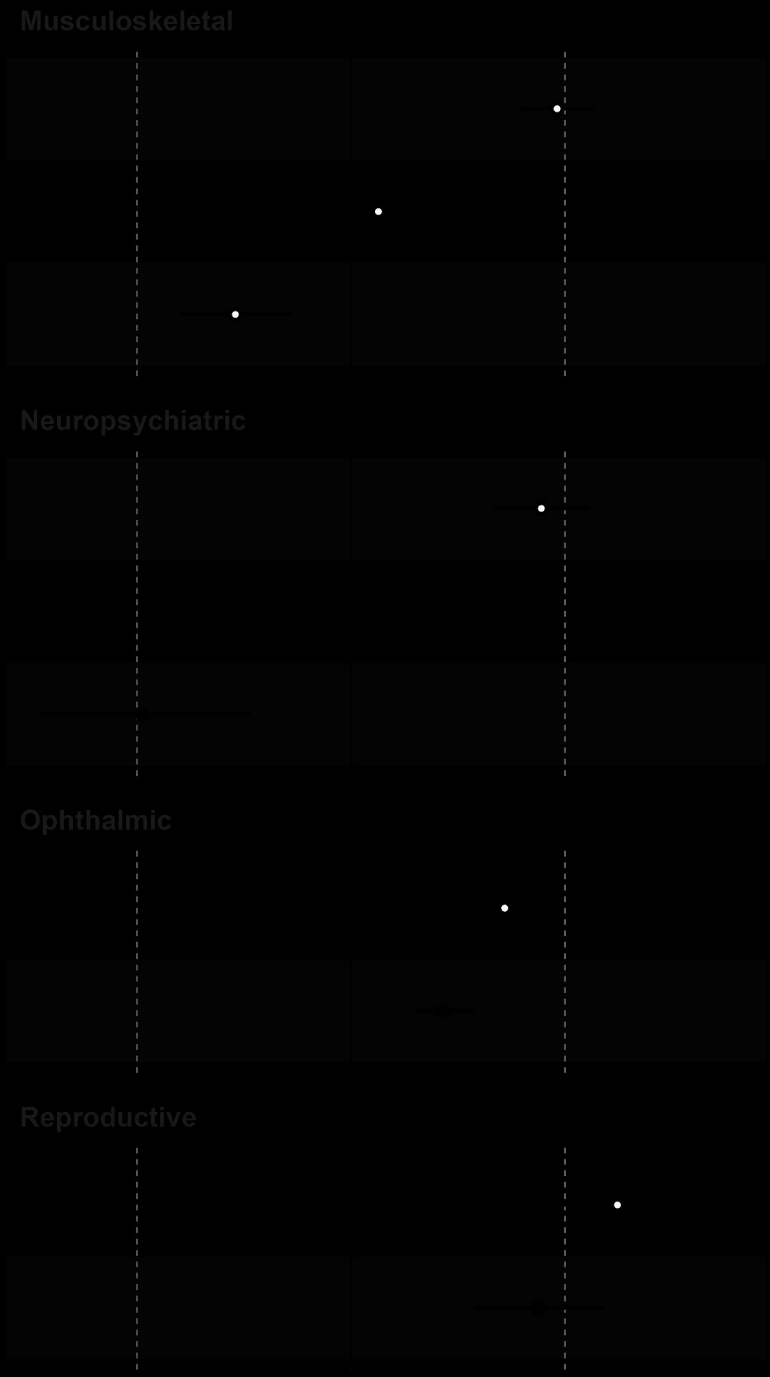
54. Darrous, L., et al., PheWAS-based clustering of Mendelian Randomisation instruments reveals distinct mechanism-specific causal effects between obesity and educational attainment. Nat Commun, 2024. **15**(1): p. 1420.
55. Foley, C.N., et al., MR-Clust: clustering of genetic variants in Mendelian randomization with similar causal estimates. Bioinformatics, 2021. **37**(4): p. 531-541.
56. Liang, X., et al., Using clustering of genetic variants in Mendelian randomization to interrogate the causal pathways underlying multimorbidity from a common risk factor. Genet Epidemiol, 2024.
57. Mills, M.C. and C. Rahal, The GWAS Diversity Monitor tracks diversity by disease in real time. Nature Genetics, 2020. **52**(3): p. 242-243.
58. Myers, T.A., S.J. Chanock, and M.J. Machiela, LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations. Frontiers in Genetics, 2020. **11**.
59. Skrivankova, V.W., et al., Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. JAMA, 2021. **326**(16): p. 1614-1621.
60. Sanderson, E., et al., Mendelian randomization. Nature Reviews Methods Primers 2022 2:1, 2022. **2**(1): p. 1-21.
61. Hemani, G., et al., The MR-base platform supports systematic causal inference across the human phenome. eLife, 2018. **7**.
62. Mounier, N. and Z. Kutalik, Bias correction for inverse variance weighting Mendelian randomization. Genet Epidemiol, 2023. **47**(4): p. 314-331.
63. Bowden, J., G. Davey Smith, and S. Burgess, Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International Journal of Epidemiology, 2015. **44**(2): p. 512-525.
64. Verbanck, M., et al., Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet, 2018. **50**(5): p. 693-698.
65. Hemani, G., K. Tilling, and G. Davey Smith, Orienting the causal relationship between imprecisely measured traits using GWAS summary data. PLOS Genetics, 2017. **13**(11): p. e1007081-e1007081.
66. Burgess, S. and J.A. Labrecque, Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. Eur J Epidemiol, 2018. **33**(10): p. 947-952.
67. Higgins, J.P. and S.G. Thompson, Quantifying heterogeneity in a meta-analysis. Stat Med, 2002. **21**(11): p. 1539-58.
68. Patel, A., et al., MendelianRandomization v0.9.0: updates to an R package for performing Mendelian randomization analyses using summarized data
DrugBank 6.0: the DrugBank Knowledgebase for 2024. Wellcome Open Res, 2023. **8**(D1): p. 449.
69. Viechtbauer, W., Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software, 2010. **36**(3): p. 1 - 48.
70. Sanderson, E., W. Spiller, and J. Bowden, Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. Statistics in Medicine, 2021. **40**(25): p. 5434-5452.
71. Price, A.L., et al., Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet, 2006. **38**(8): p. 904-9.

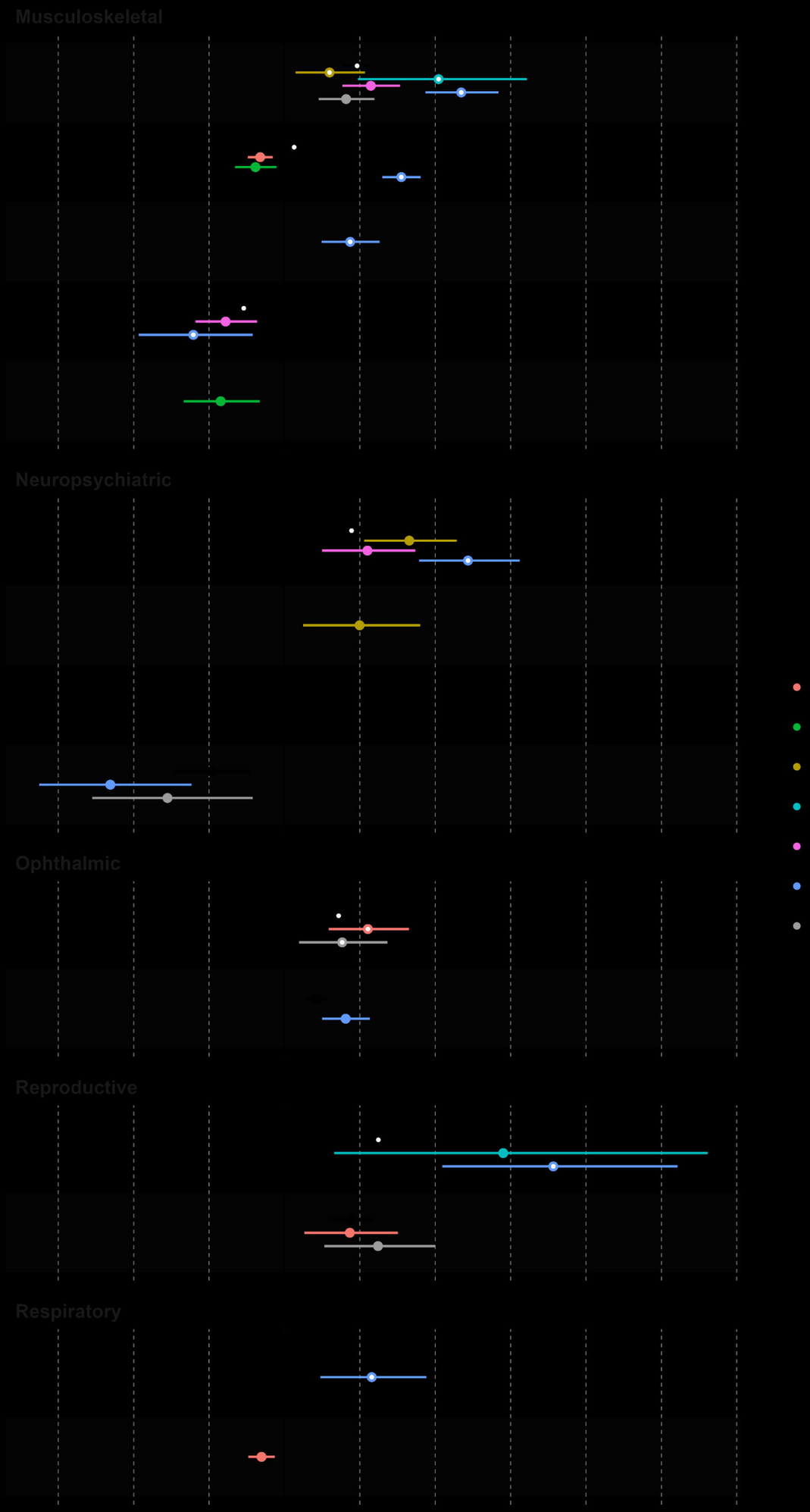
72. Chang, C.C., et al., Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience, 2015. **4**: p. 7.
73. Tran, T.C., et al., PheWAS analysis on large-scale biobank data with PheTK. medRxiv, 2024.
74. Ochoa, D., et al., The next-generation Open Targets Platform: reimagined, redesigned, rebuilt. Nucleic Acids Research, 2022. **51**(D1): p. D1353-D1359.
75. Knox, C., et al., DrugBank 6.0: the DrugBank Knowledgebase for 2024. Nucleic Acids Res, 2024. **52**(D1): p. D1265-d1275.

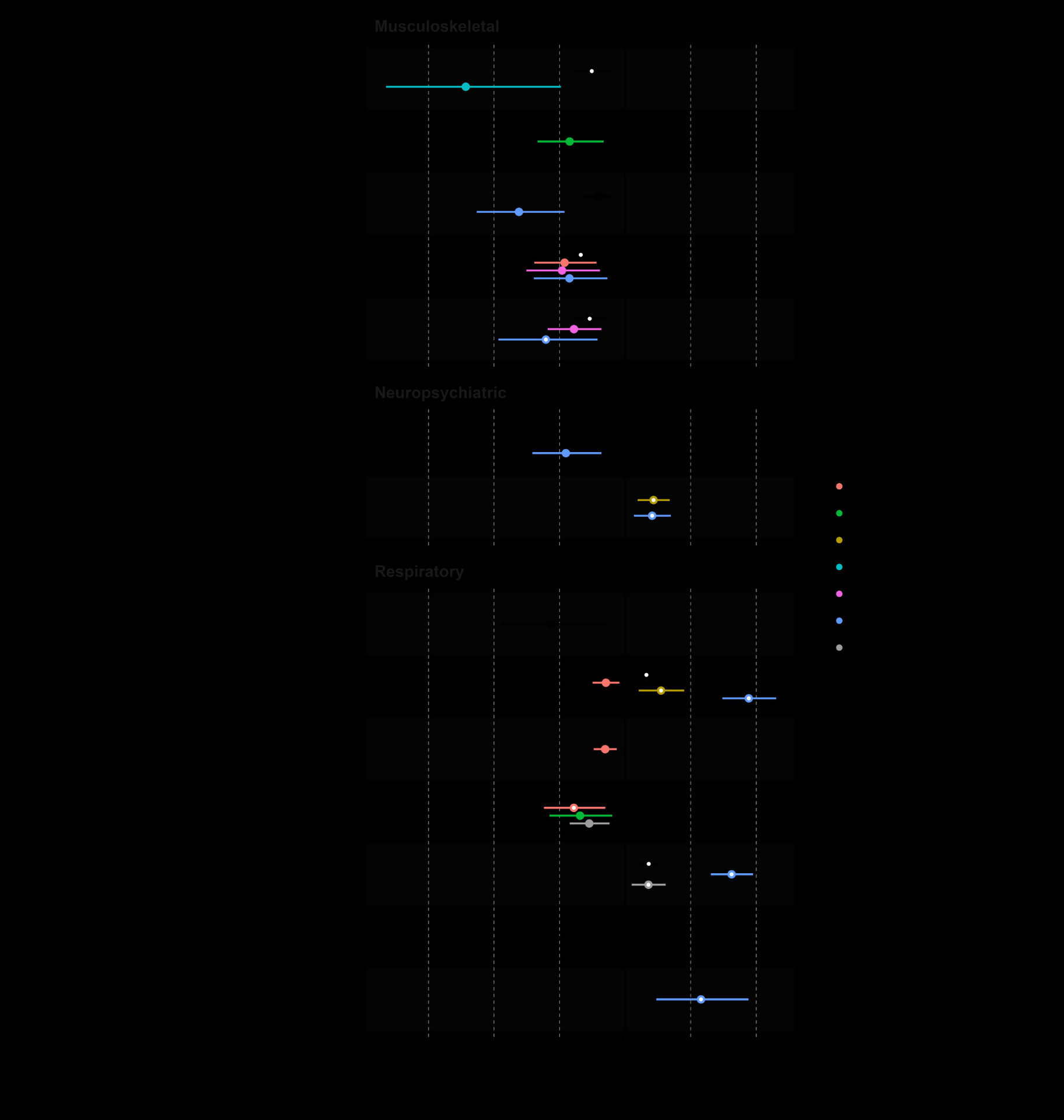

Respiratory
Neuropsychiatric
Musculoskeletal


Respiratory
Reproductive
Ophthalmic
Neuropsychiatric
Musculoskeletal


Classes of T2D comorbidities
Musculoskeletal
Neuropsychiatric
Ophthalmic
Respiratory
Reproductive
Forward MR analysis
Confounder
(2)
T2D risk variants
(1)
T2D
Comorbidity
(3)
Reverse MR analysis
Confounder
(2)
Comorbidity risk variants
(1)
Comorbidity
T2D
(3)
Cluster-stratified MR analysis
Confounder
(2)
Cluster-stratified T2D risk variants
(1)
T2D
Comorbidity
(3)
MR sensitivity analyses
(1) Relevance assumption
• IVs with F-statistic > 10
• Alternative approaches to select T2D IVs
(2) Independence assumption
• Multivariable MR analysis with cardiometabolic traits
T2D risk variants
T2D
Comorbidity
Cardiometabolic trait risk variants
Cardiometabolic trait
(3) Exclusion restriction assumption
• MR-Egger intercept test P > 0.05
• Steiger directionality test (no reverse causation)
• Heterogeneity I² < 0.5
• Same direction of effect for different MR methods
Validation analysis
Cluster-stratified phenome-wide association study (PheWAS) using the independent cohort All of Us


Musculoskeletal
Neuropsychiatric
Ophthalmic
Reproductive


Musculoskeletal
Neuropsychiatric
Ophthalmic
Reproductive
Respiratory


Musculoskeletal
Neuropsychiatric
Respiratory