

Université de Montréal

Vérification de la pléiotropie en randomisation mendélienne :
*Évaluation méthodologique et application à l'estimation de l'effet causal de
l'adiposité sur la pression artérielle*

Par

Fiston Ikwa Ndol Mbutiwi

Département de Médecine Sociale et Préventive
École de Santé Publique

Thèse présentée en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.) en santé publique
Option épidémiologie

Juillet, 2023

© Fiston Ikwa Ndol Mbutiwi, 2023

Université de Montréal
Département de Médecine Sociale et Préventive
École de Santé Publique

Cette thèse intitulée

Vérification de la pléiotropie en randomisation mendélienne :
*Évaluation méthodologique et application à l'estimation de l'effet causal de
l'adiposité sur la pression artérielle*

Présentée par

Fiston Ikwa Ndol Mbutiwi

A été évaluée par un jury composé des personnes suivantes

Helen Trottier
Président-rapporteur

Marie-Pierre Sylvestre
Directeur de recherche

Stéphane Buteau
Membre du jury

Celia Greenwood
Examineur externe

Résumé

Introduction

La randomisation mendélienne (RM) est une approche de plus en plus populaire dans les études observationnelles qui utilise des variants génétiques (habituellement des polymorphismes mononucléotidiques ou *single-nucleotide polymorphisms*, *SNPs*) associés à une exposition (hypothèse 1 ou pertinence) comme instruments pour estimer l'effet causal de cette exposition sur une issue, en assumant l'absence de confusion entre l'instrument et l'issue (hypothèse 2 ou indépendance) et l'absence d'un effet de l'instrument sur l'issue en dehors de l'exposition (hypothèse 3 ou restriction d'exclusion). Cependant, la validité des résultats de la RM est menacée par la pléiotropie, phénomène biologique par lequel un SNP affecte distinctement l'exposition et l'issue, qui est l'une des principales causes de violation de la restriction d'exclusion. Cette thèse examine certains défis méthodologiques pratiques de la RM relatifs à la vérification de la restriction d'exclusion et à la validité des résultats à travers trois principaux objectifs : 1) cartographier comment les chercheurs en RM préviennent, détectent et/ou contrôlent, et discutent des violations potentielles de la restriction d'exclusion dues notamment à la pléiotropie ; 2) évaluer la performance de la méthode basée sur la confusion positive, qui compare les estimés ponctuels de l'effet de l'exposition sur l'issue obtenus par la RM et par la régression conventionnelle, dans la détection des instruments invalides dans plusieurs contextes pratiques d'études de RM ; et 3) examiner l'impact des méthodes courantes de gestion de la médication antihypertensive dans les études de RM modélisant la pression artérielle (PA) sur l'estimation de l'effet causal et la détection des violations potentielles de la restriction d'exclusion.

Méthodes

Pour l'objectif 1, une revue de littérature de 128 études de RM ayant utilisé au moins un SNP sur le gène *FTO* (*fat mass and obesity-associated*) comme instrument de l'indice de masse corporelle (IMC) a été réalisée. La façon dont les auteurs préviennent, évaluent ou contrôlent, et discutent des violations potentielles de la restriction d'exclusion dues notamment à la pléiotropie a été examinée. Pour l'objectif 2, une étude de simulation statistique considérant des contextes d'études de RM utilisant comme instrument un SNP ou un score de risque génétique (*genetic risk score*, *GRS*), une issue continue ou binaire, dans des scénarios évaluant l'impact de la taille de l'échantillon et du

type de pléiotropie (indirect ou direct), a été réalisée. La performance de la méthode basée sur la confusion positive a été définie comme le pourcentage de jeux de données simulés dans lesquels la méthode détectait des instruments invalides. Pour l'objectif 3, une étude de RM de l'association entre l'IMC et la PA systolique (PAS) a été réalisée. Les méthodes de gestion de la médication antihypertensive examinées étaient : (i) pas de correction, (ii) inclure la médication dans les modèles comme une covariable d'ajustement, (iii) exclure de l'analyse les sujets traités aux antihypertenseurs, (iv) ajouter une valeur constante de 15 mm Hg aux valeurs mesurées de la PAS chez les sujets traités aux antihypertenseurs, et (v) utiliser comme issue un indicateur binaire de l'hypertension.

Résultats

Il existe une pléthore de méthodes utilisées dans les études de RM dont certaines peuvent être sous-optimales à prévenir, détecter ou contrôler le biais dû à l'inclusion des SNPs pléiotropiques. Les simulations statistiques montrent qu'en RM utilisant un SNP comme instrument, la méthode basée sur la confusion positive est performante à détecter l'invalidité de l'instrument lorsque la pléiotropie est directe plutôt qu'indirecte, indépendamment de l'issue, mais la performance de la méthode s'améliore avec l'augmentation de taille de l'échantillon. En revanche, la méthode est moins performante à détecter l'invalidité lorsque l'instrument est un GRS, mais sa performance augmente avec la proportion des SNPs invalides inclus dans le GRS. Enfin, les estimations de la RM varient énormément selon la stratégie de gestion de la médication antihypertensive choisie, contrairement à la détection des violations de la restriction d'exclusion qui n'en est pas affectée.

Conclusion

Cette thèse met de l'avant certaines difficultés méthodologiques dans les applications de la RM et l'importance de la triangulation de plusieurs méthodes dans la vérification des hypothèses de RM. Le champ de la RM est en plein essor, et des nouvelles méthodes sont souvent proposées. Il devient important non seulement de les évaluer, mais aussi d'en détailler l'utilisation et les hypothèses sous-jacentes pour une utilisation optimale en complément aux méthodes existantes.

Mots-clés : Randomisation mendélienne, pléiotropie, restriction d'exclusion, méthode de falsification, adiposité, indice de masse corporelle, pression artérielle, hypertension, médication antihypertensive, gène *FTO*.

Abstract

Introduction

Mendelian randomization (MR) is an increasingly popular technique in observational studies that uses genetic variants (usually single-nucleotide polymorphisms, SNPs) associated with an exposure (Assumption 1 or relevance) as instruments to estimate the causal effect of that exposure on an outcome, assuming no confounding between the instrument and the outcome (Assumption 2 or independence) and no effect of the instrument on the outcome outside of its association with the exposure (Assumption 3 or exclusion restriction). However, the validity of the MR results is challenged by pleiotropy, the biological phenomenon whereby a SNP distinctly affects the exposure and the outcome, which is one of the leading causes of violation of the exclusion restriction assumption. This thesis examines some practical MR methodological challenges related to the assessment of the exclusion restriction and the validity of MR results through three main objectives: 1) to examine how MR researchers prevent, detect, and/or control for, and discuss potential violations of the exclusion restriction due especially to pleiotropy; 2) to evaluate the performance of the leveraging positive confounding (LPC) method that compares the MR and the conventional point estimates in detecting invalid instruments in several practical MR settings; and 3) to examine the impact of commonly used methods of accounting for antihypertensive medication in MR studies modeling blood pressure (BP) on the estimation of the causal effect and the detection of potential violations of the exclusion restriction.

Methods

For Objective 1, a literature review of 128 MR studies that used at least one SNP in the fat mass and obesity-associated (*FTO*) gene as an instrument for body mass index (BMI) was conducted to examine how the authors prevent, detect, or control, and discuss potential violations of the exclusion restriction, especially due to pleiotropy. For Objective 2, a simulation study considering MR analyse settings using single SNP or genetic risk score (GRS) as an instrument, continuous or binary outcome, in scenarios evaluating the impact of sample size and type of pleiotropy (indirect vs. direct) was performed. The performance of the LPC method was assessed as the percentage of simulated datasets in which the LPC method detected invalid instruments. For Objective 3, an MR study of the association between BMI and systolic BP (SBP) was performed. The methods for

accounting for antihypertensive medication examined were: (i) no adjustment, (ii) include medication in the models as an adjustment covariate, (iii) exclude from the analysis subjects treated with antihypertensive medication, (iv) add a constant value of 15 mm Hg to the measured values of SBP in subjects using antihypertensive medication, and (v) use as outcome a binary indicator of hypertension.

Results

There exists a plethora of methods used in MR studies, some of which may be suboptimal for preventing, detecting, or controlling for bias due to the inclusion of pleiotropic SNPs. Statistical simulations show that in MR using single SNP as an instrument, the LPC method performs better at detecting invalidity of the instrument when the pleiotropy is direct than indirect, regardless of the outcome, although the performance of the method improves with increasing sample size. In contrast, the method performs less well in detecting invalidity when the instrument is a GRS, but its performance increases with the proportion of invalid SNPs included in the GRS. Finally, MR estimates change greatly depending on the chosen strategy of accounting for antihypertensive medication in contrast to the detection of exclusion restriction violations which is not impacted.

Conclusion

This present thesis highlights some of the methodological challenges in MR applications and the importance of triangulating multiple methods when assessing the MR assumptions. The MR field is booming, and new methods are often proposed. Therefore, it is important to evaluate these methods as well as to detail their application and underlying assumptions for optimal use as a complement to existing methods.

Keywords: Mendelian randomization, pleiotropy, exclusion restriction, falsification method, adiposity, body mass index, blood pressure, hypertension, antihypertensive medication, *FTO* gene.

Table des matières

Résumé	3
Abstract	5
Table des matières	7
Liste des tableaux	11
Liste des figures	13
Liste des sigles et abréviations	14
Remerciements	17
Chapitre 1 – Introduction	19
1.1. Problématique	19
1.2. Objectifs de recherche.....	23
1.2.1. Objectif général.....	23
1.2.2. Objectifs spécifiques.....	23
1.3. Plan de la dissertation.....	24
Chapitre 2 – Revue de la littérature	25
2.1. Randomisation mendélienne	25
2.1.1. Définition, historique, approche.....	25
2.1.2. Types d’instruments en RM et méthodes de sélection des SNPs	28
2.1.2.1. Types d’instruments en RM	28
2.1.2.2. Méthodes de sélection des SNPs	30
2.1.3. Hypothèses de la RM	30
2.1.3.1. Hypothèse 1 ou ‘pertinence’ (<i>relevance</i>).....	31
2.1.3.2. Hypothèse 2 ou ‘indépendance’ (<i>independence</i>).....	32
2.1.3.3. Hypothèse 3 ou ‘restriction d’exclusion’ (<i>exclusion restriction</i>)	34
2.1.3.4. Hypothèse 4 ou condition d'identification de l'estimation ponctuelle	35
2.1.4. Estimation de l’effet causal en RM.....	37
2.1.4.1. Nature de données exploitées en RM	37
2.1.4.2. Méthodes d’estimation de l’effet causal en RM.....	37

2.1.5. Considérations diverses dans l'interprétation des résultats de la RM.....	41
2.1.5.1. Hypothèse d'équivalence gène-environnement.....	41
2.1.5.2. Linéarité.....	41
2.1.5.3. RM et exposition variant dans le temps.....	42
2.1.5.4. Autres types de biais (collision, biais de sélection, erreur de classification)	42
2.1.5.5. Causalité inverse.....	45
2.1.5.6. RM et exposition binaire dichotomisée.....	46
2.1.5.7. RM et épigénétique.....	47
2.1.5.8. RM et taille de l'échantillon	47
2.2. Pléiotropie en RM	48
2.2.1. Définition et types de pléiotropie en RM.....	48
2.2.2. Source et ampleur de pléiotropie dans le génome humain.....	50
2.2.3. Impact et défi de la gestion de la pléiotropie en RM	50
2.2.4. Prévention, détection ou correction de la pléiotropie en RM	51
2.2.5. Illustration de la problématique de pléiotropie en RM : gène <i>FTO</i> comme instrument de l'adiposité.....	51
2.2.5.1. Adiposité, gène <i>FTO</i> et association <i>FTO</i> -adiposité.....	51
2.2.5.2. Mécanismes biologiques de <i>FTO</i> sur l'adiposité	52
2.2.5.3. Suspicion de pléiotropie de <i>FTO</i>	53
Chapitre 3 – Méthodes.....	55
3.1. Méthodes pour l'étude 1 (revue de littérature).....	55
3.1.1. Recherche documentaire	55
3.1.1.1. Source et opérationnalisation de la recherche d'articles.....	55
3.1.1.2. Critères d'inclusion des études	58
3.1.1.3. Critères d'exclusion des études	58
3.1.1.4. Processus de sélection des articles.....	58
3.1.2. Extraction des données.....	59
3.1.2.1. Données générales	59

3.1.2.2. Approches pour prévenir la pléiotropie (sélection et combinaison des instruments)	60
3.1.2.3. Approches pour détecter et/ou gérer la pléiotropie (analyses complémentaires) ...	60
3.1.2.4. Discussion des résultats en tenant compte de l'évaluation de la pléiotropie	61
3.1.2.5. Processus d'extraction et de validation des données	61
3.1.3. Analyse des données	61
3.2. Méthodes pour l'étude 2 (étude de simulation statistique)	61
3.2.1. Description de la méthode basée sur la confusion positive	62
3.2.1.1. Approche originale	62
3.2.1.2. Extension de la méthode basée sur la confusion positive aux issues binaires.....	63
3.2.2. Simulations.....	63
3.2.2.1. Cas 1 : un SNP comme instrument et issue continue	64
3.2.2.2. Cas 2 : un SNP comme instrument et issue binaire	65
3.2.2.3. Cas 3 : GRS comme instrument et issue continue.....	65
3.2.2.4. Cas 4 : GRS comme instrument et issue binaire	66
3.2.3. Comparaison de la méthode basée sur la confusion positive avec des méthodes alternatives dans l'évaluation de la validité des instruments dans des études de RM sélectionnées	67
3.3. Méthodes pour l'étude 3 (analyse de RM).....	67
3.3.1. Population de l'étude	68
3.3.2. Génotypage	68
3.3.3. Variables	69
3.3.3.1. Exposition.....	69
3.3.3.2. Issues	69
3.3.3.3. Instruments	69
3.3.3.4. Covariables	70
3.3.4. Analyses statistiques	72
3.3.4.1. Analyses principales	72
3.3.4.2. Évaluation des hypothèses de la RM.....	73

3.3.5. Considérations éthiques	75
Chapitre 4 – Résultats	76
4.1. Article 1	77
4.2. Article 2	153
4.3. Article 3	180
Chapitre 5 – Discussion et conclusion	221
5.1. Discussion des principaux résultats de la thèse	222
5.1.1. Prévention du biais dû aux SNPs pléiotropiques dans les études de RM	222
5.1.2. Évaluation de la plausibilité de l’hypothèse de restriction d’exclusion dans les études de RM	223
5.1.3. Performance de la méthode basée sur la confusion positive à détecter la pléiotropie des instruments dans les études de RM	225
5.1.4. Impact des méthodes de gestion de la médication antihypertensive sur l’estimation de l’effet causal dans les études de RM modélisant la PA	226
5.1.5. Utilisation spécifique de <i>FTO</i> comme instrument de l’IMC dans nos études	228
5.2. Impact de la thèse en santé publique	229
5.3. Forces et limites de notre recherche	230
5.4. Conclusion	237
Références bibliographiques	238
Annexes	262
Annexe 1. Certificats d’éthique	262
Annexe 2. Autorisation de BMJ Publishing Group Ltd	266
Annexe 3. Auteurs des articles de la revue de littérature contactés	268

Liste des tableaux

Article 1

Table 1. Characteristics of included studies.....	101
Table 2. Approaches cited for preventing pleiotropy in the selection of IVs.....	102
Table 3. Methods cited for detection or accounting the violation of the independence and/or restriction exclusion assumptions.....	103
Table 4. Discussion of the independence and/or exclusion restriction assumptions.....	105
Table S1. List of the 128 studies included.....	111
Table S2. Frequency of use of 31 <i>FTO</i> single-nucleotide polymorphisms as instruments for body mass index.....	115
Table S3. Characteristics of included studies.....	117
Table S4. Approaches cited for preventing pleiotropy in the selection of IVs.....	119
Table S5. Methods cited for detection or accounting the violation of the independence and/or restriction exclusion assumptions.....	120
Table S6. Summary description of methods cited for detection or accounting the violation of the independence and/or restriction exclusion assumptions in the review.....	125
Table S7. Discussion of the independence and/or exclusion restriction assumptions.....	130
Table S8. List of 21 studies that suspected IV pleiotropy/invalidity.....	132

Article 2

Table 1. Simulation Cases 1 (left panel) and 2 (right panel): Performance of the LPC method in assessing the validity of a single SNP as IV in three scenarios allowing for different sample sizes and a null causal effect ($\beta_{YX} = 0$).....	171
Table 2. Simulation Cases 3 (left panel) and 4 (right panel): Performance of the LPC method in assessing the validity of the genetic risk score (GRS) as IV in three scenarios allowing different sample sizes and proportion of pleiotropic SNPs in the GRS, and a null causal effect ($\beta_{YX} = 0$).....	172
Table 3. Comparison of the LPC method with alternative approaches in assessing the validity of the instruments in selected MR studies.....	174
Table S1. Summary of parameters used in the simulations.....	176

Table S2. Simulation Cases 1 (left panel) and 2 (right panel): Performance of the LPC method in assessing the validity of a single SNP as IV in three scenarios allowing different sample sizes, and positive and negative causal effect.....	177
Table S3. Simulation Cases 3 (left panel) and 4 (right panel): Performance of the LPC method in assessing the validity of the genetic risk score (GRS) as IV in three scenarios allowing different sample sizes and proportion of pleiotropic SNPs in the GRS, and positive and negative causal effect.....	178
 Article 3	
Table 1. Baseline characteristics of participants by study inclusion status, CLSA Comprehensive cohort, 2011-2018.....	200
Table 2. Conventional and MR estimates of the association between body mass index (BMI) with blood pressure outcomes in different scenarios of accounting for antihypertensive medication effects, CLSA Comprehensive cohort, 2011-2018.....	201
Table 3. Assessment of the validity of the BMI-associated instruments in different scenarios of accounting for antihypertensive medication using MR-Egger intercept test and the leveraging positive confounding method, CLSA Comprehensive cohort, 2011-2018.....	202
Table 4. Associations between the instruments and blood pressure outcomes in different scenarios of accounting for antihypertensive medication, antihypertensive medication covariate, and potential confounders, CLSA Comprehensive cohort, 2011-2018.....	203
Table S1. Associations of 97 BMI-associated SNPs with BMI, CLSA Comprehensive cohort, 2011-2018 (n = 20,430).....	206
Table S2. MR estimates of the association between BMI with blood pressure outcomes in different scenarios of accounting for antihypertensive medication using the GRS based on 97 BMI-associated SNPs as instrument, CLSA Comprehensive cohort, 2011-2018.....	209
Table S3. Associations of 87 BMI-associated SNPs included in the genetic risk score with BMI and with SBP/hypertension, CLSA Comprehensive cohort, 2011-2018 (n = 20,430)..	210
Table S4. Associations of 87 BMI-associated SNPs included in the GRS with BMI and with measured SBP both adjusted for antihypertensive medication covariate, CLSA Comprehensive cohort, 2011-2018 (n = 20,430).....	214
Table S5. Associations of 87 BMI-associated SNPs included in the GRS with BMI and with measured SBP after exclusion of participants using antihypertensive medication, CLSA Comprehensive cohort, 2011-2018 (n = 14,337).....	217

Liste des figures

Thèse

Figure 1. Graphique acyclique orienté de la randomisation mendélienne.....	19
Figure 2. Parallélisme entre le plan d'étude de randomisation mendélienne et celui d'un essai clinique randomisé.....	27
Figure 3. Illustration de trois types d'instruments utilisés en randomisation mendélienne..	29
Figure 4. Illustration de l'hypothèse 1 de la randomisation mendélienne.....	31
Figure 5. Illustration de l'hypothèse 2 de la randomisation mendélienne.....	33
Figure 6. Illustration de l'hypothèse 3 de la randomisation mendélienne.....	35
Figure 7. Illustration de la collision en randomisation mendélienne.....	43
Figure 8. Illustration de la pléiotropie en randomisation mendélienne.....	49

Article 1

Figure 1. Flowchart of selection of studies includes in the review.....	106
Figure 2. Body mass index instrumental variables (IVs) used in the main analysis by article publication year.....	107

Article 3

Figure S1. Flowchart of selection of study participants, CLSA Comprehensive cohort, 2011-2018.....	205
---	-----

Liste des sigles et abréviations

ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
ATE	Average treatment effect
CAUSE	Causal analysis using summary effect estimates
ÉLCV/CLSA	Étude longitudinale canadienne sur le vieillissement/Canadian longitudinal study on aging
GLM	Generalized linear model
GRS	Genetic risk score
FTO	Fat mass and obesity-associated
GWAS	Genome-wide association study
HLA	Human Leukocyte Antigen
ICBP	International Consortium for Blood Pressure
IMC/BMI	Indice de masse corporelle/Body mass index
IVW	Inverse-variance weighted
InSIDE	Instrument Strength Independent of Direct Effect
LATE	Local average treatment effect
LPC	Leveraging positive confounding
MRMix	Mendelian randomization analysis using mixture-model
OLS	Ordinary least squares
PA/BP	Pression artérielle/Blood pressure
PAD/DPB	Pression artérielle diastolique/Diastolic blood pressure
PAS/SBP	Pression artérielle systolique/Systolic blood pressure
PASE	Physical Activity Scale for the Elderly
RM/MR	Randomisation mendélienne/Mendelian randomization
MR-PheWAS	Mendelian randomization-phenome-wide association study
RUNX1T1	Runt-related transcription factor 1
SNP	Single-nucleotide polymorphism

SRSF2	Serine and arginine rich splicing factor 2
VI/IV	Variable instrumentale/Instrumental variable
2SLS	Two stage least squares
2SPS	Two-stage predictor substitution
2SRI	Two-stage residual inclusion

*« Mon âme, bénis l'Éternel, et n'oublie aucun de ses bienfaits! »
(Psaumes 103:2)*

*À ma bienaimée épouse !
Mimitha Kangufu*

*À nos merveilleux enfants !
Pavi, Daniela, Kevin, Mifi et Raphaël*

Merci de tout cœur pour votre amour et votre soutien.

Remerciements

Du plus profond de mon cœur, je voudrais remercier ma directrice de recherche, Professeure Marie-Pierre Sylvestre, une scientifique aguerrie, et une personne humaine exceptionnelle, qui a cru en moi dès notre première rencontre. Votre encadrement scientifique et votre soutien inconditionnel tant financier qu'humain ont été déterminants pour l'aboutissement de ce projet de recherche. Ma thèse est le fruit d'un travail rigoureux et ardu de recherche et d'application, mais aussi de remise en question, bref, d'une construction de savoir, brique après brique. J'ai apprécié à juste titre, non seulement l'immensité de vos connaissances scientifiques et votre passion pour la recherche, mais aussi vos qualités d'humilité et de persévérance dans la quête de l'excellence. Vous avez toujours été là quand j'avais besoin de vous. Vous avez instauré autour de moi un environnement d'encadrement sain, convivial et stimulant. Quelle grâce de vous avoir eue comme directrice de recherche ! Merci infiniment pour tout ce que vous avez fait pour moi.

Mes remerciements s'adressent également aux différentes organisations qui m'ont apporté du soutien financier par l'attribution des bourses durant ma formation : le Fonds de recherche du Québec - Santé (FRQS) pour la bourse de formation de doctorat, la Faculté des études supérieures et postdoctorales (FESP) de l'Université de Montréal pour la bourse de la réussite étudiante, l'École de Santé Publique de l'Université de Montréal (ESPUM) pour les bourses de soutien financier, le Centre Hospitalier de l'Université de Montréal (CRCHUM, bourse de la directrice de recherche).

Je voudrais exprimer toute ma gratitude au CRCHUM qui m'a offert le cadre et l'environnement idéal pour mes travaux de recherche, et aussi aux responsables de l'Étude Longitudinale Canadienne sur le Vieillissement (ÉLCV) qui ont mis gratuitement à ma disposition les bases de données pour la réalisation du troisième article de ma thèse. Je remercie tous les participants de l'ÉLCV pour leur consentement à l'utilisation de leurs données.

J'exprime ma gratitude envers l'Université de Montréal, l'ESPUM et tous mes Professeurs et encadreurs pour la qualité de la formation reçue. Je remercie tous les membres de notre dynamique équipe au CRCHUM, avec une mention spéciale aux Professeures Jennifer O'Loughlin et Isabelle Doré.

Je remercie le jury pour les remarques et suggestions formulées qui ont amélioré ce travail. Je tiens à saluer la contribution substantielle de tous les co-auteurs des articles inclus dans ma thèse. Je remercie particulièrement M. David Boucher-Roy pour son précieux soutien dans la programmation des simulations statistiques incluses dans le deuxième article de la thèse. Mes sincères remerciements s'adressent également au Professeur Bernard-Simon Leclerc (ESPUM), à M. Georges-Charles Thiebaut (chargé du cours SPU7007, automne 2015, ESPUM) et au Professeur Dr Philippe Meert (Université Catholique de Louvain, en Belgique) qui ont cru en moi et ont recommandé mon admission au doctorat à l'UdeM.

Je voudrais exprimer ma reconnaissance envers l'Université de Kikwit (UNIKIK) en République Démocratique du Congo, le comité de gestion de l'UNIKIK et le Recteur, Professeur Alphonse Kapumba Nkodi, la faculté de médecine de l'UNIKIK et son Doyen, Professeur Dr Jean-Bosco Kasiam, pour leur soutien.

Je salue la mémoire de Professeur Dr François Lepira Bompeka, ancien Doyen de la faculté de médecine de l'UNIKIK, qui nous a quitté en 2020. Le Professeur Lepira était un scientifique hors pair, épris d'humilité et d'humanisme ; il a été un mentor et m'a encadré dans ma pratique de médecine et dans l'enseignement. Paix à son âme ! Je remercie de tout cœur le Professeur Dr Jean-Bosco Kasiam pour sa contribution déterminante dans ma formation. Je pense aussi aux Professeurs Drs Jean-Marie Kayembe, Nazaire Nseka, Ernest Sumaili, Justine Bukabau, Jean Robert Makulo, Vieux Mokoli, Jean-Jacques Malemba, Antoine Molua, à Dre Angèle Mbongo, ainsi qu'à tous mes encadreurs aux Cliniques Universitaires de Kinshasa.

À mes parents, Gilbert Ikwa Mbutiwi et Béatrice Kasongo Kunumakana ; à mes frères et sœurs, Michel, Chantal, Célestin, Pablo, Taty et Nais ; à mes oncles, tantes, cousins, cousines ; au Rvd Père Augustin Ankwanda ; à la famille de Dr Patty Makaula ; à ma belle-famille, spécialement ma belle-mère, Jeannine Munganga ; à tous, je dis merci pour tout.

Enfin, je voudrais remercier tous les amis et connaissances qui m'ont soutenu d'une manière ou d'une autre. Je pense particulièrement à Espérant et Rose Tuta, et aussi à Pascal Dubé, Andrew Papadopoulos, Nicholas Major, ainsi qu'aux autres membres de nos équipes de Hockey Saint-Laurent qui ont accompagné mon fils dans ses parties de hockey durant cette saison 2022-2023 alors que j'étais occupé à finaliser ma thèse.

Chapitre 1 – Introduction

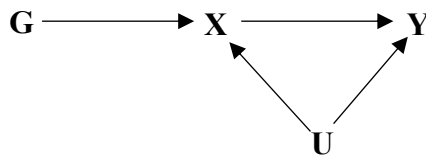
1.1. Problématique

L'un des problèmes majeurs menaçant la validité interne des études épidémiologiques observationnelles est le biais dû à des facteurs de confusion non mesurés.¹ Ce problème est géré dans les essais cliniques randomisés grâce au processus de randomisation qui permet de balancer les groupes comparés pour l'ensemble des facteurs de confusion mesurés et non mesurés. Mais pour des raisons éthiques et pratiques, les essais cliniques randomisés ne sont toujours pas réalisables.^{2,3} Des études observationnelles doivent donc être menées, et dans bien des cas, certains facteurs de confusion ne sont pas mesurés, voire ne sont pas connus. Les méthodes de variables instrumentales, qui incluent la randomisation mendélienne (RM), comptent parmi les alternatives méthodologiques les plus utilisées pour gérer la confusion par des facteurs non mesurés affectant les études observationnelles.⁴

L'approche de RM utilise un variant génétique associé à une exposition (hypothèse 1 de la RM ou pertinence) comme un instrument pour estimer l'effet causal de cette exposition sur une issue (**Figure 1**), en assumant l'absence de confusion entre l'instrument et l'issue (hypothèse 2 de la RM ou indépendance) et l'absence d'un effet de l'instrument sur l'issue en dehors de l'exposition (hypothèse 3 de la RM ou restriction d'exclusion ou absence de pléiotropie).⁵⁻⁸

Figure 1

Graphique acyclique orienté de la randomisation mendélienne.



Note. G est l'instrument, X l'exposition, Y l'issue et U un facteur de confusion de l'association entre X et Y.

En dépit de l'essor important de la RM depuis sa présentation initiale en 2003,^{9,10} l'application pratique de cette approche est particulièrement défiée par l'impossibilité de démontrer formellement, avec les données sous analyse, la satisfaction des hypothèses d'indépendance et de restriction d'exclusion sus-évoquées,^{5,11} ce qui demeure une préoccupation vitale pour la validité des résultats. Plus spécifiquement, les violations de la restriction d'exclusion, dues principalement à la pléiotropie horizontale (phénomène biologique par lequel un variant génétique utilisé comme instrument affecte distinctement l'exposition et l'issue d'intérêt¹² – nous parlerons simplement de pléiotropie pour alléger le texte), représentent les principales menaces à la validité des résultats d'une analyse de RM.^{13,14}

La présente thèse s'intéresse à certains défis méthodologiques rencontrés en pratique lorsqu'on utilise de manière rigoureuse la méthode de RM, particulièrement en rapport avec la vérification de l'hypothèse de restriction d'exclusion et la validité des résultats.

Le premier volet de cette thèse cartographie les méthodes appliquées par les chercheurs en RM pour détecter des violations de la restriction d'exclusion et/ou d'en atténuer l'impact. En effet, différentes études sur l'utilisation de la RM¹⁵⁻²³ suggèrent que l'hypothèse de restriction d'exclusion n'est pas systématiquement évaluée ou discutée. Par exemple, Boef et al.²³ ont noté dans leur revue que seulement 111 sur 178 études (62,4 %) ont discuté de la plausibilité de la restriction d'exclusion. De plus, aucune étude n'a jusqu'à présent examiné comment, en pratique, les auteurs préviennent ou minimisent le biais potentiel lié à l'utilisation des instruments pléiotropiques dans les études de RM. Cette évaluation est importante car non seulement la pléiotropie est omniprésente dans le génome humain,^{24,25} mais aussi l'absence de pléiotropie dans une étude de RM ne peut être prouvée empiriquement.⁵ Aussi, même si plusieurs méthodes visant à suggérer une violation de la restriction d'exclusion et/ou à fournir des estimations de RM robustes à la pléiotropie ont été proposées et comparées,^{7,12,26-34} leur utilisation en pratique, notamment dans la triangulation des méthodes n'a pas été documentée dans les études. A la lumière des lignes directrices de la RM^{35,36} qui recommandent d'évaluer la robustesse des résultats à la violation des hypothèses de la RM, la première étude de notre thèse examine spécifiquement comment le biais potentiel dû à une violation de la restriction d'exclusion est géré dans la littérature.

Le deuxième aspect examiné dans la présente thèse concerne l'évaluation de la performance d'une méthode de détection de violation de la restriction d'exclusion proposée par Glymour et al.⁵ en

2012 dénommée “falsification des hypothèses de variable instrumentale (*instrumental variable*, IV) en se basant sur les hypothèses causales antérieures” que nous appellerons par la suite ‘méthode basée sur la confusion positive’.⁷

Cette méthode propose simplement de comparer les estimations ponctuelles de l’effet d’une exposition sur une issue continue obtenues par la RM et par la régression linéaire (méthode des moindres carrés ordinaires ou *ordinary least squares*, OLS), et de suspecter que l’instrument est invalide si l’estimation de RM est plus grande que l’estimation OLS, en assumant certains prérequis.⁵ Le principal avantage de cette méthode est qu’elle est simple et peut être appliquée par un simple lecteur d’une étude de RM en comparant simplement les coefficients publiés dans un article sans avoir à manipuler les données. Toutefois, même si la méthode a été utilisée comme telle dans l’évaluation des hypothèses de RM dans la littérature,³⁷⁻⁴⁰ sa performance n’a pas été étudiée de manière approfondie.

Peu de temps après la proposition de Glymour et al.,⁵ Burgess,⁴¹ utilisant des simulations statistiques, a suggéré que la méthode basée sur la confusion positive est adéquate pour détecter l’invalidité de l’instrument lorsque celui-ci affecte directement l’issue (pléiotropie directe), mais moins satisfaisante si l’instrument affecte l’issue via les facteurs de confusion de l’association entre l’exposition et l’issue (pléiotropie indirecte). En outre, la méthode a détecté des instruments invalides dans 11% des scénarios où tous les instruments simulés étaient pourtant valides (résultats faux positifs).⁴¹ Toutefois, la taille d’échantillon considérée par Burgess⁴¹ était relativement modeste pour les études de RM (n=2000), et les simulations n’ont pas été élargies à d’autres scénarios, comme ceux allouant une issue binaire (contexte dans lequel la méthode basée sur la confusion positive a déjà été appliquée³⁸), ou utilisant comme instruments des scores de risque génétique (*genetic risk scores*, GRS) qui sont d’usage courant dans les études de RM.⁴² D’où l’intérêt de notre deuxième étude de la thèse d’étendre l’évaluation de la performance de la méthode basée sur la confusion positive dans différentes applications pratiques de RM.

Le troisième volet de la thèse s’intéresse à l’impact des méthodes de prise en compte de la médication antihypertensive sur les résultats d’une analyse de RM modélisant la pression artérielle (PA). En effet, l’une des principales caractéristiques de la littérature relative à la modélisation de la PA en RM est la prise en compte incohérente de la variable médication antihypertensive (nous parlerons souvent simplement de ‘médication’ pour alléger le texte) dans les études. Par exemple,

sur les 10 études de RM de l'association entre l'indice de masse corporelle (IMC) et la PA identifiées dans notre récente revue de littérature (article 1 de la présente thèse),⁴² cinq études ont simplement modélisé la PA mesurée en ignorant la question de la médication,⁴³⁻⁴⁷ deux ont ajusté pour la médication comme une covariable dans les modèles,^{48,49} et les trois autres ont imputé la valeur de la PA mesurée chez les sujets traités avec les antihypertenseurs en ajoutant une valeur constante de PA systolique (PAS) et de PA diastolique (PAD) respectivement de 15 mm Hg et 10 mm Hg^{50,51} ou 10 mm Hg et 5 mm Hg.⁵²

La prise en compte correcte de la médication dans les études de la PA est primordiale car la PA mesurée chez les personnes traitées par des médicaments antihypertenseurs est probablement inférieure à la valeur qu'elle aurait été si ces personnes n'étaient pas traitées.^{53,54} Par conséquent, ignorer la médication et modéliser simplement la PA mesurée peut fausser l'estimation d'une association avec un facteur de risque.⁵³ Une telle distorsion est susceptible d'être importante considérant l'usage très courant des médicaments antihypertenseurs⁵⁵ dans un contexte de prévalence mondiale élevée de l'hypertension.^{56,57}

Dans la littérature, l'étude de simulations de Tobin et al.⁵³ a suggéré que l'imputation de la valeur mesurée de la PA chez les sujets traités aux antihypertenseurs, notamment en ajoutant une valeur constante de 15 mm Hg, était une meilleure méthode pour gérer la médication comparativement à d'autres stratégies comme exclure les sujets traités de l'analyse, ignorer la question de la médication ou ajuster les modèles pour la variable médication. Cependant, à ce jour, aucune étude n'a évalué l'impact de ces méthodes de prise en compte de la médication sur les résultats des études de RM modélisant la PA. Et, pourtant, les résultats des analyses par des régressions conventionnelles peuvent ne pas se transposer aux études de RM. Par exemple, les études de RM nécessitent généralement des échantillons de plus grande taille car elles utilisent, comme instruments, des variants génétiques qui n'expliquent souvent qu'une faible proportion de la variance de l'exposition.⁵⁸ Par conséquent, la méthode qui exclut de l'analyse les personnes traitées aux antihypertenseurs peut affecter les études de RM dans une plus large mesure que les analyses conventionnelles. Aussi, la variable médication peut être associée aux instruments utilisés en RM et ainsi entraîner une violation de l'hypothèse de restriction d'exclusion. Par exemple, le gène *FTO* (*fat mass and obesity-associated*), le plus fort signal associé avec l'IMC⁵⁹ et souvent utilisé comme instrument dans les études de RM de l'impact de l'adiposité sur la PA, a été associé à des problèmes

de santé mentale (par exemple, la démence)^{37,39,60} qui peuvent à leur tour affecter l'adhésion à la médication,^{61,62} ce qui créerait une association entre l'instrument (*FTO*) et l'issue (*PA*) indépendante de l'exposition (adiposité), violant ainsi la restriction d'exclusion. Même s'il existe, comme précédemment mentionné, des méthodes pouvant détecter des violations de la restriction d'exclusion dans les études de RM,⁴² on sait peu de choses sur leur performance en conjonction avec les méthodes de prise en compte de la médication.

1.2. Objectifs de recherche

1.2.1. Objectif général

La présente thèse avait pour objectif général d'examiner certains défis méthodologiques pratiques relatifs à la vérification de l'hypothèse de restriction d'exclusion et leur incidence sur la validité des résultats de la RM.

1.2.2. Objectifs spécifiques

Trois objectifs spécifiques ont été formulés, chacun ayant fait l'objet d'une étude spécifique :

1. Cartographier comment les chercheurs en RM préviennent, détectent et/ou contrôlent, et discutent des violations potentielles de la restriction d'exclusion dues notamment à la pléiotropie ;
2. Évaluer la performance de la méthode basée sur la confusion positive à détecter des instruments invalides dans divers contextes pratiques d'études de RM ;
3. Examiner l'impact de méthodes courantes de gestion de la médication antihypertensive dans les études de RM modélisant la *PA* sur l'estimation de l'effet causal et la détection des violations potentielles de la restriction d'exclusion.

1.3. Plan de la dissertation

La présente thèse est une thèse par articles, structurée autour de cinq chapitres. Outre ce premier chapitre d'introduction résumant les principaux défis abordés et les objectifs poursuivis, le deuxième chapitre présente une recension de la littérature sur l'approche de la RM dans ses aspects historiques et méthodologiques, incluant la problématique de la pléiotropie en RM. Le troisième chapitre décrit l'approche méthodologique utilisée pour atteindre l'objectif de chaque étude. Le quatrième chapitre présente les résultats de chacune des trois études sous forme d'articles. Et enfin le cinquième chapitre est une discussion générale des résultats des trois études, en épinglant la contribution originale de la thèse mais aussi des limites de notre démarche, ainsi que des perspectives en vue des recherches futures.

Chapitre 2 – Revue de la littérature

Ce chapitre fait une recension de la littérature sur la méthode de RM. La première partie décrit la RM en épinglant la définition, l’historique et l’approche, les types d’instruments utilisés et les méthodes de sélection des SNPs, les hypothèses de la RM, les méthodes d’estimation de l’effet causal ainsi que diverses considérations utiles dans l’interprétation des résultats de la RM. La deuxième partie approfondie particulièrement la problématique de la pléiotropie en RM, une cause importante d’invalidation des résultats et qui est abordée dans les trois articles de la présente thèse. Plus spécifiquement, cette deuxième partie décrit les différents types de pléiotropie en RM, documente les sources et l’ampleur de la pléiotropie dans le génome humain, présente l’impact de la pléiotropie et le défi de sa gestion dans une étude de RM. Elle présente aussi les méthodes de prévention, détection ou correction de la pléiotropie en RM qui font l’objet d’un examen plus approfondi dans l’article 1 de cette thèse qui est une revue de littérature. Enfin, cette deuxième partie illustre la problématique de la pléiotropie en pratique en considérant l’utilisation du gène *FTO* comme un instrument de l’adiposité en RM.

2.1. Randomisation mendélienne

2.1.1. Définition, historique, approche

La RM est une approche de plus en plus utilisée dans les études épidémiologiques observationnelles⁹; elle exploite les techniques de VI, communément utilisées en économétrie,^{63,64} pour inférer une association causale entre une exposition (facteur de risque non génétique) et une issue en utilisant un variant génétique comme un instrument ou VI (**Figure 1**).⁶⁵⁻⁶⁷ La RM utilise donc un variant génétique associé à une exposition comme un substitut de cette exposition dans le but d’estimer la composante causale (‘non confondue’) de l’association observée entre l’exposition et l’issue.^{66,68-70} En tant que telle, la méthode de RM peut être doublement utile : (i) apporter la preuve de l’existence d’une association causale entre une exposition et une issue et (ii) permettre une estimation précise de l’ampleur de cet effet.⁷¹

La RM est le fruit d'une évolution méthodologique depuis l'approche initiale proposée par Katan en 1986 qui n'exploitait pas les techniques de VI à proprement parlé.^{72,73} L'approche de Katan consistait simplement à utiliser l'évidence d'une association entre le génotype (Apolipoprotéine E) et l'issue (risque de cancer) pour tester la présence d'un effet causal d'un facteur de risque modifiable (cholestérol sérique bas) sur l'issue, étant donné l'existence d'une association entre ce facteur de risque et le génotype dans la population.⁷⁴⁻⁷⁶ D'ailleurs, certaines études ont attribué le terme RM à des analyses de l'association entre le génotype et l'issue d'intérêt.^{77,78}

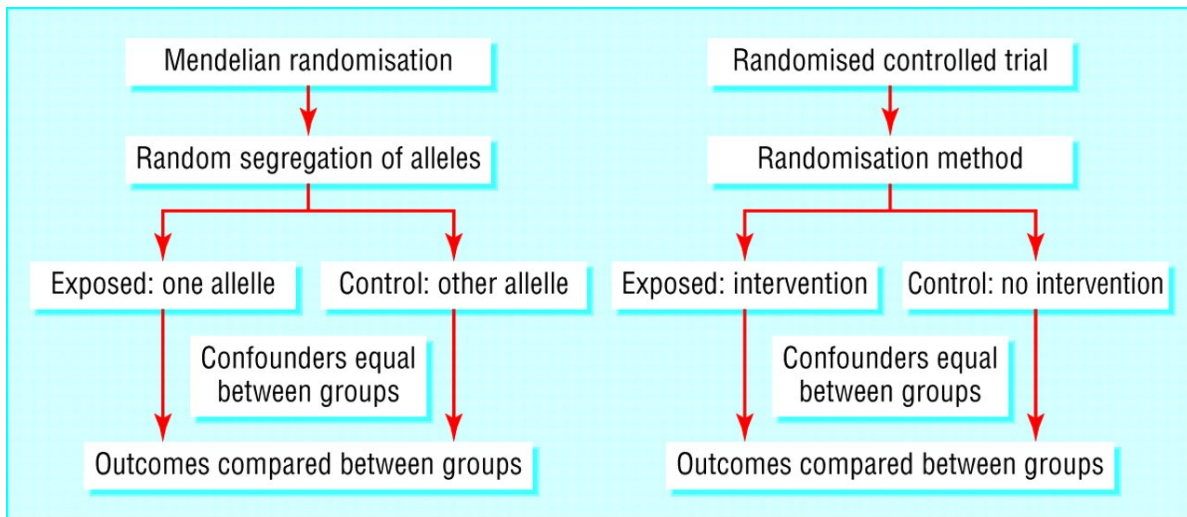
C'est véritablement en 1991 que le terme RM a été utilisé pour la première fois par Gray et Wheatley⁷⁹ dans le cadre d'un essai clinique visant à évaluer la survie des patients recevant une greffe allogénique de la moelle osseuse par rapport à la chimiothérapie dans le traitement de la leucémie. La comparaison des deux groupes de traitement était difficile du fait qu'il n'était pas éthique de 'refuser' à un patient de bénéficier de la greffe de la moelle osseuse lorsqu'un donneur était disponible. Ainsi, ces chercheurs ont résolu de comparer la survie des patients ayant reçu une greffe et qui avait un frère ou une sœur HLA (*Human Leukocyte Antigen*) compatible, de celle des patients ayant reçu une greffe et qui n'avaient pas un frère ou une sœur HLA compatible.^{70,79,80} Ces auteurs assumaient que le statut HLA, qui est déterminé par un assortiment aléatoire des gènes lors de la gamétogénèse et de la conception, ne serait pas relié à des facteurs de confusion.⁸⁰

L'approche actuelle de la RM résulte donc de l'évolution de l'épidémiologie génétique et a été initialement exposée de manière détaillée par Smith et Ebrahim en 2003.¹⁰ Elle implique l'utilisation des variants génétiques, habituellement des SNPs, comme instruments ou substituts de l'exposition.^{65,66,81} La RM tire son essence de la ségrégation aléatoire des allèles/génotypes transférés des parents à la progéniture à la conception (2^{ème} loi de Mendel),^{80,82} conférant à la RM une certaine protection contre le biais dû à des facteurs de confusion. En effet, puisque le génotype est fixé dès la conception et précède donc chronologiquement les variables observées, il est plus difficile que des facteurs de risque post-conception causent le génotype (protection contre la causalité inverse) ou causent à la fois le génotype et l'issue (protection contre la confusion).^{67-69,80} Aussi, l'allocation des allèles au cours de la méiose se faisant de manière aléatoire fait que les SNPs ne sont souvent pas influencés par des facteurs de confusion autre que l'ascendance.⁶⁷ Cette protection de la RM contre la confusion grâce à l'utilisation des SNPs comme instruments a des similitudes avec celle que procure le processus de randomisation dans les essais cliniques

randomisés, laissant entrevoir un certain parallélisme entre la RM et les essais cliniques randomisés (Figure 2).⁸³

Figure 2

Parallélisme entre le plan d'étude de randomisation mendélienne et celui d'un essai clinique randomisé.



Note. Tirée de Smith et Ebrahim.⁸⁴ Figure reproduite avec la permission de BMJ Publishing Group Ltd.

L'approche de la RM connaît un essor important et son utilisation a explosé passant de 2 publications en 2003 à plus de 800 en 2020.^{9,85} L'approche s'est de plus en plus complexifiée, par exemple, elle peut permettre de décomposer l'effet causal total d'une exposition sur une issue en estimant les portions directe et indirectes via d'autres phénotypes intermédiaires dans un devis de RM multivariée,^{86,87} comme on le ferait dans une analyse de médiation traditionnelle. Des lignes directrices encadrant l'utilisation de la méthode de la RM ont été édictées et sont régulièrement mises à jour avec la contribution de la communauté scientifique.^{35,85}

2.1.2. Types d'instruments en RM et méthodes de sélection des SNPs

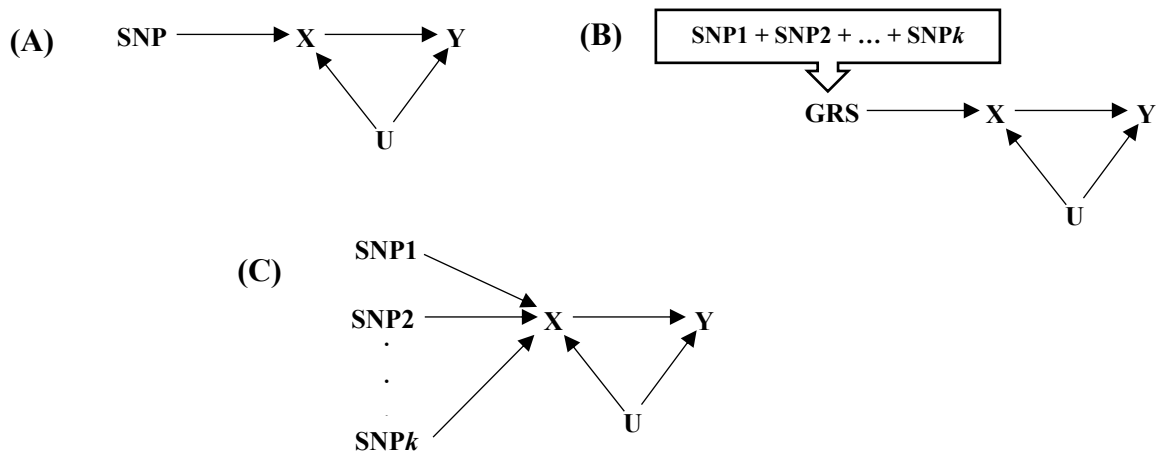
2.1.2.1. Types d'instruments en RM

Comme précédemment évoqué, la RM utilise habituellement des SNPs comme des instruments.^{81,88} Les SNPs sont des variations de la séquence de l'ADN (acide désoxyribonucléique) les plus simples et les plus communes, et se rencontrent en moyenne toutes les 1000 bases dans le génome.^{89,90} Un SNP consiste en une substitution de l'une ou l'autre paire de bases nucléiques à un emplacement particulier du génome (appelé *locus*).⁸⁹ Cette substitution peut consister en un remplacement d'une base purique (Adénine (A) ou Guanine (G)) par une autre base purique ou d'une base pyrimidique (Thymine (T) ou Cytosine (C)) par une autre base pyrimidique ; on parle de *transition*. Si par contre une base purique est remplacée par une base pyrimidique ou vice-versa, on parle de *transversion*.⁹⁰ La plupart des SNPs comportent deux allèles (on parle de SNPs bi-alléliques) correspondants aux deux différentes bases nucléiques d'un locus⁸⁹ : l'allèle dit majeur ou sauvage, habituellement le plus fréquent, et l'allèle dit mineur ou de risque.^{81,89}

La **Figure 3** illustre les différents types d'instruments utilisés en RM. On peut utiliser un seul SNP (**Figure 3A**) ou plusieurs SNPs indépendants comme instruments. Dans le cas où plusieurs SNPs indépendants sont sélectionnés, ceux-ci peuvent être utilisés sous-forme de GRS (**Figure 3B**) ou d'instruments multiples (**Figure 3C**).⁹¹⁻⁹³ Un GRS est une combinaison de plusieurs SNPs indépendants en une variable unique correspondant à la somme des allèles de risque des SNPs pour chaque individu pondérée ou non par les coefficients des associations SNPs–exposition.^{81,93,94} Les instruments multiples réfèrent à l'utilisation d'au moins deux SNPs indépendants comme des instruments distincts dans un seul modèle ou à la combinaison de ≥ 2 effets causaux estimés chacun en utilisant un SNP unique en un seul effet causal en utilisant des techniques de méta-analyse.⁹³

Figure 3

Illustration de trois types d'instruments utilisés en randomisation mendélienne.



Note. SNP = *single-nucleotide polymorphism*, GRS = *genetic risk score*, X est l'exposition, Y est l'issue et U un facteur de confusion de l'association entre X et Y. (A) Un seul SNP est utilisé comme instrument. (B) Plusieurs SNPs indépendants (1 à k) sont préalablement additionnés pour constituer un GRS qui est ensuite utilisé comme un seul instrument. (C) Plusieurs SNPs indépendants (1 à k) sont utilisés chacun comme un instrument distinct de X pour estimer un seul effet causal de X sur Y.

Un SNP unique est facile à utiliser. Mais, étant donné que les effets génétiques sur des phénotypes sont souvent faibles,⁵⁸ utiliser un seul SNP qui n'explique qu'une faible variation de l'exposition peut être problématique en RM (voir sous-section 2.1.3.1 sur l'hypothèse 1 de la RM).⁹⁵ Le GRS s'utilise facilement comme un seul SNP, mais contrairement à celui-ci, le GRS explique plus de variabilité de l'exposition et augmente donc la puissance statistique.^{58,93} L'utilisation des SNPs sous-forme d'instruments multiples, contrairement à celle d'un seul instrument (un SNP ou un GRS) facilite l'évaluation de la plausibilité des hypothèses de la RM par l'utilisation des méthodes robustes qui requièrent d'avoir plusieurs instruments indépendants (voir article 1 de la présente thèse pour un inventaire détaillé de ces méthodes).^{35,96,97}

2.1.2.2. Méthodes de sélection des SNPs

La littérature décrit deux principales approches de sélection des SNPs à utiliser comme instruments en RM, l'approche biologique et l'approche statistique.³⁵ L'approche biologique consiste à sélectionner les SNPs dans la région de gène qui code spécifiquement pour l'exposition. Cela requiert d'avoir une bonne connaissance du mécanisme biologique spécifique des gènes. L'approche biologique est souvent de mise pour des expositions de nature protéique qui disposent d'un gène codant proximal spécifique bien connu.^{35,98} Cette approche a l'avantage d'augmenter la crédibilité de la plausibilité des hypothèses de la RM, mais si les SNPs provenant d'une même région du gène sont corrélés, l'application des méthodes robustes serait limitée dans la mesure où celles-ci assument l'indépendance entre les SNPs.³⁵ De plus, dans plusieurs situations, la connaissance de la fonction biologique de gènes n'est toujours pas disponible,^{98,99} et les chercheurs recourent bien souvent à l'approche statistique qui consiste à sélectionner comme instruments les SNPs qui ont une association statistique robuste (souvent définie par une valeur de $p < 5 \times 10^{-8}$) avec l'exposition, habituellement dans les études d'association pangénomique (*genome-wide association studies*, GWAS).³⁵ L'approche statistique est la plus utilisée en pratique, particulièrement dans les études de RM où l'exposition est un trait polygénique (ex. IMC).⁴² Toutefois, sélectionner des SNPs sur la seule base d'une association statistique avec l'exposition augmente la probabilité d'inclusion des SNPs pléiotropiques (c.à.d. affectant l'issue en dehors de l'exposition) susceptibles d'invalider les résultats de la RM (voir section 2.2 sur la pléiotropie).^{28,100}

Les lignes directrices en RM suggèrent également une approche nuancée consistant à débiter la sélection des SNPs par l'approche statistique, puis exclure par la suite les SNPs connus ou suspectés de pléiotropie pour l'issue d'intérêt avant de procéder à l'analyse principale.³⁵

2.1.3. Hypothèses de la RM

L'interprétation des résultats de la RM dépend d'au moins quatre hypothèses.⁶⁻⁸ Les trois premières hypothèses sont souvent qualifiées d'hypothèses principales¹⁰¹ et sont énoncées comme suit : (i) l'instrument doit être associé à l'exposition (*relevance assumption*), (ii) il ne doit exister aucune cause commune entre l'instrument et l'issue (*independence assumption* ou *exchangeability assumption*) et (iii) l'instrument affecte l'issue exclusivement via l'exposition (*exclusion restriction assumption*).^{6,7,34,88,102} Ces trois hypothèses permettent de définir qu'un

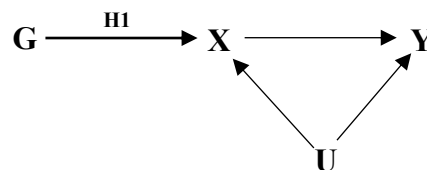
instrument est valide et doivent être satisfaites pour inférer l'existence d'un effet causal de l'exposition sur l'issue.⁶⁻⁸ Une hypothèse additionnelle (hypothèse d'homogénéité ou hypothèse de monotonie) est nécessaire pour définir clairement la magnitude de l'estimation causale ponctuelle et les intervalles de confiance.^{7,103,104}

2.1.3.1. Hypothèse 1 ou 'pertinence' (*relevance*)

L'hypothèse 1 de la RM stipule que l'instrument doit être associé à l'exposition (**Figure 4**). Il n'est pas nécessaire que l'association entre l'instrument et l'exposition soit causale.⁷ En revanche, il est requis que cette association soit forte.^{34,71}

Figure 4

Illustration de l'hypothèse 1 de la randomisation mendélienne.



Note. G est l'instrument, X l'exposition, Y l'issue et U un facteur de confusion de l'association entre X et Y. L'hypothèse 1 (H1) de la randomisation mendélienne requiert que G soit associé à X.

L'hypothèse 1 est la seule des trois premières hypothèses qui peut être vérifiée avec les données.⁵ L'association entre l'instrument et l'exposition est habituellement estimée par la régression de l'exposition sur l'instrument, et on utilise la statistique F de cette régression pour mesurer la force de l'instrument.⁹⁵ La statistique F est fonction de la proportion de la variation de l'exposition expliquée par l'instrument (R^2), de la taille de l'échantillon (n) et du nombre d'instruments (k), selon la formule ci-après : $F = \left(\frac{n-k-1}{k}\right) \left(\frac{R^2}{1-R^2}\right)$.⁹⁵ Des valeurs de $F \geq 10$ sont généralement considérées comme indicatives d'instrument fort.⁶⁶

Un SNP qui n'est pas fortement associé à l'exposition est dit instrument faible.^{92,105} Il est admis que les estimateurs de VI, comme les moindres carrés en deux étapes (*two stage least squares*, *2SLS*), sont asymptotiquement sans biais, mais sont biaisés lorsque l'échantillon est fini.⁹¹ Ce biais est dans la direction de l'estimation conventionnelle confondue c.à.d. du biais de l'estimateur OLS.

L'amplitude du biais de 2SLS est approximativement $1/F$ du biais de l'estimateur OLS de sorte qu'une valeur de $F = 10$ indique que le biais de l'estimateur 2SLS est approximativement égal à 10% celui de l'estimateur OLS.^{81,91,95}

L'utilisation en RM d'un instrument faible peut biaiser l'estimation de l'effet causal (on parle de biais d'instrument faible), augmenter l'erreur de type I, et voire amplifier les violations des hypothèses 2 et 3 de la RM.^{95,106,107} En effet, comme l'effet causal en RM peut être estimé comme le rapport entre l'effet de l'instrument sur l'issue (numérateur) divisé par l'effet de l'instrument sur l'exposition (dénominateur),⁹² un instrument faible conduit à un petit dénominateur de l'estimateur de VI.¹¹ Ainsi, les biais affectant le numérateur (pouvant résulter d'une violation de l'hypothèse 2 en présence d'une confusion génétique entre l'instrument et l'issue ou d'une violation de l'hypothèse 3 lorsque l'instrument affecte l'issue en dehors de l'exposition) ou les biais résultant d'un échantillon fini dans le dénominateur seront grandement amplifiés et peuvent induire une estimation causale de la RM plus biaisée que l'estimation non ajustée.¹¹ Toutefois, un instrument faible n'est pas automatiquement un instrument invalide car il peut être valide pour tester l'hypothèse d'absence d'un effet causal, mais manquer de puissance pour détecter la présence d'un effet causal.⁹²

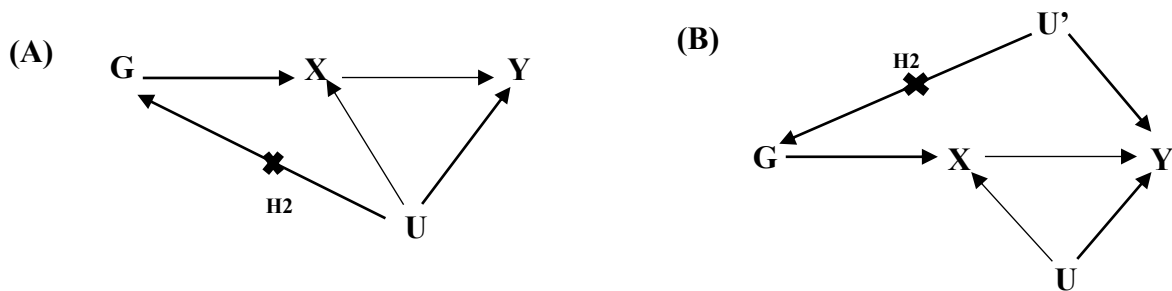
Les stratégies suggérées pour limiter le biais d'instruments faibles incluent l'augmentation de la valeur de F , notamment en augmentant la taille d'échantillon, en diminuant le nombre de paramètres dans le modèle (ex. en utilisant un GRS au lieu de multiples SNPs) ou en excluant les SNPs qui ne contribuent que très peu à expliquer la variation de l'exposition, et l'ajustement des modèles de la RM pour les facteurs de confusion de l'association entre l'exposition et l'issue.^{91,95} Des méthodes d'estimation robuste en présence d'instruments faibles ont été proposées.¹⁰⁸⁻¹¹⁰

2.1.3.2. Hypothèse 2 ou 'indépendance' (*independence*)

Aussi appelée 'échangeabilité' (*exchangeability*), l'hypothèse 2 de la RM requiert qu'il ne doive exister aucune cause commune entre l'instrument et l'issue, en d'autres termes, l'effet de l'instrument sur l'issue doit être non confondu.^{6,7,11,34,88,102} Dans la **Figure 5** illustrée ci-dessous, si la variable U (facteur de confusion de l'association entre X et Y, partie **A**) ou U' (un facteur de risque de Y qui n'est pas un facteur de confusion de l'association entre X et Y, partie **B**) cause l'instrument G, alors l'hypothèse 2 de la RM serait violée.

Figure 5

Illustration de l'hypothèse 2 de la randomisation mendélienne.



Note. G est l'instrument, X l'exposition, Y l'issue, U un facteur de confusion de l'association entre X et Y, et U' un facteur de risque de Y qui n'est pas un facteur de confusion de l'association entre X et Y. L'hypothèse 2 (H2) de la randomisation mendélienne requiert qu'il n'y ait pas de causes communes entre G et Y, et donc les variables U (A) et U' (B), qui sont des causes de Y, ne doivent pas aussi être des causes de G.

Dans la littérature, cependant, l'hypothèse 2 est aussi formulée comme l'instrument n'est pas associé aux (ou est indépendant des-) facteurs de confusion de l'association entre X et Y,^{66,111} ce qui ne spécifie pas le sens de causalité entre G et U. Sous cette formulation, certains auteurs ont dessiné le graphique acyclique orienté de la RM en pointant la flèche entre G et U en direction de U,¹¹¹ ce qui laisse entrevoir une certaine superposition des hypothèses 2 et 3 (l'hypothèse 3 est abordée dans la sous-section 2.1.3.3). D'autres auteurs encore ont relié G et U par une barre fléchée de deux cotés,¹¹² ce qui dans un sens poserait le problème de superposition des hypothèses 2 et 3 précédemment mentionné.

Au niveau populationnel, c'est l'ascendance ou la stratification de la population qui est la cause de violation de l'hypothèse 2. La stratification de la population est l'association génétique entre le SNP et un phénotype due à la structure sous-jacente de la population.⁸ Elle survient lorsqu'il existe des sous-groupes (liés par exemple à des phénotypes sociaux ayant des distributions géographiques différentes) de la population de l'étude qui diffèrent à la fois en fonction de la distribution des fréquences alléliques et de l'issue d'intérêt, confondant ainsi l'association entre l'instrument et l'issue.^{34,113}

Au niveau familial, en revanche, l'allocation aléatoire des allèles des parents vers la progéniture à la conception fait que les associations génétiques sont souvent indépendantes des facteurs de confusion traditionnels.^{34,67} Toutefois, on décrit deux effets familiaux susceptibles de violer l'hypothèse 2 de la RM, à savoir les 'effets dynastiques' (*dynastic effects*) lorsque les génotypes des parents affectent directement des phénotypes de la progéniture, et l' 'accouplement assortative' (*assortative mating*) lorsque les individus sélectionnent leurs partenaires en fonction d'un phénotype particulier.^{34,114}

De manière pratique, il n'existe aucune façon de vérifier formellement l'hypothèse 2 de la RM avec les données.^{5,11} Toutefois, il existe des stratégies de falsification des hypothèses de la RM qui peuvent capturer des violations de l'hypothèse 2,⁷ incluant la méthode comparant les effets conventionnel et causal estimés proposée par Glymour et al.⁵ (abordée en détail dans l'article 2 de la présente thèse) ainsi que l'utilisation des variables de 'contrôle négatif' (variables non censées être associées à l'exposition mais censées être influencées par les mêmes facteurs de confusion de l'association entre l'exposition et l'issue d'intérêt) comme issues (*negative control outcomes*).¹¹³ Cette dernière approche consiste à comparer les associations estimées entre l'exposition avec l'issue d'intérêt et avec l'issue de contrôle négatif, et suspecter une confusion génétique (violation de l'hypothèse 2) en présence d'une association entre l'exposition et l'issue de contrôle négatif.¹¹³ Toutefois, ces deux méthodes (méthode proposée par Glymour et méthode de contrôle négatif) ne sont pas spécifiques à l'hypothèse 2 et peuvent capturer aussi des violations de l'hypothèse 3 de la RM.^{5,7,35}

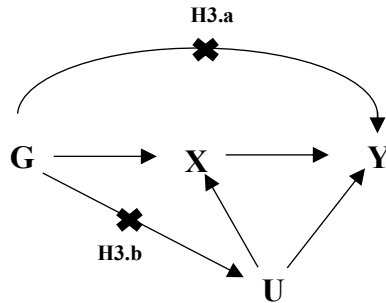
Dans la plupart des cas, la violation de l'hypothèse 2 de la RM est limitée en restreignant l'analyse aux sous-groupes ethniques de la population qui sont génétiquement homogènes et en contrôlant les modèles de la RM pour les composantes principales génomiques de l'ascendance.^{115,116} Des analyses de RM intra-familiales (*within-family MR analyses*) ont aussi été proposées.¹¹³

2.1.3.3. Hypothèse 3 ou 'restriction d'exclusion' (*exclusion restriction*)

Aussi appelée 'absence de pléiotropie',²⁶ l'hypothèse 3 requiert que l'instrument n'affecte pas l'issue en dehors de son effet sur l'exposition.^{6,7,34,88} Autrement dit, il ne doit exister aucune autre voie par laquelle l'instrument affecte l'issue en dehors de l'exposition, soit directement (**Figure 6, H3a**) ou indirectement via U (**Figure 6, H3b**).^{34,92}

Figure 6

Illustration de l'hypothèse 3 de la randomisation mendélienne.



Note. G est l'instrument, X l'exposition, Y l'issue, U un facteur de confusion de l'association entre X et Y. L'hypothèse 3 (H3) de la randomisation mendélienne requiert que G n'affecte pas Y en dehors de son association avec X, soit directement (H3.a), soit indirectement via U (H3.b).

En RM, l'hypothèse 3 est la plus critique et sa violation représente la principale menace à la validité des instruments.^{13,14} VanderWeele et al.⁷⁵ ont recensé plusieurs causes potentielles de violation de l'hypothèse 3 de la RM. Toutefois, les deux mécanismes principaux sont la pléiotropie horizontale, phénomène biologique par lequel un SNP affecte distinctement l'exposition et l'issue, ainsi que le déséquilibre de liaison (*linkage disequilibrium*) ou corrélation entre un SNP utilisé comme instrument et un autre SNP qui affecte l'issue d'intérêt par une voie autre que via l'exposition d'intérêt.^{12,14,66} La section 2.2 de la présente thèse aborde plus en profondeur la problématique de la pléiotropie en RM.

Comme pour l'hypothèse 2, l'hypothèse 3 de la RM ne peut être formellement vérifiée avec les données.^{5,11} La connaissance des mécanismes biologiques des SNPs utilisés comme instruments est utile pour justifier la plausibilité de l'hypothèse 3.⁷ Néanmoins, il existe plusieurs méthodes utilisées dans les analyses complémentaires pour tenter d'apporter une preuve d'une plausible violation de l'hypothèse 3. Ces méthodes sont abordées dans la section 2.2 et plus en détail dans l'article 1 de la présente thèse (voir **Tableau S6** de l'article 1).

2.1.3.4. Hypothèse 4 ou condition d'identification de l'estimation ponctuelle

Comme précédemment indiqué, les trois principales hypothèses ci-haut décrites permettent de tester l'hypothèse de l'existence (ou non) d'une association causale entre l'exposition et l'issue. En revanche, elles ne permettent pas de dériver une estimation ponctuelle de la magnitude de l'effet

de l'exposition sur l'issue.⁸ Il existe deux variantes communes de la condition d'identification de l'estimateur ponctuel (hypothèse d'homogénéité et hypothèse de monotonie) qui impliquent des interprétations quelque peu différentes de l'effet de l'exposition sur l'issue en RM.^{6-8,117} Les chercheurs sont donc invités à choisir la variante de l'hypothèse 4 qui serait la plus pertinente dans le contexte de leur étude.⁸

Hypothèse d'homogénéité

Cette hypothèse requiert que l'instrument ne modifie pas l'effet de l'exposition sur l'issue sur l'échelle additive.^{7,117} Elle permet d'évaluer si l'effet causal estimé est homogène (uniforme) dans toute la population, dans lequel cas il est interprété comme l'effet moyen de traitement (*average treatment effect*, ATE) dans la population.^{34,117} L'hypothèse d'homogénéité serait peu plausible advenant des facteurs de confusion non mesurés de l'association entre l'exposition et l'issue qui soient aussi des modificateurs d'effet.⁷ Lorsqu'il existe des covariables modifiant significativement l'effet estimé, certains auteurs ont suggéré d'ajuster les modèles pour ces covariables afin de récupérer l'effet causal moyen.⁷

Hypothèse de monotonie

Non testable,¹¹⁷ l'hypothèse de monotonie requiert que l'instrument affecte l'exposition dans la même direction pour tous les individus.^{7,8} Autrement dit, on ne devrait pas observer par exemple que l'augmentation du nombre d'allèles de risque augmente le niveau de l'exposition chez certains individus, mais le diminue chez d'autres.^{7,34} Lorsqu'on assume la monotonie en lieu et place de l'homogénéité, et particulièrement pour un instrument binaire, l'effet causal identifié représente l'effet local moyen du traitement (*local average treatment effect*, LATE) dans la sous-population des sujets compliants (*compliers*),^{7,81} c.à.d., les sujets dont le niveau d'exposition dépend de l'instrument.^{118,119} Cet effet est interprété comme le changement moyen dans l'issue pour un changement dans l'exposition parmi les sujets dont le niveau de l'exposition dépend de leur valeur de l'instrument.⁸¹ Toutefois, l'estimation de cet effet ne ferait pas l'unanimité parmi les chercheurs, certains estimant que le sous-groupe des sujets compliants n'est pas identifié et l'interprétation de cet effet n'est pas claire, par exemple quand l'instrument n'est pas binaire.⁷

2.1.4. Estimation de l'effet causal en RM

2.1.4.1. Nature de données exploitées en RM

Les analyses de RM peuvent être réalisées sur des données individuelles ou résumées. On parle de données individuelles (*individual level data*) lorsque l'analyse de RM exploite les variables SNPs, exposition et issue mesurées dans un seul échantillon. En revanche, l'analyse avec les données résumées exploite uniquement les statistiques sommaires (coefficients, erreurs types) respectivement des associations SNP–exposition et SNP–issue provenant d'un seul échantillon (*one-sample using summary data*), soit de deux échantillons différents (*two-sample using summary data*).¹²

En présence d'instruments faibles, l'estimation causale d'une analyse de RM à un échantillon est biaisée dans la direction de l'association conventionnelle entre l'exposition et l'issue.^{120,121} En revanche, dans une analyse de RM à deux échantillons, lorsqu'il n'y a pas de chevauchement des sujets entre l'échantillon ayant fourni les associations SNP–exposition et celui ayant fourni les associations SNP–issue, l'estimation causale est moins biaisée, et si biais il y avait, il serait vers la valeur nulle.¹²⁰ Mais en présence d'un chevauchement de certains sujets dans les deux échantillons, le biais va vers la valeur nulle ou vers l'association conventionnelle en fonction de l'importance du chevauchement.¹²

2.1.4.2. Méthodes d'estimation de l'effet causal en RM

Les méthodes pour estimer l'effet causal en RM sont nombreuses, et peuvent varier selon divers éléments, incluant la nature de données (individuelle ou résumée), le type d'instruments (simple ou multiple), et la nature de l'issue d'intérêt (continue ou binaire).

En présence des données individuelles, plusieurs méthodes d'estimation de l'effet causal sont applicables incluant la méthode de ratio des coefficients (méthode de Wald), la méthode en deux étapes (ex. 2SLS), les méthodes basées sur la vraisemblance (ex. méthodes Bayésiennes) et les méthodes semi-paramétriques (ex. méthode des moments généralisés).⁸¹ Lorsqu'on a des données résumées, l'estimation de l'effet causal peut se faire en utilisant le ratio de Wald pour un seul SNP, ou en cas de plusieurs SNPs, en combinant les ratios des associations SNP–issue et SNP–exposition par la méthode de pondération par l'inverse de la variance (*inverse-variance weighted, IVW*) ou la

méthode basée sur la vraisemblance.^{81,122} Il existe aussi plusieurs méthodes robustes utilisées comme analyses secondaires qui permettent d'estimer l'effet causal lorsque les hypothèses de RM sont peu satisfaites, en l'occurrence en présence de la pléiotropie (ex. MR–Egger).^{27-29,31-33,97,123,124}

Il est important d'indiquer qu'étant donné que les instruments n'expliquent souvent que peu de variation de l'exposition, les estimations de la RM ont souvent des intervalles de confiance plus larges que les estimations correspondantes obtenues par les régressions conventionnelles.¹²⁵

Burgess et Thompson⁹² et Burgess et al.^{81,122} offrent une description détaillée de différentes méthodes d'estimation de l'effet causal en RM. Ainsi donc, dans ce qui suit, nous nous limitons à décrire brièvement les méthodes les plus communes, soit la méthode de ratio des coefficients, la méthode en deux étapes et IVW, en assumant que l'exposition d'intérêt est une variable continue.

Méthode de ratio des coefficients (Méthode de Wald)

C'est une méthode simple, applicable en RM avec un seul instrument (un SNP ou un GRS), dans laquelle l'effet causal est estimé comme le rapport du coefficient l'instrument obtenu dans la régression de l'issue sur l'instrument divisé par celui obtenu dans la régression de l'exposition sur l'instrument.⁸¹ La méthode de Wald peut être appliquée dans un devis de RM à deux échantillons dans lequel l'association instrument–issue est estimée dans un échantillon, l'association instrument–exposition dans un autre échantillon.⁸¹

Si l'issue est continue, la régression linéaire est appliquée pour l'association instrument–issue. L'estimation causale est dite 'estimateur de Wald', et est asymptotiquement non biaisée c.à.d. en mesure que la taille d'échantillon augmente, le biais tend vers zéro.¹²⁶⁻¹²⁸

Si l'issue est binaire, un modèle non-linéaire, en l'occurrence log-linéaire (estimant un log ratio des risques) ou logistique (estimant un log ratio des cotes) est utilisé pour l'association instrument–issue, et l'estimation causale est dite 'estimateur de type Wald'.^{81,126,128} L'estimateur de type Wald allouant une régression logistique pour l'association instrument–issue peut être asymptotiquement biaisée advenant que l'effet causal réel n'est pas nul.¹²⁶⁻¹²⁸ En revanche, si c'est un modèle log-linéaire qui est appliqué pour l'association instrument–issue, l'estimateur de type Wald est cohérent avec le ratio des risques causal advenant que les conditions de linéarité et d'absence de modification d'effet dans le modèle log de la probabilité de l'issue sur l'exposition soient satisfaites.⁸¹ Si la prévalence de l'issue est rare, l'estimateur de type Wald allouant une régression

log-linéaire et celui avec une régression logistique dans l'association instrument–issue sont approximativement similaires, car sous cette condition (maladie rare), le ratio des cotes approxime le ratio des risques.^{128,129}

Plusieurs méthodes existent pour dériver les intervalles de confiance pour le ratio de Wald incluant les techniques de rééchantillonnage (*bootstrapping*).⁸¹

Méthode en deux étapes

C'est une méthode utilisable avec un ou plusieurs instruments et dans laquelle l'effet causal est estimé en deux étapes. Dans la première étape, on régresse l'exposition observée sur l'instrument et on génère la variable exposition prédite, et dans la deuxième étape, on régresse l'issue sur l'exposition prédite de la première régression.⁹²

Lorsque l'issue est continue, l'effet causal est estimé par la méthode 2SLS allouant des modèles de régression linéaire aux deux étapes. La méthode 2SLS peut également être appliquée dans un devis de RM à 'échantillon partagé' (*split-sample IV analysis*) dans lequel on subdivise un même échantillon en deux parties ; la régression de la 1^{ère} étape est effectuée en utilisant une partie de l'échantillon et celle de la deuxième étape en utilisant l'autre partie de l'échantillon.¹²¹ La RM à échantillon partagé serait moins vulnérable au biais d'instruments faibles comparé au 2SLS classique.³⁹ Aussi, contrairement au 2SLS dont le biais d'échantillon fini va dans la direction de celui de OLS, le biais de l'estimateur de RM à échantillon partagé va dans la direction de la valeur nulle.¹²¹

Lorsque l'issue est binaire, la méthode en deux étapes alloue un modèle non linéaire (ex. régression logistique ou log-linéaire) à la deuxième étape.⁸¹ Deux approches considérant un modèle non linéaire pour l'issue sont décrites, l'approche de 'substitution des prédicteurs en deux étapes' (*two-stage predictor substitution*, 2SPS) et l'approche d' 'inclusion des résidus en deux étapes' (*two-stage residual inclusion*, 2SRI), aussi appelée approche de la 'fonction de contrôle' (*control function approach*).^{130,131}

L'approche 2SPS est une simple extension du 2SLS classique dans laquelle la régression de l'exposition sur l'instrument à la première étape génère l'exposition prédite qui est incluse comme variable indépendante dans le modèle non linéaire de l'issue à deuxième étape, en remplacement

de l'exposition observée.¹³² L'analyse par l'approche 2SPS allouant un modèle logistique à la deuxième étape est conditionnelle sur l'instrument mais marginale sur toutes les autres variables, y compris l'exposition si elle est continue.⁸¹ L'estimation est donc un ratio des cotes causal moyen de la population.¹³⁰ Les biais de l'estimateur 2SPS sont similaires à ceux de l'estimateur de type Wald.¹³³

Dans le but de minimiser le biais de l'estimateur de VI lorsque le modèle de l'issue est non linéaire, il a été suggéré un autre estimateur, le 2SRI, qui inclut, dans le modèle de l'issue, les résidus de la régression de la première étape comme une variable indépendante additionnelle en plus de l'exposition prédite.^{131,133} Les résidus de la première régression capturent une partie de la variabilité des facteurs confondants non mesurés et du terme d'erreur de l'exposition.¹³³ En ajustant un modèle logistique de l'issue pour ces résidus, l'estimation causale par l'approche 2SRI est conditionnelle sur ces résidus et est plus proche du ratio des cotes spécifique au sujet.^{81,130} Globalement, même si le ratio des cotes est une mesure d'association non-collapsible (c.à.d. dont la valeur varie avec l'ajustement du modèle pour des covariables),¹³⁴ l'estimateur de VI allouant une régression logistique à la deuxième étape fournit toujours un test valide de l'hypothèse nulle.^{81,130}

Pondération par l'inverse de la variance (inverse-variance weighted, IVW)

IVW est la méthode conventionnelle la plus utilisée pour estimer l'effet causal dans un devis de RM à deux échantillons où l'on ne dispose que des données résumées (coefficients et erreurs types) des associations SNP–issue et SNP–exposition.³⁵

Assumant que les SNPs sont non corrélés (en équilibre de liaison), la méthode IVW estime l'effet causal ($\hat{\beta}_{IVW}$) comme une moyenne pondérée des estimations des ratios de Wald spécifique de chaque SNP dans un modèle de méta-analyse à effet fixe (*fixed-effect meta-analysis*), le poids étant l'inverse de la variance des associations SNP–issue,^{29,122,135} selon la formule suivante :

$$\hat{\beta}_{IVW} = \frac{\sum_{k=1}^K \hat{\beta}_{Xk} \hat{\beta}_{Yk} \sigma_{Yk}^{-2}}{\sum_{k=1}^K \hat{\beta}_{Xk}^2 \sigma_{Yk}^{-2}}$$

où K représente les SNPs (1 à K), $\hat{\beta}_{Xk}$ le coefficient de l'association entre le SNP k et l'exposition, $\hat{\beta}_{Yk}$ le coefficient de l'association entre le SNP k et l'issue et σ_{Yk}^2 la variance de l'association entre

le SNP k et l'issue.¹³⁵ Lorsqu'il existe une hétérogénéité importante entre les effets causaux spécifiques des SNPs, l'estimation IVW par un modèle de méta-analyse à effet aléatoire est plus appropriée du fait que celle par modèle à effet fixe tend à avoir des intervalles de confiance faussement étroits.³⁵

De manière équivalente, l'estimation causale par IVW serait la pente de la régression linéaire du coefficient de l'association SNP–issue sur le coefficient de l'association SNP–exposition pondéré par l'inverse de la variance de l'association SNP–issue en allouant un terme d'intercept nul.²⁹

La méthode IVW assume que tous les SNPs soient valides. Toutefois, l'estimation par IVW est aussi consistante s'il existe de la pléiotropie balancée (c.à.d. l'effet pléiotropique moyen est zéro) et l'hypothèse InSIDE (*Instrument Strength Independent of Direct Effect*, ou indépendance entre les effets pléiotropiques directs et les associations SNPs–exposition) est satisfaite.⁹⁶

2.1.5. Considérations diverses dans l'interprétation des résultats de la RM

2.1.5.1. Hypothèse d'équivalence gène-environnement

Cette hypothèse requiert que les changements dans l'exposition induits par l'instrument produisent les mêmes résultats de santé que les changements dans l'exposition induits par les manipulations environnementales ou thérapeutiques. La condition d'équivalence gène–environnement devrait être considérée, car d'un point de vue de santé publique, les résultats de la RM peuvent être exploités pour orienter des stratégies d'action (intervention) sur l'exposition.^{6,8}

2.1.5.2. Linéarité

Dans la grande majorité des cas, l'analyse de l'association entre une exposition et une issue en RM génère une seule estimation causale et assume implicitement que la relation entre cette exposition et l'issue est linéaire pour différentes valeurs de l'exposition.^{8,136} Une telle estimation ne serait pas informative si l'association entre l'exposition et l'issue est non-linéaire, comme rapporté dans plusieurs études de RM.^{8,137-139} Différentes stratégies existent dans la littérature pour gérer des associations non linéaires entre exposition et issue en RM, incluant la régression linéaire par morceaux (*piecewise linear method*).^{136,140}

2.1.5.3. RM et exposition variant dans le temps

Comme précédemment mentionné, la RM exploite les SNPs comme substituts de l'exposition pour estimer un effet causal de cette exposition sur une issue.⁶⁶ Étant donné que les effets génétiques commencent à la conception et se maintiennent tout au long de la vie des individus, les estimations de RM peuvent être vues comme des effets à vie des niveaux d'exposition génétiquement conférés.¹⁴¹⁻¹⁴³ Si donc l'association génétique avec l'exposition varient dans le temps au cours de la vie (information souvent non complètement disponible), l'interprétation des résultats d'une analyse de RM ayant considéré une mesure ponctuelle de l'exposition et des méthodes d'analyse instrumentale traditionnelles peuvent être peu fiables.³⁵ Shi et al.¹⁴² donnent une bonne description des types d'effets causaux pouvant être ciblés dans une analyse de RM avec une exposition variant dans le temps ainsi que les hypothèses additionnelles à satisfaire pour que ces effets soient valides : l'effet ponctuel (effet de l'exposition en seul point temporel sur l'issue), l'effet période (effet de l'exposition pendant une période sur l'issue) et l'effet à vie (effet de l'exposition tout au long de la vie sur le résultat). En pratique, différentes méthodes sont proposées pour estimer l'effet causal en RM lorsque l'association génétique varie dans le temps, incluant des modèles structurels marginaux avec estimation g (*g-estimation of structural mean models*)¹⁴² et l'analyse de RM multivariée.^{35,144}

2.1.5.4. Autres types de biais (collision, biais de sélection, erreur de classification)

Si la méthode de RM est reconnue pour être protégée contre le biais due aux facteurs de confusion,^{67-69,80} elle peut néanmoins être vulnérable à d'autres types de biais dont ceux dus à des erreurs de mesure (erreurs de classification) et le biais de sélection incluant le biais de collision.^{6,145}

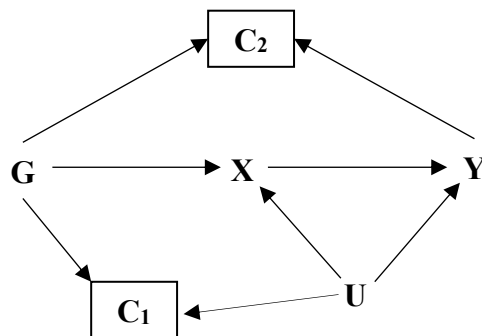
Biais de collision

Une variable est dite 'collisionneur' (*collider*) lorsqu'elle est causée par au moins deux autres variables.¹⁴⁶ Les deux variables causant le collisionneur sont marginalement indépendantes mais seront reliées si on conditionne pour le collisionneur.³⁵ Le biais de collision est une distorsion dans l'estimation de l'effet d'une exposition sur issue résultant d'une restriction ou d'une stratification sur – ou d'un ajustement pour un collisionneur dans un modèle de régression.^{35,146}

La collision peut être source de violation des hypothèses de RM. La **Figure 7** illustre deux exemples de mécanisme de collision pouvant biaiser l'estimation causale (voir Hughes et al.¹⁴⁷ et Coscia et al.¹⁴⁸ pour un recueil plus exhaustif de différents mécanismes de collision en RM). La variable C_1 est un collisionneur entre G et U, et C_2 l'est entre G et Y. Stratifier sur - ou conditionner pour la variable C_1 ou C_2 créerait une association génétique indépendante de X respectivement entre G et U ou entre G et Y susceptible d'invalidier les résultats de la RM.¹⁴⁹

Figure 7

Illustration de la collision en randomisation mendélienne.



Note. G est l'instrument, X l'exposition, Y l'issue, U un facteur de confusion de l'association entre X et Y, C_1 et C_2 sont des collisionneurs respectivement entre G et U et entre G et Y. Le conditionnement de l'analyse pour C_1 ou C_2 créerait une association génétique respectivement entre G et U ou entre G et Y indépendante de X biaisant l'estimation de l'effet causal de X sur Y.

Pour minimiser le biais de collision, il est recommandé de limiter l'ajustement des modèles de RM à certaines variables comme l'âge, le sexe, les composantes principales génomiques de l'ascendance et des variables techniques.³⁵ Dans le cas de RM à deux échantillons, il est important pour les chercheurs de savoir si les statistiques sommaires ont été obtenues à partir des modèles ajustés pour des covariables et si tel était le cas, s'assurer si ces covariables peuvent être des collisionneurs et biaiser potentiellement les estimations des associations SNP–exposition et/ou SNP–issue.³⁵ Certaines techniques ont été suggérées pour gérer le biais de collision en RM à un¹⁴⁸ ou deux échantillons.¹⁵⁰

Biais de sélection

L'effet estimé d'une exposition sur une issue dans un échantillon peut systématiquement différer de l'effet réel dans la population cible lorsque les sujets analysés ne sont pas représentatifs de cette population cible, du fait de la sélection des participants, de l'attrition au cours du suivi, de l'analyse de sous-groupes ou des données manquantes.^{147,151}

Le biais de sélection peut résulter d'un biais de collision lorsque la sélection des participants dans l'échantillon de l'étude dépend d'un collisionneur sur une voie non causale de l'exposition sur l'issue.^{35,145,148,149} Dans un diagramme acyclique orienté classique de la RM (**Figure 1**), l'exposition et l'issue sont des collisionneurs entre l'instrument (G) et le facteur de confusion de l'association entre X et Y (U). Ainsi, en RM, le biais de sélection peut être présente, par exemple, si la sélection des participants est dépendante de X et/ou Y¹⁴⁹ ou des variables que X et Y influencent, car une variable qui est la conséquence d'un collisionneur est aussi un collisionneur.¹⁵² Hughes et al.¹⁴⁷ ont décrit les différents mécanismes par lesquels la sélection des participants dans un échantillon sous analyse peut biaiser l'estimateur 2SLS. Ces auteurs ont montré que la sélection des participants ne biaise pas l'estimation de l'effet de X sur Y en RM si elle est totalement aléatoire ou si elle est fonction de G seulement ou de U seulement.¹⁴⁷

La méthode de pondération des participants de l'étude par la probabilité inverse de sélection dans l'échantillon a été proposée pour ajuster pour le biais de sélection.^{145,149} Toutefois, les simulations ont montré que cette méthode fonctionne si le biais de sélection est fort et le modèle de probabilité est correctement spécifié, alors que dans le cas contraire, la pondération peut même augmenter le biais.¹⁴⁹

Erreurs de mesure ou de classification

Dans une étude de RM, les erreurs de mesure concernent essentiellement les variables exposition et issue plutôt que les variants génétiques pour lesquels des techniques modernes de génotypage et des procédures de contrôle de qualité habituellement mises en place permettent une mesure plus précise.¹⁵³ Des travaux antérieurs ont suggéré qu'en présence des erreurs de mesure aléatoires de l'exposition, les estimateurs de VI pour modèles linéaires sont robustes et demeurent asymptotiquement non biaisés.^{153,154} En revanche, si l'erreur de mesure de l'exposition est

différentielle, l'estimation de RM peut être biaisée,⁷⁵ mais un test d'hypothèse nulle pour l'absence d'un effet causal restera tout de même valide.¹⁵³

Pierce et VanderWeele¹⁵³ ont utilisé des simulations statistiques pour examiner l'impact de deux types d'erreurs de mesure non différentielles d'une exposition continue et d'une issue (continue ou binaire) en RM, en utilisant un cadre théorique de régression de la variable (exposition ou issue) mesurée sur la vraie variable : (i) l'erreur de 'discrimination', capturée par le coefficient R^2 de cette régression, qui est l'erreur de mesure classique c.à.d. la mesure dans laquelle les sujets ayant des valeurs mesurées de l'exposition (ou de l'issue) plus élevées ont tendance à avoir aussi leurs vraies valeurs plus élevées, et (ii) l'erreur de 'calibration', capturée par la pente de la régression, qui est la sensibilité de la valeur mesurée (de l'exposition ou de l'issue) à la variation de la vraie valeur.¹⁵³ L'étude a montré que, pour les issues continues, les estimations de RM n'étaient pas biaisées par l'erreur de discrimination de l'exposition (divergence avec OLS) ou de l'issue (similarité avec OLS), mais l'augmentation de l'erreur de discrimination de l'issue diminue sensiblement la précision et la puissance. En revanche, les erreurs de calibration de l'exposition et de l'issue biaisaient l'estimation de RM si elles étaient différentielles (c.à.d. non égales). Pour les issues binaires, les erreurs de discrimination de l'exposition n'induisaient pas de biais substantiel contrairement aux erreurs de calibration qui biaisaient l'estimation de l'effet causal si celui-ci était non nul.¹⁵³

2.1.5.5. Causalité inverse

On parle de causalité inverse lorsque l'issue d'intérêt à un moment antérieur ou lorsqu'un précurseur de l'issue affecte causalement la valeur mesurée de l'exposition.⁶⁷ La causalité inverse peut être perçue comme de la confusion lorsque l'exposition et l'issue d'intérêt sont toutes deux influencées par un précurseur de l'issue.⁶⁷ Une étude de simulation sur l'impact de la causalité inverse dans un modèle de régression conventionnel a montré que la causalité inverse n'était généralement pas un problème pratique. Mais lorsqu'il y avait un biais, celui-ci se traduisait généralement par une sous-estimation (atténuation) de la vraie association.¹⁵⁵

Comme précédemment mentionné, l'un des avantages présentés des techniques de RM par rapport aux analyses par les régressions conventionnelles est la protection contre le biais dû à la causalité inverse du fait que les variants génétiques sont fixés dès la conception et précèdent

chronologiquement toutes les variables observées (en dehors de l'ascendance).⁶⁸ Toutefois, la protection de la RM contre le biais due à la causalité inverse n'est pas absolue.⁶⁷ En effet, VanderWeele et al.⁷⁵ ont montré que la causalité inverse pouvait être une source de violation des hypothèses de RM, préoccupation particulièrement inquiétante dans un devis cas-témoins dans lequel l'exposition est évaluée rétrospectivement après la mesure de l'issue. Dans une autre étude, Burgess et al.⁶⁷ ont noté que la causalité inverse pouvait biaiser l'estimation de l'effet (magnitude de l'effet) ou invalider l'inférence statistique sur la présence d'un effet causal en RM. C'est notamment le cas (i) lorsque l'association entre le variant génétique et l'exposition n'est pas primaire, mais médiée par un précurseur de l'issue, (ii) lorsque le variant génétique affecte directement l'exposition, qui affecte l'issue, mais que l'issue affecte encore l'exposition à un moment ultérieur (variables variant dans le temps), ou (iii) à travers des effets transgénérationnels lorsqu'une issue parentale affecte l'issue et les facteurs de confusion de l'association entre l'exposition et l'issue dans la génération cible.⁶⁷

2.1.5.6. RM et exposition binaire dichotomisée

Lorsque l'exposition d'intérêt est une variable binaire, il y a souvent un facteur de risque continu sous-jacent à partir duquel celle-ci a été dérivée ou dichotomisée (ex. obésité et IMC).¹⁵⁶ L'utilisation en RM d'une exposition binaire dichotomisée peut être problématique. Un premier défi concerne la satisfaction de l'hypothèse principale d'absence de pléiotropie en RM²⁶ selon laquelle l'instrument n'affecte pas l'issue en dehors de son effet sur l'exposition.^{7,34} Lorsque l'exposition d'intérêt est une variable binaire dichotomisée, le variant génétique peut affecter l'issue d'intérêt via le facteur de risque continu sous-jacent en l'absence d'une variation de l'exposition binaire, ce qui violerait l'hypothèse d'absence de pléiotropie. Ainsi, l'évaluation des hypothèses de RM dans une analyse où l'exposition est une variable binaire dichotomisée devrait aussi tenir compte du facteur de risque continu sous-jacent.¹⁵⁶ D'autres défis incluent le manque de clarté dans l'interprétation des estimations de RM due au fait que l'exposition binaire dichotomisée ne capture pas complètement la vraie relation causale, ainsi que le calcul de la puissance statistique qui a tendance à être sous-estimée. Face à ces problèmes, il est conseillé, en présence d'une exposition binaire dichotomisée, de simplement rapporter l'association entre le variant génétique et l'issue d'intérêt.¹⁵⁶

2.1.5.7. RM et épigénétique

L'épigénétique est un domaine en émergence qui étudie les modifications dans l'expression des gènes consécutives à des altérations structurelles et biochimiques de la chromatine sans modification de la séquence nucléotidique de l'ADN.¹⁵⁷ Les principaux mécanismes sous-tendant les effets épigénétiques comprennent la méthylation de l'ADN, des modifications des histones et les ARN (acides ribonucléiques) non codants.^{157,158} Certaines modifications épigénétiques sont transmissibles d'une génération à l'autre, d'autres pas.¹⁵⁹

Les effets épigénétiques peuvent induire une violation des hypothèses de RM et en invalider les résultats.¹⁶⁰ En effet, puisqu'au travers les changements dans l'expression des gènes, les effets épigénétiques hérités des parents peuvent modifier les associations entre les génotypes et des phénotypes, les analyses de RM devraient assumer, en plus de l'assignation aléatoire des allèles à la conception, une distribution aléatoire des changements épigénétiques à la conception pour que les hypothèses principales de la RM restent valides.¹⁶¹ Ogbuanu et al.¹⁶² ont utilisé des simulations statistiques pour examiner l'impact d'ignorer les effets épigénétiques hérités des parents à la conception sur l'estimation de l'effet causal en RM. Les auteurs ont montré que les effets épigénétiques affectaient l'association entre l'instrument et l'exposition qui biaisait systématiquement l'effet causal estimé (biais qualifié de 'biais épigénétique').¹⁶²

2.1.5.8. RM et taille de l'échantillon

Le calcul de la taille d'échantillon est une préoccupation centrale en RM. En effet, l'utilisation en RM des instruments génétiques qui n'expliquent souvent que faiblement la variation de l'exposition fait que la puissance statistique pour détecter un effet causal non nul peut être faible et des échantillons de taille plus grande sont souvent requis.^{58,163}

Pour une issue (Y) continue et un instrument simple (G , un SNP ou un GRS), la taille d'échantillon (n_{RM}) requise en RM pour détecter un effet causal non nul (β_X) de l'exposition X sur Y est approximativement égale à celle requise pour détecter un effet OLS de même amplitude divisée par le carré de la corrélation entre X et G (ρ_{XG}^2 , pouvant être estimé par le coefficient R^2 de la régression de X sur G).⁵⁸ Elle peut être calculée en utilisant la formule suivante :

$$n_{RM} = \frac{(Z_{\alpha} + Z_{\beta})^2 \cdot V}{\beta_X^2 \cdot \rho_{XG}^2}$$

où α est le niveau de signification, β le risque d'erreur de 2^{ème} espèce et $V = Var(Y|X)/Var(X)$ le rapport de la variance échantillonnale de Y due à des facteurs autres que X divisé par la variance échantillonnale de X .¹⁶⁴

Pour un niveau de signification $\alpha = 0,05$ ($Z_{\alpha} = 1,96$) et un risque $\beta = 0.20$ ($Z_{\beta} = 0,842$), la taille d'échantillon pour tester un changement de β_X écart-type dans Y par augmentation d'un écart-type dans X est donnée par la formule : $n_{RM} = \frac{7,848}{\beta_X^2 \cdot \rho_{XG}^2}$.⁵⁸

Dans le cas d'une issue binaire avec $P(Y = 1)$ la probabilité que l'issue soit présente et $P(Y = 0)$ la probabilité qu'elle soit absente, un niveau de signification de 0,05, et une puissance de 0,80, la taille d'échantillon est calculée par la formule ci-après :

$$n_{RM} = \frac{7,848}{\beta_X^2 \cdot \rho_{XG}^2 P(Y=1) P(Y=0)} \text{ où } \beta_X \text{ est le log ratio de cotes.}^{58}$$

Comme on peut le noter, la taille d'échantillon pour maintenir une puissance de 80% sera plus importante pour une issue binaire comparativement lorsque l'issue est continue.⁵⁸

Des formules ont été proposées pour le calcul de la puissance statistique dans un devis de RM à deux échantillons.¹⁶⁵

2.2. Pléiotropie en RM

2.2.1. Définition et types de pléiotropie en RM

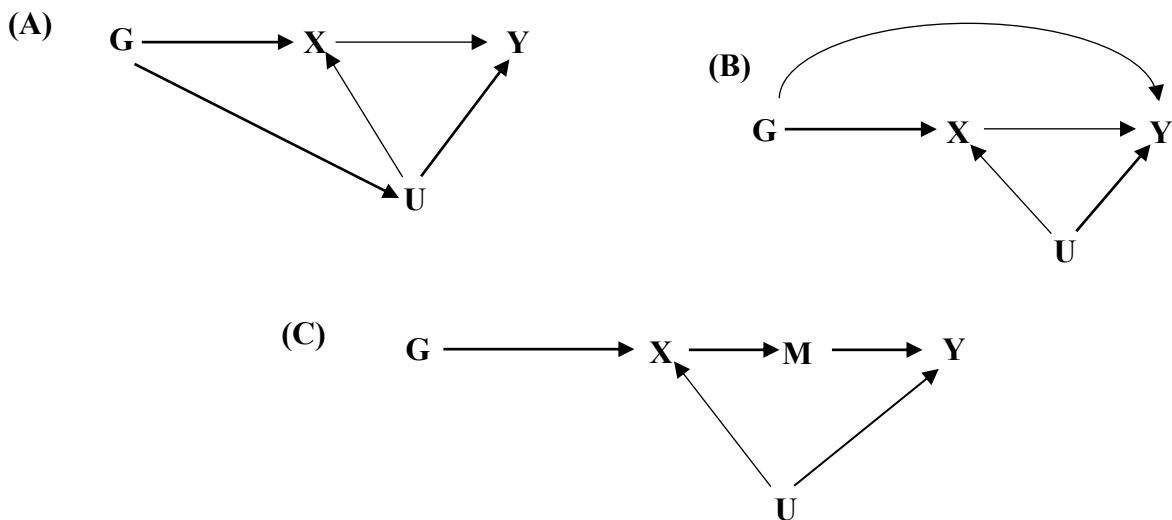
La pléiotropie est un concept génétique se référant au phénomène biologique par lequel un SNP influence plus qu'un phénotype.^{12,166} Dans le contexte de la RM, il est important de distinguer la pléiotropie horizontale de la pléiotropie verticale, car seule la pléiotropie horizontale invalide l'instrument.¹⁶⁷

La pléiotropie horizontale, aussi appelée pléiotropie biologique ou pléiotropie de type I, survient lorsque l'instrument associé à l'exposition affecte également l'issue par une voie autre que celle

via l'exposition.^{12,26,168} Elle peut être indirecte (ou pléiotropie corrélée) lorsque les effets pléiotropiques de G sur Y sont corrélés aux effets de G sur X (c'est notamment le cas lorsque le chemin pléiotropique passe par U, **Figure 8, partie A**) ou directe (ou pléiotropie non corrélée) lorsque les effets pléiotropiques de G sur Y ne sont pas corrélés aux effets de G sur X (c'est le cas où le chemin pléiotropique est directe de G à Y, **Figure 8, partie B**).^{34,123} En revanche, la pléiotropie verticale, aussi appelée pléiotropie médiée ou pléiotropie de type II, fait référence à une succession d'associations sur le même chemin causal de sorte qu'un SNP influence un phénotype lequel influence à son tour d'autres phénotypes (**Figure 8, partie C**).^{12,168,169} Pour la brevété de langage, nous utilisons simplement le terme pléiotropie pour désigner la pléiotropie horizontale.

Figure 8

Illustration de la pléiotropie en randomisation mendélienne.



Note. G est l'instrument, X l'exposition, Y l'issue, U un facteur de confusion de l'association entre X et Y, et M une variable sur le chemin causale de X à Y. **(A)** Pléiotropie horizontale corrélée : G est associé à Y indépendamment de X via U. **(B)** Pléiotropie horizontale non corrélée : G est associé à Y indépendamment de X de manière directe. **(C)** Pléiotropie verticale : G influence X qui influence M lequel influence Y sur le même chemin causal.

2.2.2. Source et ampleur de pléiotropie dans le génome humain

La pléiotropie est reconnue comme un phénomène ubiquitaire, largement distribuée dans le génome humain.¹⁷⁰ Elle s'observe au niveau des loci fonctionnels (et donc détectables par des GWAS), mais principalement au niveau des SNPs ayant des effets faibles sur plusieurs phénotypes.^{24,170} La pléiotropie est favorisée par le déséquilibre de liaison (corrélation entre plusieurs SNPs) et la nature polygénique de plusieurs traits.¹⁷⁰ En effet, contrairement aux biomarqueurs de nature protéique qui sont souvent associés à un gène codant proximal,^{96,98} diverses phénotypes étudiés, incluant des paramètres anthropométriques (ex, IMC), traits cardiométaboliques (ex. pression artérielle), etc., sont des traits complexes, influencés en amont par plusieurs gènes qui affectent à leur tour d'autres phénotypes, ce qui accroît les possibilités de pléiotropie.¹⁷⁰⁻¹⁷²

L'ampleur de la pléiotropie a été soulignée par Chesmore et al.²⁵ qui ont analysé 1094 phénotypes et 14459 gènes dans le catalogue des GWAS. Ces chercheurs ont identifié 1 à 53 phénotypes associés par gène, et 44% des gènes rapportés étaient associés avec plus d'un phénotype. De même, Watanabe et al.¹⁷³ ont analysé systématiquement 4155 GWAS couvrant 2965 traits uniques à partir de 295 études. Ces auteurs ont montré que la pléiotropie était largement distribuée dans le génome humain et identifiés à l'échelle des loci, des gènes, des SNPs et des groupes des gènes. En effet, 93% des loci, 81% des gènes, 60% des SNPs et 57% des groupes de gènes associés ont démontré de la pléiotropie.¹⁷³

2.2.3. Impact et défi de la gestion de la pléiotropie en RM

Comme précédemment évoqué, la pléiotropie horizontale est une cause principale de violation de l'hypothèse 3 de la RM,¹⁴ et peut entraîner un biais dans l'estimation de l'effet causal, une inflation de l'erreur de type I et une perte de la puissance statistique.^{26,29,167} En revanche, la pléiotropie verticale n'est pas problématique pour la RM. Bien au contraire, elle est l'essence même de l'approche de RM qui utilise un instrument qui influence une exposition pour estimer l'effet de cette exposition sur une issue qui lui est en aval.¹⁶⁸

La pléiotropie représente un défi majeur pour la validité des études de RM, non seulement du fait de son ampleur dans le génome humain, mais aussi du fait qu'il n'existe aucune façon de prouver

l'absence de pléiotropie au moyen des données sous analyse.^{5,11} L'unique façon pour attester la plausibilité de l'absence de pléiotropie d'un SNP est à travers la connaissance de son mécanisme biologique spécifique.^{88,174,175} Or, cette connaissance n'est pas toujours disponible pour plusieurs SNPs utilisés comme instruments en RM,^{98,99} ce qui fait que plusieurs chercheurs recourent souvent au critère d'existence d'une association statistique robuste entre le SNP et l'exposition pour être sélectionné.⁴² Cette approche augmente les possibilités d'inclusion des SNPs pléiotropiques et donc de biaiser les résultats de la RM.^{28,100}

2.2.4. Prévention, détection ou correction de la pléiotropie en RM

Même s'il n'est pas possible de prouver que la pléiotropie n'existe pas dans une étude de RM, il est, en revanche, possible de falsifier les hypothèses de RM et tenter d'apporter l'évidence d'une plausible violation des hypothèses de RM et/ou de contrôler le biais.^{7,34,42,88}

Les lignes directrices de la RM^{35,85} recommandent de trianguler différentes approches qui opèrent sous différentes hypothèses sur la pléiotropie pour augmenter la crédibilité des résultats de la RM.^{176,177} La littérature de RM rapporte certaines stratégies souvent utilisées par les chercheurs pour prévenir la pléiotropie dans une étude de RM. Aussi, plusieurs méthodes et stratégies pour tenter de détecter la pléiotropie et/ou fournir des estimations robustes à la pléiotropie ont été proposées et comparées,^{12,26-33,75} mais leur utilisation en pratique n'avait pas été documentée. Le premier article de la présente thèse est une revue de littérature qui recense les différentes méthodes et stratégies utilisées par les auteurs des études de RM pour prévenir, détecter ou corriger pour le biais dû à l'utilisation des instruments pléiotropiques en RM. Nous référons donc le lecteur à cet article pour la description de ces différentes méthodes.

2.2.5. Illustration de la problématique de pléiotropie en RM : gène *FTO* comme instrument de l'adiposité

2.2.5.1. Adiposité, gène *FTO* et association *FTO*-adiposité

Le surpoids et l'obésité sont des facteurs de risque du développement de diverses comorbidités chroniques (maladies cardiovasculaires, diabète de type 2, divers cancers, maladies mentales etc.) responsables d'une morbi-mortalité et des dépenses économiques importantes.^{178,179} Au Canada, la

proportion de la population adulte de 18 ans et plus en surpoids ou obèse est passée de 61,9% en 2015 à 63,1% en 2018.¹⁸⁰

Plusieurs études de RM estimant l'effet causal de l'adiposité sur diverses issues de santé utilisent des SNPs sur le gène *FTO* comme des instruments car *FTO* est reconnu comme le signal génétique le plus fortement associé à l'adiposité.^{17,59} C'est en 2007 que, pour la première fois, Frayling¹⁸¹ rapportait une association génétique entre le SNP rs9939609 sur le gène *FTO* et l'adiposité dans une GWAS incluant des participants d'ascendance Européenne. Par la suite, des associations entre différents SNPs sur *FTO* et l'adiposité ont été documentées dans plusieurs autres GWAS, dont celles impliquant des populations d'ascendance non Européenne.¹⁸²⁻¹⁸⁶

Le gène *FTO* est situé sur le chromosome 16q12.2 sur un locus s'étendant sur environ 417 979 paires de bases (chromosome humain) et contenant 9 régions codantes ou exons.¹⁸⁷ Les associations génétiques les plus robustes avec l'obésité polygénique identifiées dans des GWAS impliquent des SNPs localisés sur les introns (régions non codantes) 1 et 2 de *FTO* comprenant environ 89 SNPs.^{187,188}

2.2.5.2. Mécanismes biologiques de *FTO* sur l'adiposité

L'influence génétique de *FTO* sur la régulation de la graisse corporelle s'exerce tout au long de la vie humaine.¹⁸⁹ Certaines études ont indiqué que l'association entre *FTO* et l'IMC devient détectable vers l'âge de trois ans ; elle augmente fortement dès l'enfance puis progressivement au fur et en mesure du vieillissement de l'individu.¹⁸⁹⁻¹⁹¹ Cependant, la fonction biologique spécifique de *FTO* sur l'adiposité est encore largement méconnue,¹⁹² mais plusieurs mécanismes ont été évoqués, particulièrement l'influence sur l'apport alimentaire et la dépense énergétique.^{189,192}

En effet, il a été admis que certains SNPs sur le gène *FTO* influenceraient le poids corporel par leurs effets sur l'augmentation d'apports alimentaires, hypothèse soutenue notamment par une expression importante de *FTO* au niveau des régions hypothalamiques régulant l'apport alimentaire et la balance énergétique.¹⁹³ D'autres chercheurs ont évoqué que *FTO* augmentait, indépendamment de l'adiposité, la concentration sérique de leptine,¹⁹⁴ une hormone peptidique régulant entre autres l'apport alimentaire et la masse corporelle.¹⁹⁵ Mais cette association entre *FTO* et la leptine circulante n'a pas été observée dans une autre étude.¹⁹⁶ Ainsi, il a été suggéré que

les effets de *FTO* sur l'augmentation de l'apport alimentaire seraient plutôt négligeables chez l'humain.¹⁸⁹ Par conséquent, la régulation de la dépense énergétique apparaît comme le mécanisme principal par lequel *FTO* influence l'accumulation adipocytaire dans l'organisme humain.

A ce sujet, les travaux de Claussnitzer et al.¹⁹⁷ ont montré que *FTO* perturbe les mécanismes régulant la différenciation précoce des cellules adipocytaires induisant une surexpression des gènes de la famille des Iroquois (*IRX3* et *IRX5*), avec pour conséquences une réduction de la dissipation de la chaleur par la thermogenèse mitochondriale et une augmentation du stockage des lipides dans les adipocytes blancs au détriment de la synthèse des adipocytes bruns, dissipateurs de chaleur. D'autres études ont aussi suggéré que *FTO* pourrait favoriser la formation des cellules adipeuses en régulant l'activité d'un facteur de transcription adipogène dénommé *RUNX1T1* (*Runt-related transcription factor 1*).¹⁸⁷ Selon ce mécanisme, *FTO* régule l'épissage des ARN messager en contrôlant l'activité de liaison du facteur d'épissage *SRSF2* (*Serine and arginine rich splicing factor 2*) qui épisse *RUNX1T1*, avec pour conséquence une surexpression de l'isoforme S de *RUNX1T1* responsable de la formation des adipocytes.^{187,198}

2.2.5.3. Suspicion de pléiotropie de *FTO*

Divers SNPs sur le gène *FTO* ont été suspectés de pléiotropie sur un éventail de phénotypes de santé incluant des traits/maladies cardiovasculaires (pression artérielle, hypertension, maladie coronarienne, etc.), des maladies métaboliques (diabète de type 2, syndrome des ovaires polykystiques, stéatose hépatique non alcoolique, etc.), des cancers et des maladies mentales (démence, maladie d'Alzheimer).^{60,199-207} Cette suspicion était basée sur la persistance d'une association statistique significative dans des modèles de régression de ces phénotypes sur *FTO* après ajustement pour l'IMC. En outre, des études de RM évaluant l'association entre l'IMC et certains phénotypes de santé mentale (troubles mentaux courants incluant la dépression, symptômes d'anxiété phobique, bien-être subjectif) ont incriminé des SNPs sur *FTO* d'être pléiotropiques après triangulation des certaines méthodes d'évaluation des hypothèses de RM.^{37,39,40,208}

Outre ces suspicions empiriques de pléiotropie, certains mécanismes biologiques plausibles à travers lesquels *FTO* pourrait influencer certains phénotypes indépendamment de l'adiposité ont été évoqués. Par exemple, l'étude de Pausova et al.²⁰¹ a suggéré que *FTO* pourrait favoriser une

augmentation de la PA et donc l'apparition de l'hypertension en régulant à la hausse le tonus vasomoteur sympathique. Aussi, la suspicion de pléiotropie de *FTO* sur des phénotypes de santé mentale semble soutenue par des études animales. C'est le cas notamment de l'étude de Sun et al.²⁰⁹ qui a suggéré que *FTO* pourrait influencer des comportements de type anxiété et dépression par le biais d'altérations du microbiote intestinal chez des souris.

Ainsi donc, puisque *FTO* peut être pléiotropique sur de multiples phénotypes, son utilisation comme instrument de l'adiposité en RM peut représenter un défi majeur sur la validité des résultats.

Dans cette thèse, nous exploitons justement le potentiel de pléiotropie de *FTO* comme instrument de l'IMC en RM comme une opportunité méthodologique pour examiner les approches utilisées pour gérer la pléiotropie en RM (étude 1 et illustration de l'étude 2). Nous examinons aussi l'effet causal de l'IMC sur la PAS, un phénotype vis-à-vis duquel *FTO* est suspecté de pléiotropie,²⁰¹ dans le but de voir l'impact des stratégies communément préconisées pour gérer la baisse de la pression induite par la médication antihypertensive sur l'estimation de l'effet causal et la détection de l'invalidité des instruments en RM (étude 3).

Chapitre 3 – Méthodes

Chacun des trois objectifs spécifiques de cette thèse a fait l'objet d'un article. Dans les lignes qui suivent, nous décrivons tour à tour les aspects méthodologiques spécifiques à chacune des trois études de la thèse. Le texte de ce chapitre est une copie bonifiée de celui apparaissant dans la section méthodes de chacun des trois articles de la thèse.

3.1. Méthodes pour l'étude 1 (revue de littérature)

Le premier objectif de cette thèse était de cartographier comment les chercheurs en RM préviennent, détectent ou contrôlent, et discutent des violations potentielles de l'hypothèse de restriction d'exclusion dues notamment à la pléiotropie. Pour atteindre cet objectif, nous avons réalisé une revue de littérature des études de RM.

3.1.1. Recherche documentaire

3.1.1.1. Source et opérationnalisation de la recherche d'articles

La recherche d'articles a été effectuée en anglais dans quatre moteurs de recherche, à savoir, Pubmed, Medline, Embase et Web of Science, en utilisant les mots-clés "*Mendelian randomization*", "*adiposity*" et leurs termes connexes ou équivalents. Les termes connexes ou équivalents pour *Mendelian randomization* ont été empruntés de Boef et al..²³

PubMed

Requête :

```
("Mendelian Randomization Analysis"[Mesh] OR "Mendelian randomisation"[all fields] OR "Mendelian randomization"[all fields] OR (Mendelian[all fields] AND randomi*[all fields]) OR "genetic instrumental variable"[all fields] OR "genetic instrumental variables"[all fields] OR "genetic instrument"[all fields] OR "genetic instruments"[all fields] OR "genes as instruments"[all fields] OR "gene as instrument"[all fields] OR "genes as instrument"[all fields] OR "gene as instruments"[all fields] OR (instrument*[ti] AND (gene[ti] OR genes[ti] OR genetic*[ti] OR mendel*[ti])) OR (("instrumental variable"[all fields] OR "instrumental variables"[all fields] OR
```

"instrumented analysis"[all fields] OR "instrumented analyses"[all fields] OR "instrumental variable analysis"[all fields] OR "instrumental variable analyses"[all fields] OR "instrumental variables analysis"[all fields] OR "instrumental variables analyses"[all fields]) AND (gene OR genes OR genetics OR mendel OR mendelian)) OR ("mendelian"[all fields] AND ("randomisation"[all fields] OR "randomization"[all fields] OR "randomising"[all fields] OR "randomizing"[all fields])) AND ("adiposity"[Mesh] OR "adiposity"[all fields] OR "obesity"[all fields] OR "body mass index"[Mesh] OR "body mass index"[all fields] OR "BMI"[all fields] OR "fat mass"[all fields] OR "body fat"[all fields] OR "body fat percentage"[all fields] OR "obesity, abdominal"[Mesh] OR "waist-hip ratio"[Mesh] OR "waist-hip ratio"[all fields] OR "waist circumference"[Mesh] OR "waist circumference"[all fields])

Medline

Requête :

(Mendelian Randomization Analysis/ OR "Mendelian randomisation".af OR "Mendelian randomization".af OR "genetic instrumental variable".af OR "genetic instrumental variables".af OR "genetic instrument".af OR "genetic instruments".af OR "mendel randomise" OR "mendel randomize" OR "mendel randomization" OR "mendel randomisation" OR "random Mendelian" OR "genes as instruments".af OR "gene as instrument".af OR "genes as instrument".af OR "gene as instruments".af) OR (instrument*.ti AND (gene.ti OR genes.ti OR genetic*.ti OR mendel*.ti)) OR (("instrumental variable".af OR "instrumental variables".af OR "instrumented analysis".af OR "instrumented analyses".af OR "instrumental variable analysis".af OR "instrumental variable analyses".af OR "instrumental variables analysis".af OR "instrumental variables analyses".af) AND (gene OR genes OR genetics OR mendel OR mendelian).af) OR ("mendelian".af AND ("randomisation".af OR "randomization".af OR "randomising".af OR "randomizing".af)) AND (adiposity/ OR "adiposity".af OR "obesity".af OR body mass index/ OR "body mass index".af OR "BMI".af OR "fat mass".af OR "body fat".af OR "body fat percentage".af OR obesity, abdominal/ OR waist-hip ratio/ OR "waist-hip ratio".af OR waist circumference/ OR "waist circumference".af)

Embase

Restriction: no conference abstracts.

Requête :

(Mendelian Randomization Analysis/ OR "Mendelian randomisation".af OR "Mendelian randomization".af OR "genetic instrumental variable".af OR "genetic instrumental variables".af OR "genetic instrument".af OR "genetic instruments".af OR "mendel randomise" OR "mendel randomize" OR "mendel randomization" OR "mendel randomisation" OR "random Mendelian" OR "genes as instruments".af OR "gene as instrument".af OR "genes as instrument".af OR "gene as instruments".af) OR (instrument*.ti AND (gene.ti OR genes.ti OR genetic*.ti OR mendel*.ti)) OR (("instrumental variable".af OR "instrumental variables".af OR "instrumented analysis".af OR "instrumented analyses".af OR "instrumental variable analysis".af OR "instrumental variable analyses".af OR "instrumental variables analysis".af OR "instrumental variables analyses".af) AND (gene OR genes OR genetics OR mendel OR mendelian).af) OR ("mendelian".af AND ("randomisation".af OR "randomization".af OR "randomising".af OR "randomizing".af)) AND (adiposity/ OR "adiposity".af OR "obesity".af OR body mass index/ OR "body mass index".af OR "BMI".af OR "fat mass".af OR "body fat".af OR "body fat percentage".af OR obesity, abdominal/ OR waist-hip ratio/ OR "waist-hip ratio".af OR waist circumference/ OR "waist circumference".af)

Web of Science

Restrictions: no review, meeting abstract, editorial material, correction, letter, or book chapter.

Requête :

ALL FIELDS: (("Mendelian randomisation" OR "Mendelian randomization" OR "genetic instrumental variable" OR "genetic instrumental variables" OR "genetic instrument" OR "genetic instruments" OR "mendel randomise" OR "mendel randomize" OR "mendel randomization" OR "mendel randomisation" OR "random Mendelian" OR "gene as instrument" OR "gene as instruments" OR "instrumental genetic variable" OR "instrumental genetic variables" OR "instrumental variable" OR "instrumental variable analysis") AND ("adiposity" OR "obesity" OR "body mass index" OR "BMI" OR "fat mass" OR "body fat" OR "body fat percentage" OR "abdominal obesity" OR "waist-hip ratio" OR "waist circumference"))

3.1.1.2. Critères d'inclusion des études

Nous avons inclus les articles publiés avant le 21 mai 2020 qui répondaient aux trois critères suivants : (i) l'analyse principale était la RM, (ii) l'exposition principale était l'adiposité évaluée par l'IMC, et (iii) au moins un SNP sur le gène *FTO* était sélectionné comme instrument de l'IMC. Nous n'avons mis aucune restriction sur l'issue d'intérêt.

Notre choix de restreindre l'inclusion aux études ayant utilisé *FTO* comme instrument de l'IMC était motivé par le fait que, non seulement qu'il n'aurait pas été très pratique de réviser l'ensemble des études de RM publiées, mais surtout par le fait que nous nous attendions à ce que les études retenues fournissent une discussion plus riche sur la violation de la restriction d'exclusion, dans la mesure où *FTO* est suspecté d'affecter un large éventail d'issues de santé indépendamment de l'IMC, incluant des affections cardiovasculaires et métaboliques, des cancers et des problèmes de santé mentale.^{37,199-202,205-207,210}

3.1.1.3. Critères d'exclusion des études

Nous avons exclu les études dans lesquelles la RM était une analyse secondaire afin de nous assurer que l'étude fournissait une évaluation détaillée des hypothèses de la RM, les études non rédigées en anglais ou en français, les résumés de conférences, les lettres, les rapports, les commentaires, les corrections, les points de vue, les nouvelles de recherche, les perspectives, les protocoles d'études, les errata, les articles de forum, les éditoriaux, les chapitres de livres, les revues, les méta-analyses des études de RM, les études publiées dans des périodiques d'économie de la santé, les articles théoriques ou méthodologiques avec des exemples d'analyses de RM, les études phénoMIques de RM (*MR-PheWAS*, *MR phenome-wide association studies*), les études de RM avec une exposition composite combinant l'IMC et un ou plusieurs autres facteurs de risque, ainsi que les études impliquant des couples mère-enfants (dans lesquelles l'instrument et l'exposition sont évalués chez la mère et l'issue chez l'enfant). Lorsqu'une étude de RM était bidirectionnelle, nous n'avons considéré que l'analyse qui traitait l'IMC comme l'exposition.

3.1.1.4. Processus de sélection des articles

Le processus de sélection des articles a été conduit par un investigateur (Fiston Ikwa Ndol Mbutiwi) et s'est déroulé en plusieurs étapes. L'étape initiale a consisté à importer puis fusionner toutes les

publications identifiées via les quatre moteurs de recherche dans le logiciel EndNote X8, ce qui a facilité d'identifier les doublons qui ont été supprimés. Les prochaines étapes ont consisté à examiner successivement les titres, puis les résumés, et enfin les manuscrits entiers. Une publication qui ne remplissait pas tous les critères d'inclusion ou qui présentait un critère d'exclusion était exclu de l'examen de la prochaine étape.

3.1.2. Extraction des données

Dans chaque article retenu, nous avons collecté certaines données générales, puis des données autour de trois thèmes qui décrivent la façon dont la pléiotropie des instruments est traitée dans le processus analytique, à savoir (i) la sélection et la combinaison des instruments où l'on peut prévenir d'inclure des SNPs pléiotropiques, (ii) les analyses complémentaires où l'on applique les méthodes pour détecter et/ou gérer la pléiotropie et (iii) la discussion des résultats en tenant compte de l'évaluation de la pléiotropie.

3.1.2.1. Données générales

Les données générales collectées comprenaient le plan d'étude de RM, le type d'instruments utilisés, et le(s) SNP(s) *FTO* sélectionné(s).

Nous avons documenté si le plan d'étude de RM était à un ou à deux échantillons. Nous avons considéré comme RM à un échantillon (i) les études qui ont effectué la RM en utilisant des données au niveau individuel sur le(s) SNP(s), l'exposition et l'issue, et (ii) celles qui ont utilisé des statistiques sommaires sur les associations SNP-exposition et SNP-issue obtenues à partir d'un même échantillon.¹² La RM à deux échantillons a été retenue lorsque l'étude a utilisé les statistiques sommaires sur les associations SNP-exposition et SNP-issue provenant respectivement de deux échantillons distincts.¹² Le type d'instruments utilisés pouvait être un seul SNP, un GRS, ou des instruments multiples. Un GRS a été défini comme la fusion de plusieurs SNPs en une seule variable correspondant à la somme pondérée ou non pondérée d'allèles de risque des différents SNPs.^{93,94} Les instruments multiples ont été définis par l'utilisation de ≥ 2 SNPs en tant qu'instruments distincts dans un seul modèle ou à la combinaison des effets estimés en utilisant ≥ 2 SNPs en un seul effet causal en utilisant des techniques de méta-analyse.⁹³ Nous avons enregistré le(s) SNP(s) spécifique(s) sur *FTO* utilisé(s) dans chacun de types d'instruments ci-haut décrits.

3.1.2.2. Approches pour prévenir la pléiotropie (sélection et combinaison des instruments)

Selon lignes directrices de la RM,³⁵ les SNPs peuvent être sélectionnés dans les régions des gènes qui codent spécifiquement l'exposition (approche biologique) ou sur la base de leur association statistique avec l'exposition (approche statistique). Les SNPs connus ou suspectés d'être pléiotropiques peuvent être exclus après leur sélection initiale avant d'effectuer l'analyse principale.³⁵

Dans chaque article inclus, nous avons rapporté l'approche utilisée par les auteurs pour sélectionner les SNPs. Nous avons également documenté si les auteurs ont évoqué avoir utilisé plusieurs SNPs indépendants ou un GRS comme un moyen pour atténuer l'effet de la pléiotropie en supposant que les effets pléiotropiques potentiels des SNPs s'annuleraient mutuellement.²¹¹ Enfin, nous avons rapporté toute autre stratégie se rapportant à la sélection des instruments explicitement présentée comme servant à minimiser la présence de pléiotropie.

3.1.2.3. Approches pour détecter et/ou gérer la pléiotropie (analyses complémentaires)

Les lignes directrices de la RM exigent aux auteurs des études de RM de rapporter les méthodes utilisées dans les analyses complémentaires pour évaluer les hypothèses de la RM, ce qui inclut l'évaluation du biais dû à la pléiotropie.^{35,36}

Nous avons enregistré les méthodes utilisées pour évaluer les hypothèses d'indépendance et de restriction d'exclusion, à l'exception de celles utilisées pour la stratification de la population, en soulignant si l'évaluation de la pléiotropie était explicitement mentionnée pour justifier l'utilisation de ces méthodes. Nous avons classé les méthodes en quatre catégories, à savoir (i) les méthodes d'estimation instrumentale robustes (par exemple MR-Egger) et autres (par exemple MR multivariable), (ii) les méthodes permettant de détecter l'hétérogénéité entre les effets causaux estimés avec les différents instruments (par exemple les tests statistiques d'hétérogénéité entre les effets causaux estimés spécifiques aux SNP), (iii) les méthodes permettant de détecter ou de prendre en compte les associations entre les instruments et l'issue en dehors de l'exposition (par exemple l'analyse de médiation), et (iv) d'autres méthodes. Les méthodes robustes fournissent des estimations de l'effet causal avec un ensemble d'hypothèses plus faibles que les méthodes conventionnelles.^{35,96}

3.1.2.4. Discussion des résultats en tenant compte de l'évaluation de la pléiotropie

Nous avons vérifié si les auteurs ont discuté les hypothèses d'indépendance et de restriction d'exclusion, en distinguant les études qui faisaient explicitement référence au terme "pléiotropie" de celles qui ne le faisaient pas. Lorsque les études suspectaient une invalidité de l'instrument, nous avons vérifié en outre si les auteurs rapportaient l'impact de l'invalidité de l'instrument sur les résultats de la RM, et si des SNPs sur *FTO* étaient incriminés.

3.1.2.5. Processus d'extraction et de validation des données

Les articles retenus ont été examinés en entier (titre, résumé, manuscrit, matériel supplémentaire) afin d'identifier les informations d'intérêt. Les données ont été extraites par un investigateur (Fiston Ikwa Ndol Mbutiwi), par la suite, un deuxième investigateur (Marie-Pierre Sylvestre) a validé les méthodes recensées dans les articles. Nous avons contacté certains auteurs (listés dans l'annexe 3) lorsque des informations n'étaient pas ou peu clairement rapportées dans l'article. Nous avons documenté les méthodes telles qu'elles étaient explicitement énoncées dans chaque article, indépendamment de leur applicabilité ou de leur pertinence. Une évaluation critique de l'utilisation de certaines méthodes est fournie dans la discussion de l'article 1, en plus d'une description sommaire de toutes les méthodes recensées, incluant leurs principales hypothèses et limites, qui est présentée dans le **Tableau S6** du même article.

3.1.3. Analyse des données

Les données ont été tabulées à l'aide de Microsoft Excel® 2016 et décrites à l'aide du logiciel Stata/IC version 14.2 (StataCorp, College Station, Texas, USA).

3.2. Méthodes pour l'étude 2 (étude de simulation statistique)

Le deuxième objectif de cette thèse était d'évaluer la performance de la méthode basée sur la confusion positive à détecter des instruments invalides dans divers contextes pratiques d'études de RM. Pour atteindre cet objectif, nous avons conduit une étude de simulation statistique et avons illustré l'application de cette méthode dans les études de RM retenues dans notre revue de littérature (étude 1).

Nous avons privilégié de concevoir la structure et les scénarios de nos simulations sur la base des études de simulations précédemment publiées^{41,94} afin de nous assurer que les données générées concordent avec la littérature, limitant ainsi des erreurs, et assurer la comparabilité de nos résultats avec la littérature existante. Par la suite, nous avons envisagé des valeurs des paramètres et des scénarios supplémentaires.

Dans les lignes qui suivent, nous présentons tout d’abord le fonctionnement de la méthode basée sur la confusion positive, puis décrivons nos simulations, avant d’illustrer l’application de cette méthode dans les articles de RM publiés.

3.2.1. Description de la méthode basée sur la confusion positive

3.2.1.1. Approche originale

La méthode basée sur la confusion positive avait originalement été proposée dans un contexte d’analyse de RM utilisant un SNP comme instrument et une issue continue.⁵ Soit X une exposition continue, G un SNP qui est l’instrument de X , Y une issue continue et i l’indice des individus. L’estimation ponctuelle conventionnelle de l’association entre X et Y est obtenue par le coefficient $\hat{\beta}_{CO}$ dans le modèle linéaire $Y_i = \beta_0 + \beta_{CO}X_i + \varepsilon_{Yi}$ qui utilise la méthode OLS. L’estimation ponctuelle de la RM ($\hat{\beta}_{MR}$) de l’association entre X et Y peut être obtenue en utilisant la méthode 2SLS. La première étape correspond à la régression linéaire de X sur G , alors qu’à la deuxième étape on régresse Y sur les valeurs prédites de X obtenues de la régression de la première étape. Sous l’hypothèse de confusion nette positive de l’association entre X et Y (c.à.d. que X et Y ont une cause commune qui les affecte chacune positivement, ou plusieurs causes communes telles que l’effet confondant net de l’association entre X et Y est positif), les situations pour lesquelles $\hat{\beta}_{MR} > \hat{\beta}_{CO}$ suggèrent que G est un instrument invalide pour X dans son association avec Y .⁵

Il est important de mentionner que cette méthode compare les estimés ponctuels sans tenir compte de la variabilité échantillonnale affectant les coefficients estimés, ni du biais d’échantillon fini qui affecte l’estimation de RM.¹⁰⁶ Par conséquent, les situations où $\hat{\beta}_{MR}$ and $\hat{\beta}_{CO}$ sont proches devraient être interprétées avec prudence dans la mesure où elles pourraient refléter davantage l’imprécision des estimations plutôt que l’invalidité des instruments.

3.2.1.2. Extension de la méthode basée sur la confusion positive aux issues binaires

Bien que la méthode basée sur la confusion positive ait été initialement présentée pour une issue continue, elle a néanmoins déjà été appliquée dans une analyse de RM avec issue binaire.³⁸

Lorsque l'issue est binaire, un modèle linéaire généralisé (*generalized linear model, GLM*), comme la régression logistique, est utilisé à la place de la méthode OLS pour l'analyse conventionnelle ainsi qu'à la deuxième étape de la méthode 2SLS, une technique appelée approche de 'substitution de prédicteurs en deux étapes' (*two-stage predictor substitution, 2SPS*). Il existe un autre estimateur de variable instrumentale pour une issue binaire qui est similaire au 2SPS, mais qui inclut en plus les résidus de la régression de première étape comme une variable indépendante supplémentaire dans la régression de la deuxième étape, connu sous le nom d' 'inclusion des résidus en deux étapes' (*two-stage residual inclusion, 2SRI*) ou 'fonction de contrôle' (*control function*).^{81,131,132}

Nos simulations se sont concentrées sur des issues binaires rares, qui ont été modélisées en utilisant la régression logistique pour l'analyse conventionnelle et l'estimateur 2SRI allouant un modèle logistique à la deuxième étape pour la RM. Toutefois, dans les situations où les issues binaires ne sont pas rares, les modèles log-linéaires (par ex. la régression robuste de Poisson) estimant les ratios des risques doivent être préférés aux modèles logistiques, car les ratios des cotes estimés dans une régression logistique surestiment les ratios des risques.²¹²⁻²¹⁴

3.2.2. Simulations

Nous avons conduit une étude de simulation pour évaluer la performance de la méthode basée sur la confusion positive à détecter l'invalidité des instruments dans des contextes d'études de RM considérant (i) des issues continues et binaires, et (ii) des SNPs uniques et GRS comme instruments. Dans les quatre cas examinés et décrits ci-dessous, nous avons considéré différents scénarios investiguant l'impact de la taille d'échantillon, du type de pléiotropie (directe et indirecte) et de la proportion des SNPs pléiotropiques dans les cas où le GRS était l'instrument. Le GRS était non pondéré et incluait chaque fois 25 SNPs simulés. Nous avons conçu nos scénarios en nous basant sur des études de simulations précédemment publiées (l'étude de Burgess (2012)⁴¹ pour la RM utilisant un SNP comme instrument et l'étude de Burgess and Thompson (2013)⁹⁴ pour

l'utilisation de GRS en RM), que nous avons étendues par la suite en considérant des valeurs de paramètres et scénarios additionnels. Les paramètres considérés dans chacun des quatre cas sont résumés dans le **Tableau S1** de l'article 2.

Dans chacun des quatre cas de simulation, nous avons rapporté les paramètres suivants : (i) la moyenne des statistiques F de la régression de l'exposition sur l'instrument entre les itérations comme une mesure de la force de l'instrument⁹⁵; (ii) la moyenne des $\hat{\beta}_{MR}$ et des $\hat{\beta}_{CO}$ entre les jeux de données simulés ; (iii) la moyenne des erreurs types de $\hat{\beta}_{MR}$ (SE_{MR}) et de $\hat{\beta}_{CO}$ (SE_{CO}) entre les itérations pour évaluer la précision des deux estimations ; et (iv) la performance de la méthode basée sur la confusion positive comme le pourcentage de jeux de données simulés dans lesquels la méthode détecte des instruments invalides, c.à.d. où $\hat{\beta}_{MR} > \hat{\beta}_{CO}$.

Toutes les simulations ont été réalisées à l'aide du logiciel R version 4.2.1.

3.2.2.1. Cas 1 : un SNP comme instrument et issue continue

Nous avons avant tout répliqué les simulations de Burgess⁴¹ qui ont évalué la performance de la méthode basée sur la confusion positive dans une étude de RM de 2000 sujets en utilisant un SNP comme instrument et une issue continue.

Création des données

Nous avons généré une exposition continue X comme la somme linéaire d'un SNP binaire (G , codé 0, 1, et divisant la population en deux sous-groupes égaux), d'un facteur de confusion non mesuré (U) et du terme d'erreur de X (\mathcal{E}_X) : $X_i = \beta_{XG}G_i + \beta_{XU}U_i + \mathcal{E}_{Xi}$ où β_{XG} et β_{XU} étaient fixés respectivement à 0,3 et 0,5. Nous avons généré Y comme dépendant à la fois de X , U et du terme d'erreur de Y (\mathcal{E}_Y) : $Y_i = \beta_{YX}X_i + \beta_{YU}U_i + \mathcal{E}_{Yi}$ où β_{YX} est l'effet causal réel de X sur Y et β_{YU} fixé à 0,5. Nous avons considéré trois différentes valeur pour le coefficient β_{YX} , soit respectivement 0 (effet causal nul), 0,4 (effet causal positif) et -0,4 (effet causal négatif). Nous avons initialement généré des échantillons de 2000 sujets comme l'a fait Burgess,⁴¹ et par la suite, avons considéré des échantillons de plus grande taille, soit 5000, 10000, 20000 et 30000 sujets. Nous avons simulé 10000 jeux de données pour chacun des trois scénarios en rapport avec la validité de G comme instrument, comme décrit ci-dessous.

Scénarios simulés

Trois scénarios en rapport avec la validité de G comme instrument ont été considérés :

Scénario 1 : G comme instrument valide, donc $U_i, \mathcal{E}_{Xi}, \mathcal{E}_{Yi} \sim N(0, 1)$ de manière indépendante.

Scénario 2 : G comme instrument invalide par association avec U (pléiotropie indirecte) : $U_i|G \sim N(\beta_{UG}G_i, 1)$.

Scénario 3 : G comme instrument invalide par association directe avec Y (pléiotropie directe) : $Y_i = \beta_{YX}X_i + \beta_{YG}G_i + \beta_{YU}U_i + \mathcal{E}_{Yi}$.

Les effets pléiotropiques indirect (β_{UG} , scénario 2) et direct (β_{YG} , scénario 3) de G ont été respectivement fixés à 0,2.

3.2.2.2. Cas 2 : un SNP comme instrument et issue binaire

Nous avons modifié les simulations décrites dans le cas 1 pour prendre en compte Y binaire. Plus particulièrement, nous avons simulé Y à partir d'une distribution binomiale avec la probabilité de maladie (P_i) de chaque participant selon les équations $\text{logit}(P_i) = \beta_0 + \beta_{YX}X_i + \beta_{YU}U_i$ dans les scénarios où G était un instrument valide ou pléiotropique par association avec U , et $\text{logit}(P_i) = \beta_0 + \beta_{YX}X_i + \beta_{YU}U_i + \beta_{YG}G_i$ dans le dernier scénario où G était pléiotropique par association directe avec Y . L'ordonnée à l'origine (β_0) était fixée de manière à avoir une prévalence de l'issue d'environ 10%.

3.2.2.3. Cas 3 : GRS comme instrument et issue continue

Pour évaluer la performance de la méthode basée sur la confusion positive dans le contexte d'une étude de RM utilisant un GRS comme instrument, nous avons partiellement répliqué deux des sept scénarios de simulations de Burgess et Thompson⁹⁴ qui ont évalué l'utilisation d'un GRS en RM. Les deux scénarios répliqués étaient (i) instruments valides avec des effets de magnitude égale et (ii) instruments invalides par association avec U . Nous avons ajouté un nouveau scénario impliquant des SNPs invalides par association directe avec Y . Dans chacun des trois scénarios, nous avons restreint les analyses en considérant comme instrument un GRS non pondéré incluant 25 SNPs simulés.

Création des données

Nous avons généré 25 SNPs indépendants (G , indexés par j , en équilibre de Hardy-Weinberg, avec une fréquence de l'allèle mineur de 0,3) codés 0, 1, 2 exprimant le nombre d'allèles de risque de chaque SNP. Nous avons généré X continue comme la somme linéaire de G_j , U et \mathcal{E}_X :

$$X_i = \sum_{j=1}^J \beta_{XG} G_{ij} + \beta_{XU} U_i + \mathcal{E}_{Xi} \text{ où } \beta_{XG} \text{ était fixé à } 0,06 \text{ pour chaque } G_j \text{ et } \beta_{XU} \text{ à } 1.$$

Nous avons généré Y continue de la même manière comme dans le cas 1, et avons fixé β_{YU} à 1. Nous avons généré des échantillons de taille variable de 2000 à 30000 sujets comme pour le cas 1, et avons considéré 1000 itérations pour chacun des trois scénarios simulés ci-dessous.

Scénarios simulés

Scénario 1: GRS comme instrument valide : U_i , \mathcal{E}_{Xi} et $\mathcal{E}_{Yi} \sim N(0, 1)$ de manière indépendante.

Scénario 2: GRS comme instrument invalide par association avec U (pléiotropie indirecte) :

$$X_i = \sum_{j=1}^J \beta_{XG} Z_j G_{ij} + \beta_{XU} U_i + \mathcal{E}_{Xi}; \text{ et } U_i \sim N\left(\sum_{j=1}^J \beta_{UG}(1 - Z_j) G_{ij}, 1\right)$$

Scénario 3: GRS comme instrument invalide par association directe avec Y (pléiotropie directe) :

$$Y_i = \beta_{YX} X_i + \beta_{YU} U_i + \left(\sum_{j=1}^J \beta_{YG}(1 - Z_j) G_{ij}\right) + \mathcal{E}_{Yi}$$

Dans les scénarios 2 et 3, Z_j représentent des variables aléatoires muettes (*random dummy variables*) qui prennent la valeur de 1 si le SNP G_j est valide et 0 s'il est pléiotropique. La probabilité d'avoir des SNPs pléiotropiques a été fixée respectivement à 0,1, 0,3 et 0,5, et l'effet pléiotropique indirect (β_{UG} , scénario 2) ou direct (β_{YG} , scénario 3) d'un SNP invalide fixé égal à β_{XG} .

3.2.2.4. Cas 4 : GRS comme instrument et issue binaire

Nous avons modifié les simulations décrites dans le cas 3 pour prendre en compte Y binaire.

3.2.3. Comparaison de la méthode basée sur la confusion positive avec des méthodes alternatives dans l'évaluation de la validité des instruments dans des études de RM sélectionnées

Nous avons illustré l'utilisation de la méthode basée sur la confusion positive dans notre récente revue (article 1 de cette thèse)⁴² qui a inclut 128 études de RM ayant utilisé au moins un SNP sur le gène *FTO* comme instrument de l'IMC.

Nous avons restreint l'application aux études qui (i) étaient à un échantillon, (ii) ont utilisé un instrument simple (un SNP ou GRS) dans l'analyse principale, et (iii) ont rapporté les estimations conventionnelles et de RM de l'association entre l'IMC et l'(les) issue(s) par une mesure d'association similaire. Pour limiter le nombre d'études, nous avons considéré les études qui avaient la PAS, la PAD, l'hypertension ou une condition de santé mentale comme issue. Dans les études retenues, nous avons rapporté le(s) instrument(s) utilisé(s), la taille d'échantillon, les effets estimés par la régression conventionnelle (OLS ou GLM) et par la RM. Nous avons appliqué la méthode basée sur la confusion positive en comparant l'estimation OLS ou GLM et l'estimation de la RM en assumant une confusion nette positive entre l'IMC et l'issue, comme précédemment indiqué. Finalement, nous avons comparé la conclusion sur la validité de l'instrument obtenue par notre application de la méthode basée sur la confusion positive avec celle des auteurs de l'article et qui est basée sur des méthodes alternatives d'évaluation des hypothèses de RM utilisées dans l'article.

3.3. Méthodes pour l'étude 3 (analyse de RM)

Le troisième objectif de cette thèse était d'examiner l'impact de méthodes courantes de gestion de la médication antihypertensive dans les études de RM modélisant la PA sur l'estimation de l'effet causal et la détection des violations potentielles de la restriction d'exclusion. Pour atteindre cet objectif, nous avons réalisé une étude de RM avec les données de l'Étude Longitudinale Canadienne sur le Vieillissement (ÉLCV) estimant l'effet causal de l'IMC (mesuré au recrutement) sur la PAS et l'hypertension (évalués au premier cycle de suivi) en utilisant le SNP rs1558902 sur *FTO* et un GRS comme les deux instruments de l'IMC.

3.3.1. Population de l'étude

Nous avons exploité trois bases de données de la cohorte globale de l'ÉLCV,²¹⁵ la base des données du recrutement (version 6.0, 2011-2015), la base de données du cycle 1 de suivi (version 3.0, 2015-2018) et la base de données génétiques (version 3.0), obtenues sous l'application numéro 2010004. En résumé, la cohorte globale de l'ÉLCV inclut 30097 participants âgés de 45-85 ans au recrutement, sélectionnés aléatoirement dans sept provinces Canadiennes, qui seront suivis tous les trois ans jusqu'en 2033 ou au décès. Les participants ont été interviewés en personne à domicile, puis, examinés physiquement au niveau des 11 sites de collecte des données retenus où des échantillons de sang et d'urines ont été collectés.

Notre échantillon analytique comprenait 24679 participants qui avaient des données phénotypiques (du recrutement et du cycle 1 de suivi) et génétiques. De ce nombre, nous avons exclus 1784 participants d'ascendance non-Européenne, 1862 qui avaient des réponses invalides ou des données manquantes/non disponibles pour les variables d'intérêt, et 630 ayant un IMC > 40 kg/m² du fait que l'association entre l'IMC et la PAS n'était pas linéaire pour des valeurs de l'IMC > 40 kg/m². En effet, bien que les analyses de RM puissent prendre en compte des relations non-linéaires,¹³⁶ elles nécessiteraient plus qu'un seul paramètre (minimalement une composante linéaire et une composante quadratique de l'effet) pour bien capturer l'association entre l'IMC et la PAS. En restreignant l'analyse aux valeurs pour lesquelles l'association entre l'IMC et la PAS est linéaire, cela nous a facilité la comparaison de méthodes de prise en compte de la médication antihypertensive que nous examinons. Le diagramme de flux décrivant l'échantillon analytique final (n=20430) est présenté dans la **Figure S1** de l'article 3.

3.3.2. Génotypage

Le génotypage de la cohorte globale de l'ÉLCV a été réalisé à l'aide de la matrice *Affymetrix UK Biobank Axiom* et a fourni des données sur 794409 marqueurs, ce qui a permis d'obtenir des données imputées de l'ensemble du génome pour environ 308 millions de variants génétiques.²¹⁶ Les procédures de contrôle de qualité des marqueurs et des échantillons ainsi que l'imputation des variants génétiques sont décrites dans Forgetta et al.²¹⁶

3.3.3. Variables

3.3.3.1. Exposition

L'IMC, mesuré au recrutement, et calculé comme le poids (kg) divisé par le carré de la taille (m²), était l'exposition d'intérêt de l'étude. Le poids et la taille ont été mesurés chez des participants déchaussés, position debout, en utilisant respectivement une pèse-personne médicale numérique de modèle *140-10 Healthweigh* (maximum à 250 kg) et le stadiomètre Seca 213.²¹⁷

3.3.3.2. Issues

Les données sur les mesures de la PA, l'histoire médicale de diagnostic de l'hypertension et l'utilisation actuelle de la médication antihypertensive ont été obtenues au premier cycle de suivi. Le diagnostic de l'hypertension et la prise actuelle des antihypertenseurs étaient auto-rapportés par les participants. La PA a été mesurée à l'aide d'un moniteur automatisé et non invasif de modèle BpTRU™ BPM200 (VSM MedTech Ltd., Coquitlam, BC, Canada).²¹⁷ Six mesures ont été prises respectivement pour la PAS et la PAD, et nous avons considéré la moyenne des cinq dernières mesures.^{218,219} Dans cette étude, la PAS était l'issue d'intérêt principal. Toutefois, nous avons aussi considéré l'hypertension comme issue binaire incluant la médication antihypertensive dans sa définition (voir les méthodes de gestion de la médication antihypertensive considérées dans la section analyses statistiques). L'hypertension a été définie par au moins un des deux critères suivants : (i) PAS \geq 140 mm Hg ou PAD \geq 90 mm Hg et/ou (iii) diagnostic auto-rapporté de l'hypertension et utilisation actuelle des médicaments antihypertenseurs.²²⁰

3.3.3.3. Instruments

Nous avons sélectionné 97 SNPs associés à l'augmentation de l'IMC au seuil de signification à l'échelle du génome ($P < 5 \times 10^{-8}$) dans une large méta-analyse de 125 études incluant 339224 individus.¹⁸³ Les 97 SNPs étaient disponibles dans la base des données génétiques de l'ÉLCV (voir **Table S1** de l'article 3). Nous avons recodé les SNPs de manière à représenter le nombre d'allèles associés à l'augmentation de l'IMC et avons harmonisé les allèles en comparant la fréquence allélique avec celle rapportée par Locke et al..¹⁸³ Dix SNPs (rs11165643, rs9641123, rs3736485, rs7715256, rs2075650, rs6477694, rs17724992, rs11847697, rs1528435 et rs6465468) ont été exclus du fait que leurs associations estimées avec l'IMC étaient négatives dans l'ÉLCV. Les 87

SNPs restants ont été utilisés pour construire un GRS non pondéré en additionnant le nombre d'allèles associés à l'augmentation de l'IMC chez chaque participant.⁹⁴ Nous avons utilisé le SNP *FTO* rs1558902 et le GRS de 87 SNPs (qui incluait *FTO* rs1558902) comme les deux instruments de l'IMC. Nous avons aussi construit un GRS non pondéré incluant l'ensemble de 97 SNPs sélectionnés que nous avons utilisé comme instrument dans l'analyse de sensibilité.

3.3.3.4. Covariables

Si les modèles de régressions conventionnelles requièrent l'ajustement pour des covariables afin de gérer leurs effets confondants sur l'association entre une exposition X et une issue Y, l'estimation de l'effet causal en RM ne nécessite pas des ajustements pour des covariables.¹⁵² En RM, la confusion par les des facteurs de risque (post-conception) est gérée grâce à l'utilisation des variants génétiques comme substituts (instruments) de X. En effet, étant donné que le génotype est fixé dès la conception et précède chronologiquement les variables observées, il est plus difficile de trouver une variable post-conception qui cause à la fois le génotype et Y, confondant ainsi l'association estimée.^{67-69,80} Cependant, la RM reste vulnérable à la confusion génétique entre l'instrument génétique et Y (communément appelée la stratification de la population) qui survient lorsqu'il existe des facteurs sous-jacents à la structure de la population (ex. phénotypes sociaux ayant des distributions géographiques différentes) qui créent des sous-groupes des personnes qui diffèrent simultanément en fonction de la distribution des fréquences alléliques et de Y.^{8,34,113} La stratification de la population est habituellement gérée en restreignant l'échantillon analytique à un sous-groupe de sujets génétiquement homogènes et en contrôlant les modèles de RM pour les composantes principales génomiques de l'ascendance.^{115,116}

Toutefois, les covariables confondantes mesurées de l'association entre X et Y peuvent être utiles dans une étude de RM, notamment dans le devis de RM à un échantillon. Les covariables mesurées peuvent servir à évaluer la plausibilité de l'hypothèse de la restriction exclusion. On peut, en effet, estimer l'association entre l'instrument et les covariables, et suspecter une violation de la restriction d'exclusion en présence d'une association significative (voir article 1 de la thèse pour une discussion plus détaillée de cette méthode d'évaluation de la pléiotropie).⁴² Lorsqu'une covariable est suspectée d'être sur un chemin pléiotropique entre l'instrument et Y, il est parfois suggéré d'ajuster l'analyse de RM pour la covariable incriminée pour tenter de bloquer le chemin pléiotropique.²²¹ Les covariables confondantes peuvent également être utilisées pour ajuster les

modèles de RM dans le but d'améliorer la précision des estimations et de limiter le biais d'instruments faibles.^{95,152} Toutefois, étant donné que l'inclusion d'une covariable dans un modèle de RM peut biaiser l'effet causal estimé si elle induit un biais de collision ou si la covariable est sur le chemin causal entre l'instrument et Y (pléiotropie verticale), il est recommandé de limiter l'ajustement des modèles de RM pour des variables comme l'âge, le sexe, les variables techniques (ex. centres de recrutement), et les composantes principales génomiques de l'ascendance.^{35,152}

Ainsi, sur base de la littérature,²²²⁻²²⁸ nous avons considéré certaines covariables connues d'être associées à l'adiposité et à la PA, et donc, susceptibles de confondre l'association entre l'IMC et la PAS. Ces covariables étaient évaluées au recrutement à l'aide d'un questionnaire auto-rapporté dans l'ÉLCV. Il s'agit de (i) l'âge, (ii) le sexe, (iii) le statut marital, (iv) le niveau d'éducation, (v) le statut tabagique, (vi) la consommation d'alcool et (vii) l'activité physique. Toutes ces covariables ont été utilisées dans les analyses complémentaires pour évaluer la plausibilité de l'hypothèse de la restriction d'exclusion. Spécifiquement, les variables âge et sexe ont été inclus dans les modèles comme des covariables d'ajustement dans le but d'améliorer la précision des estimations de la RM.

Dans l'ÉLCV,²¹⁵ la variable âge (dénommée dans le jeu de données AGE_NMBR_COM) était une variable continue, exprimées en années. Le sexe (SEX_ASK_COM) était une variable dichotomique, masculin vs. féminin. Le statut marital (SDC_MRRTL_COM) était une variable catégorielle (codée 1 = célibataire, jamais marié(e) ou jamais vécu en couple, 2 = marié(e)/conjoint(e) de fait, 3 = veuf(ve), 4 = divorcé(e), 5 = séparé(e) et 9 = refus). Nous avons recatégorisé le statut marital en 'avec partenaire' (marié(e) ou conjoint(e) de fait) vs. 'sans partenaire' (célibataire, jamais marié(e) ou jamais vécu en couple, veuf(ve), divorcé(e), séparé(e)).²²⁹ Le niveau d'éducation (ED_HSGR_COM) a été évalué comme la diplomation de l'école secondaire (1 = oui ; 2 = non ; 8 = ne sait pas/pas de réponse ; 9 = refus). Le statut tabagique a été défini en considérant les variables 'fréquence actuelle des cigarettes fumées' (SMK_CURRCG_COM, 1 = quotidiennement (au moins une cigarette par jour au cours des 30 derniers jours), 2 = occasionnellement (au moins une cigarette au cours des 30 derniers jours, mais pas tous les jours), 3 = pas du tout (vous n'avez pas du tout fumé au cours des 30 derniers jours), 8 = ne sait pas/pas de réponse, 9 = refus) et 'a déjà fumé des cigarettes quotidiennement' (SMK_EVRDL_COM, 1 = oui, 2 = non, 8 = ne sait pas/pas de réponse, 9 = refus). Nous avons défini le statut tabagique comme une variable binaire, fumeurs quotidiens actuels/anciens vs. les

autres).²³⁰ La consommation d'alcool (ALC_FREQ_COM), évaluée comme la fréquence de consommation d'alcool au cours des 12 derniers mois (01 = presque tous les jours (y compris 6 fois par semaine), 02 = 4-5 fois par semaine, 03 = 2-3 fois par semaine, 04 = une fois par semaine, 05 = 2-3 fois par mois, 06 = environ une fois par mois, 07 = moins d'une fois par mois, 96 = jamais, 98 = ne sait pas/pas de réponse, 99 = refus, 77 = manquant), a été dichotomisée en 'buveur' (au moins une fois par mois) vs. 'non-buveur' (moins d'une fois par mois).²³¹ Le niveau d'activité physique (PA2_DSCR2_MCQ) des participants, évalué au moyen de l'échelle d'activité physique pour les personnes âgées (*Physical Activity Scale for the Elderly, PASE*),²³² était une variable continue. Cette échelle comporte 12 questions auto-administrées permettant de quantifier (durée, fréquence, niveau d'effort et volume) l'activité physique effectuée au cours des 7 derniers jours par les personnes âgées de 65 ans et plus.^{232,233} Les activités évaluées comprennent les activités de loisir (marche à l'extérieur du domicile, activités sportives et récréatives légères, modérées et intenses, et exercices de force musculaire et/ou d'endurance), les activités domestiques (travaux domestiques légers, lourds ou corvées, réparations de la maison, entretien de la pelouse ou du jardin, jardinage en plein air et soins à autrui) ainsi que l'activité liée au travail (rémunéré ou bénévole).²³⁴ Les scores PASE vont de 0 à 793, et les valeurs les plus élevées sont indicatrices d'une activité physique plus importante.²³³

Outres ces covariables mesurées, nous avons également considéré les dix premières composantes principales génomiques issues de l'analyse en composantes principales du sous-ensemble des participants d'ascendance européenne afin d'ajuster nos analyses pour la stratification de la population.^{115,216}

3.3.4. Analyses statistiques

Nous avons décrit les caractéristiques des participants au recrutement sous-forme de moyenne (écart type) ou de fréquence (n, %) en distinguant les participants inclus des ceux qui ont été exclus de l'étude.

3.3.4.1. Analyses principales

Nous avons modélisé les associations conventionnelles entre l'IMC ainsi que la PAS (variable continue) et l'hypertension (variable binaire) en utilisant respectivement la régression linéaire et la

régression robuste de Poisson applicable aux événements non rares du fait de la prévalence élevée (32.1%) de l'hypertension chez nos participants.^{212,213} Nous avons ensuite effectué des analyses de RM en utilisant la méthode 2SLS pour la PAS et l'approche de la fonction de contrôle⁸¹ avec une régression robuste de Poisson à l'étape 2 lorsque l'hypertension était l'issue.²¹⁴ Les analyses ont utilisé respectivement *FTO* et le GRS comme instrument de l'IMC. Les effets estimés sont exprimés par augmentation de 1 kg/m² de l'IMC. Tous les modèles conventionnels et de RM ont été ajustés pour l'âge, le sexe et les dix premières composantes principales d'ascendance européenne. Les erreurs types ont été obtenues en utilisant l'estimateur sandwich.²³⁵

Nous avons considéré cinq méthodes de prise en compte de la médication antihypertensive décrites par Tobin et al.,⁵³ les trois premières utilisent chacune comme issue les valeurs de PAS mesurées dans l'ÉLCV. La première méthode, dénommée *Uncorrected SBP* ('PAS non corrigée') correspond à une analyse qui ignore la question de la médication antihypertensive. La deuxième méthode *Medication-adjusted SBP* ('PAS ajustée pour la médication') réfère aux modèles dans lesquels un indicateur binaire de la médication antihypertensive est incluse comme une covariable d'ajustement. La troisième méthode *Treated subjects excluded SBP* ('PAS excluant les sujets traités') est une analyse dans laquelle les participants traités avec les antihypertenseurs sont exclus de l'analyse. Les deux dernières méthodes correspondent à des transformations de la variable PAS. La quatrième méthode, *Fixed addition (15 mm Hg) SBP* ('Addition fixe de 15 mm Hg à la PAS') impute les valeurs mesurées de PAS des sujets traités avec des antihypertenseurs en ajoutant une valeur constante de 15 mm Hg. La cinquième méthode a considéré un indicateur binaire de l'hypertension qui combinait les valeurs mesurées de PAS et l'utilisation de la médication antihypertensive, comme précédemment décrit dans la section sur les variables.

3.3.4.2. Évaluation des hypothèses de la RM

La validité de la RM dépend de trois principales hypothèses : (i) l'instrument est associé à l'exposition, (ii) l'instrument n'est pas associé à l'issue due à la confusion, et (iii) l'instrument n'affecte pas l'issue indépendamment de l'exposition, hypothèse connue sous le nom de restriction d'exclusion.⁷ Pour l'hypothèse 1, nous avons évalué la force de chaque instrument par la statistique *F* de la régression linéaire de l'IMC sur l'instrument. Pour limiter les possibilités de violation de l'hypothèse 2, nous avons restreint l'analyse aux participants d'ascendance européenne et avons ajusté les modèles pour les dix premiers composantes principales.¹¹⁵

La principale source de violation de l'hypothèse 3 ou restriction d'exclusion est la pléiotropie,¹⁴ qui pourrait impliquer la variable médication antihypertensive. Nous avons évalué la plausibilité de la restriction d'exclusion en utilisant quatre méthodes qui fonctionnent selon différentes hypothèses.¹² Avant tout, nous avons recherché la présence éventuelle de pléiotropie directionnelle en utilisant le test de l'intercept de la régression MR-Egger basée sur les statistiques sommaires (coefficients betas et erreurs types) des 87 SNPs du GRS.²⁹ MR-Egger procède par la régression des coefficients des associations SNPs-*issue* sur les coefficients des associations SNPs-*exposition* pondérée par l'inverse de la variance des associations SNPs-*issue* en allouant un terme d'intercept non nul. Sous l'hypothèse d'indépendance entre les effets pléiotropiques directs et les associations SNPs-*exposition* (appelée hypothèse *InSIDE*, *Instrument Strength Independent of Direct Effect*), l'intercept de la régression MR-Egger peut être interprété comme l'effet pléiotropique moyen et la pente comme l'estimation de l'effet causal ajusté pour la pléiotropie.⁹⁶ La valeur p associée au test de l'intercept peut être utilisée pour évaluer la présence de pléiotropie directionnelle ; elle ne dépend pas de l'hypothèse *InSIDE*.⁹⁶

Par la suite, nous avons considéré trois autres méthodes applicables avec des instruments simples (un SNP ou GRS). Premièrement, nous avons appliqué la méthode basée sur la confusion positive proposée par Glymour et al.⁵ et que nous avons décrite dans les méthodes relatives à l'étude 2 de cette thèse. Deuxièmement, nous avons ajusté les associations entre les instruments et l'*issue* (PAS/hypertension) pour l'*exposition* (IMC) afin de vérifier si les coefficients estimés des instruments étaient différents de la valeur nulle, ce qui suggérerait une violation de l'hypothèse de restriction d'exclusion.³⁹ Troisièmement, enfin, nous avons utilisé une analyse de régression pour déterminer si les instruments étaient associés à l'indicateur binaire de la médication antihypertensive et/ou aux facteurs de confusion potentiels dans des modèles incluant également l'IMC. Une telle association suggérerait également une violation de l'hypothèse de restriction d'exclusion.

Les analyses ont été réalisées à l'aide des logiciels R version 4.2.1 et Stata/IC version 14.2 (StataCorp, College Station, Texas, USA) et ont inclus les poids analytiques fournis dans les jeux de données de l'ÉLCV.

3.3.5. Considérations éthiques

Tous les participants de l'ÉLCV étaient volontaires et ont donné leur consentement éclairé par écrit. Le protocole de l'ÉLCV a été approuvé par plusieurs comités d'éthique de la recherche à travers le Canada.²¹⁷ Le présent projet de thèse a été approuvé par le comité d'éthique de la recherche du Centre Hospitalier de l'Université de Montréal (numéro de projet 2021-8642, 20.148–YP).

Chapitre 4 – Résultats

4.1. Article 1

Mendelian Randomization: A Review of Methods for the Prevention, Assessment, and Discussion of Pleiotropy in Studies Using the Fat Mass and Obesity-Associated Gene as an Instrument for Adiposity

Fiston Ikwa Ndol Mbutiwi^{1,2}

Tatiana Dessy¹

Marie-Pierre Sylvestre^{1,3}

¹University of Montreal Hospital Research Centre (CRCHUM), Montreal, QC, Canada

²Faculty of Medicine, University of Kikwit, Kikwit, Democratic Republic of the Congo

³Department of Social and Preventive Medicine, University of Montreal Public Health School (ESPUM), Montreal, QC, Canada

Publié dans la revue

Frontiers in genetics 2022;13:803238. doi: 10.3389/fgene.2022.803238

Abstract

Pleiotropy assessment is critical for the validity of Mendelian randomization (MR) analyses, and its management remains a challenging task for researchers. This review examines how the authors of MR studies address bias due to pleiotropy in practice. We reviewed Pubmed, Medline, Embase and Web of Science for MR studies published before 21 May 2020 that used at least one single-nucleotide polymorphism (SNP) in the fat mass and obesity-associated (*FTO*) gene as instrumental variable (IV) for body mass index, irrespective of the outcome. We reviewed: 1) the approaches used to prevent pleiotropy, 2) the methods cited to detect or control the independence or the exclusion restriction assumption highlighting whether pleiotropy assessment was explicitly stated to justify the use of these methods, and 3) the discussion of findings related to pleiotropy. We included 128 studies, of which thirty-three reported one approach to prevent pleiotropy, such as the use of multiple (independent) SNPs combined in a genetic risk score as IVs. One hundred and twenty studies cited at least one method to detect or account for pleiotropy, including robust and other IV estimation methods ($n = 70$), methods for detection of heterogeneity between estimated causal effects across IVs ($n = 72$), methods to detect or account associations between IV and outcome outside thought the exposure ($n = 85$), and other methods ($n = 5$). Twenty-one studies suspected IV invalidity, of which 16 explicitly referred to pleiotropy, and six incriminating *FTO* SNPs. Most reviewed MR studies have cited methods to prevent or to detect or control bias due to pleiotropy. These methods are heterogeneous, their triangulation should increase the reliability of causal inference.

Keywords: *FTO*; Mendelian randomization; adiposity; body mass index; genetic risk score; instrumental variable; pleiotropy.

Introduction

Mendelian randomization (MR) is an instrumental variables (IVs) approach that exploits genetic variants (mostly single-nucleotide polymorphisms (SNPs)) as IVs of a non-genetic exposure to infer a causal relationship between this exposure and an outcome in observational studies (Lawlor et al., 2008). The validity of MR is based on three key assumptions: 1) the IV is associated with the exposure, 2) the association between the IV and the outcome is unconfounded, and 3) the IV only affects the outcome via the exposure, known as the exclusion restriction criterion (Labrecque and Swanson, 2018; Brumpton et al., 2020). Horizontal pleiotropy, the phenomenon whereby a genetic variant affects the exposure and the outcome through independent pathways and without being mediated by another (Davey Smith and Hemani, 2014; Hemani et al., 2018a; Jordan et al., 2019), is a primary cause of violation of the exclusion restriction criterion (Dixon et al., 2020). It may lead to biased causal effect estimates, reduced statistical power, and/or increased type I error (Bowden et al., 2016; Verbanck et al., 2018). We thereafter refer to “horizontal pleiotropy” as “pleiotropy” for the sake of brevity.

The increasing use of MR (Sekula et al., 2016) has prompted both subject-specific (Pingault et al., 2016; Frayling and Stoneman, 2018; Goodarzi, 2018; Lor et al., 2019; Meng et al., 2019; Guo et al., 2021) and general reviews (Bochud and Rousson, 2010; Davies et al., 2013; Boef et al., 2015) of MR studies summarizing the state of practice of MR in the last decade. These suggest that the exclusion restriction criterion is not systematically assessed or discussed. For example, in their meta-epidemiological overview on the approaches used in MR, Boef et al. (2015) noted that only 111 of 178 studies (62.4%) reported on the plausibility of the exclusion restriction criterion. However, no review thus far has focused on how authors prevent or minimize bias related to pleiotropy in MR studies. Such examination is important because while evidence suggests that pleiotropy is ubiquitous in the human genome (Boyle et al., 2017; Chesmore et al., 2018), the absence of pleiotropy in a MR study cannot be empirically proven (Glymour et al., 2012). While several approaches to detect pleiotropy and/or provide robust MR estimates have been recently proposed and compared (Bowden et al., 2015; Hartwig et al., 2017; Thompson et al., 2017; Hemani et al., 2018a; Verbanck et al., 2018; Rees et al., 2019; Burgess et al., 2020b; Zhao Q. et al., 2020; Minelli et al., 2021), their use in practice, including in sensitivity analyses and triangulation, has not been documented across studies. In light of the recent MR guidelines (Davey Smith et al.,

2019; Burgess et al., 2020a) that recommend assessing the robustness of MR results, we examine how potential bias due to pleiotropy is considered in the literature. More specifically we summarize 1) the approaches used to avoid selecting pleiotropic genetic variants, 2) the methods used to detect and account for pleiotropy in the estimation of the causal effect, and 3) how researchers discuss the exclusion restriction criterion considering their assessment of pleiotropy, including the impact on the results when pleiotropy is suspected.

Reviewing the entire body of published MR studies would not be practical. Instead, we limited our investigation to MR studies that use SNPs in the fat mass and obesity-associated (*FTO*) gene as IV to investigate the causal effect of adiposity on diverse outcomes. We expected studies that used *FTO* as an IV to provide a rich discussion of pleiotropy for several reasons. First, several SNPs in *FTO* have large and robust associations with body mass index (BMI) (Gill et al., 2019) and thus are considered strong IVs for adiposity and commonly used in MR studies. Second, unlike IVs such as variants on the C-reactive protein gene which encodes C-reactive protein (Burgess et al., 2017; Reactive Protein Coronary Heart Disease Genetics Collaboration et al., 2011), the exact biological pathways through which *FTO* affect adiposity are not fully understood, which complexifies the assessment of pleiotropy. Third, some *FTO* SNPs are suspected of pleiotropy with reported effects on a wide array of health issues ranging from cardiometabolic outcomes to cancer or mental health (Pausova et al., 2009; Delahanty et al., 2011; Hertel et al., 2011; Kivimäki et al., 2011; Li et al., 2012; Iles et al., 2013; Liu et al., 2013; Cronin et al., 2014; Aijala et al., 2015).

Methods

Search Strategy and Inclusion/Exclusion Criteria

We searched Pubmed, Medline, Embase and Web of Science to identify articles published before 21 May 2020 that met the following three criteria: 1) the primary analysis was MR, 2) the primary exposure was adiposity assessed by BMI, and 3) the IV(s) included at least one SNP in the *FTO* gene. We excluded studies in which MR was a secondary analysis to ensure that the study provided a detailed assessment of MR assumptions. We placed no restrictions on the outcome of interest. The search strategy and the specific exclusion criteria are provided in **Supplementary Methods S1, S2**, respectively.

Data Extraction and Analysis

For each study we recorded the use of a one-sample or two-sample MR design. We considered as one-sample both 1) studies that performed MR using individual level data on the SNPs, exposure and outcome and 2) studies that used summary statistics on SNP-exposure and SNP-outcome associations from the same sample (Hemani et al., 2018a). Two-sample data were defined as the use of summary statistics on SNP-exposure and SNP-outcome associations both from two distinct samples (Hemani et al., 2018a). We documented three types of IVs including 1) single IV, 2) genetic risk scores (GRS) that aggregate several SNPs into a single variable that corresponds to a weighted or unweighted sum of risk alleles (Burgess and Thompson, 2013; Burgess et al., 2016) and 3) multiple IVs. We defined multiple IVs as the use of ≥ 2 SNPs as separate IVs in a single model or to the combination of estimated effects from ≥ 2 single SNPs into one summary causal effect using meta-analytic techniques (Burgess et al., 2016). We recorded the specific *FTO* SNP(s) used in each of the types of IVs described above.

We organized the data analysis around three themes that described how pleiotropy was handled in the analytical process including 1) the selection and combinations of IVs; 2) the methods used to detect and account for pleiotropy in the estimation of causal effects, and 3) the discussion of findings considering the assessment of pleiotropy. We documented the methods as they were explicitly stated in each article irrespective of their applicability or relevance. A critical appraisal of the use of some of the methods is provided in the Discussion.

Approaches to Prevent Pleiotropy (Selection and Combination of IVs)

According to MR guidelines (Burgess et al., 2020a), SNPs can either be selected from gene regions that specifically encodes the exposure (biological approach) or on the basis of their statistical association with the exposure of interest (statistical approach). SNPs known or suspected of pleiotropy may be excluded from the initial selection before performing the main analysis (Burgess et al., 2020a). We reported the use of the biological and or statistical approach to SNP selection. We also documented the use of multiple independent SNPs as IVs as a method to attenuate the effect of pleiotropy under the assumption that the pleiotropic effects of SNPs would be balanced and thus cancel each other out (Davey Smith, 2011). Finally, we recorded any other strategy explicitly presented as pertaining to the selection of IVs to minimize the presence of pleiotropy.

Approaches to Detect and/or Account for Pleiotropy in the Estimation of the Causal Effect

MR guidelines require authors to report on the methods used to evaluate MR assumptions, which includes investigating bias due to pleiotropy (Davey Smith et al., 2019; Burgess et al., 2020a). We recorded the methods used to evaluate the independence and the exclusion restriction assumptions with the exception of those used for population stratification, highlighting whether pleiotropy assessment was explicitly stated to justify the use of these methods. We organized methods into four categories including 1) robust (e.g., MR-Egger) and other IV estimation methods (e.g., multivariable MR), 2) methods to detect heterogeneity of estimated causal effects across IVs (e.g., statistical tests of heterogeneity between the estimated SNP-specific causal effects), 3) methods to detect or account for associations between the IVs and the outcome that arise through pathways outside of the exposure (e.g., mediation analysis), and 4) other methods. Robust methods provide causal effect estimates under a weaker set of assumptions than conventional methods (Burgess and Thompson, 2017; Burgess et al., 2020a). A summary description of the methods, including their main assumptions and limitations is presented in **Supplementary Table S6**.

Discussion of Findings Considering the Assessment of Pleiotropy

We verified whether the authors discussed the independence and the exclusion restriction assumptions, distinguishing studies that explicitly referred to the term “pleiotropy” from those that did not. When studies suspected IV invalidity, we further verified whether the authors report the impact of IV invalidity on MR results, and if any *FTO* SNPs were incriminated.

The data were tabulated using Microsoft Excel[®] 2016 and described using Stata/IC version 14.2 software (StataCorp, College Station, Texas, United States).

Results

Study Selection and Characteristics of Studies

Our search identified 2,985 publications, of which 128 articles were included upon completion of the screening process (**Figure 1**). The 128 articles are listed in **Supplementary Table S1**. Included articles were published between June 2008 and May 2020, and mostly comprised one-sample MR analyses ($n = 98$, 76.6%; **Table 1** and **Supplementary Table S3**). A total of 31 *FTO* SNPs were

selected as IVs with rs1558902 being the most frequently used ($n = 64$, 50%; **Supplementary Table S2**). While 74 studies (57.8%) used a GRS to represent the IVs (**Table 1** and **Supplementary Table S3**), **Figure 2** suggests a recent decline in the use of GRS in favor of multiple IVs.

Approaches to Prevent Pleiotropy (Selection and Combination of IVs)

While all of the 128 studies selected IVs on the basis of statistical association between SNPs and BMI in the literature, 33 studies (25.8%) proceeded further in their attempt to prevent the effect of pleiotropy in the selection of the IVs (**Table 2** and **Supplementary Table S4**). Of the 33 studies, 12 excluded previously selected SNPs known or suspected of pleiotropy. Ten articles respectively cited the use of a GRS and of multiple (independent) IVs (even if six of these ten in fact used a GRS as IVs) to prevent the effect of pleiotropy, a single study justified the use of a single SNP as IV as a method to prevent pleiotropy, although these strategies do not guarantee that bias due to pleiotropy is prevented or reduced (Burgess and Thompson, 2013).

Approaches to Detect and Account for Pleiotropy in the Estimation of the Causal Effect

Overall, 120 of 128 included studies assessed the plausibility of the independence and/or exclusion restriction assumptions (**Table 3** and **Supplementary Table S5**). Of the 120 studies, 78 reported using more than one category of methods within our classification (robust and other IV estimation methods ($n = 70$), heterogeneity ($n = 72$), alternative pathways ($n = 85$) and others ($n = 5$, including the use of positive or negative control outcomes, colocalization, and verifying the concordance of MR results with those from other studies). A total of 95 studies explicitly cited pleiotropy to justify the use of such methods. MR-Egger was the most frequently reported method to assess pleiotropy ($n = 68$). Of the 68 studies, 66 used the intercept p -value as a test of the validity of IVs and 29 studies (Gao et al., 2016; Hartwig et al., 2016; Kemp et al., 2016; Mokry et al., 2016; Tyrrell et al., 2016; Censin et al., 2017; Dale et al., 2017; Lindstrom et al., 2017; Lyall et al., 2017; Noyce et al., 2017; Xu et al., 2017; Larsson et al., 2018; van den Broek et al., 2018; Wang T. et al., 2018; Bae and Lee, 2019; Brower et al., 2019; Censin et al., 2019; Gharahkhani et al., 2019; Howe et al., 2019; Taylor et al., 2019; Tyrrell et al., 2019; Xu et al., 2019; Dusingize et al., 2020; Kurz and Laxy, 2020; Larsson et al., 2020; Liu et al., 2020; Winter-Jensen et al., 2020; Zhang et al.,

2020; Zhao Y. et al., 2020) further acknowledged that MR-Egger estimates or tests may be underpowered. Two studies (Tyrrell et al., 2016; Fan et al., 2018) compared the MR-Egger slope and the conventional MR causal effect estimate while the last study (Chen et al., 2019) did not specify how the MR-Egger results were used. Forty-six studies assessed pleiotropy by evaluating the heterogeneity of estimated causal effects across IVs, whether by graphical assessment ($n = 20$) or statistical testing ($n = 7$), or by comparing estimated causal effects from GRS or multiple IVs before vs. after exclusion of suspected pleiotropic SNP(s) ($n = 20$). A total of 30 studies attempted to detect pleiotropy by estimating pathways through which the IVs were associated with the outcomes outside of that implicating the exposure. Such studies mostly reported the estimation of associations between the IVs and measured risk factors of the outcome ($n = 19$), but also adjusted the IV-outcome or IV-confounders associations for exposure ($n = 7$) or documented the associations between the IV and risk factors for the outcome in the literature ($n = 5$).

Discussion of Findings Considering the Assessment of Pleiotropy

Of 128 included articles, 108 discussed the plausibility of the independence and/or exclusion restriction assumptions of which 89 studies made an explicit reference to pleiotropy (**Table 4** and **Supplementary Table S7**). Invalid IVs were suspected in 21 studies (**Supplementary Table S8**), 16 of which cited pleiotropy as a potential source of invalidity. Eight of the 21 studies concluded that the MR results were possibly invalid, while nine studies reported that the results were robust. The remaining four studies did not discuss the impact of IV invalidity on the results. Six studies suggested that at least one *FTO* SNP (rs1558902, rs1421085, and rs17817449) was suspected of pleiotropy on the basis of sensitivity analyses (**Table 4, Supplementary Tables S7, S8**).

Discussion

Pleiotropy is considered widespread in humans (Boyle et al., 2017) and thus presents a major challenge to the validity of MR studies, especially considering the limited knowledge of the biological function of many of the SNPs used as IVs (Danchin and Fang, 2016; Swerdlow et al., 2016). We reviewed studies that used SNP(s) in the *FTO* gene as IV(s) in MR to examine the strategies employed by authors to prevent, detect or control, and discuss biases due to the use of pleiotropic IVs. Our review extends the overview of statistical approaches used in MR published

by Boef et al. (2015) by focusing on pleiotropy and by including the recent developments such as two-sample MR and the use of MR-Egger. We observed that the vast majority of studies addressed pleiotropy by using several methods that operate under different assumptions (Lawlor et al., 2016; Hemani et al., 2018b). While most authors invoked pleiotropy at the analytical stage to justify the use of detection tools and robust methods, explicit attention was also given to the prevention of pleiotropy in the selection of IVs in a fourth of the articles reviewed and pleiotropy was mentioned in the discussion in 70% of articles.

Our review highlighted three observations that merit attention for future MR studies. First, we documented an increasing use of multiple IVs over time, in addition to the exclusive use of statistical criteria to select IVs from the literature. This is largely explained by our focus on BMI as the exposure, since BMI is a polygenetic trait without a specific proximal coding gene from which to select SNPs as it is common with protein-like exposures (Swerdlow et al., 2016; Burgess and Thompson, 2017). The increasing use of multiple SNPs is also motivated by attempts to increase the strength of the instruments, the availability of large-scale genome-wide association studies to select instruments from, and by the development of MR methods that require multiple IVs to provide robust MR estimates under a less stringent set of assumptions (Burgess and Thompson, 2017; Burgess et al., 2020a). However, selecting SNPs exclusively by statistical approach increases the likelihood of including pleiotropic SNPs which may lead to biased MR results (Hartwig et al., 2017; Bowden and Holmes, 2019). Advantages of using multiple IVs independently in a MR study include the use of robust methods, such as MR-Egger or median or mode-based methods (Slob and Burgess, 2020). On the other hand, the use of GRS, which was the most frequent method to combine SNPs in the studies that we reviewed, is convenient because it leads to a single IV (Burgess et al., 2016). Further, weighted GRS with independently-derived weights lead to MR studies with similar statistical power than those using multiple independent IVs (Palmer et al., 2012). While several studies justified the use of a GRS as IV as a method to prevent bias due to pleiotropy, they have to rely on the restrictive assumption that the pleiotropic effects of SNPs cancel each other (Davey Smith, 2011), which is difficult to verify in practice. Using a GRS as IV further requires ensuring that each SNP in the GRS is itself a valid IV (Burgess and Thompson, 2013; Skaaby et al., 2018), which is limited by the low statistical power available for each SNP. Simulation studies have demonstrated that even including a small number of

pleiotropic SNPs into a GRS can lead to biased MR estimates (Burgess and Thompson, 2013). We thus recommend that robust methods be used on the SNPs that form the GRS.

Second, our review suggests that the MR-Egger intercept test from is the most frequently reported method for pleiotropy assessment. The validity of MR-Egger estimates and the interpretation of the intercept as the average pleiotropic effect of IVs require the InSIDE (Instrument Strength Independent of Direct Effect) assumption to be satisfied (Burgess and Thompson, 2017). InSIDE is not required for the use of the p -value associated with the intercept test of the validity of IVs (Burgess and Thompson, 2017). InSIDE states that the effects of the IVs on the exposure must be uncorrelated with the direct effects of the IVs on the outcome (Bowden et al., 2015; Burgess et al., 2017), which is likely to be violated in a one-sample setting (Slob et al., 2017; Minelli et al., 2021) because parameters are estimated in the same subjects. Violations of the InSIDE assumption results in increased type I error rates (Hartwig and Davies, 2016; Burgess et al., 2017) and biased estimates in the direction of the observational associations (Minelli et al., 2021). Because testing the plausibility of InSIDE assumption is still a challenge to date (Bowden, 2017), researchers should restrict the interpretation of MR-Egger estimates in two sample settings where the lack of correlation between SNPs-exposure and SNPs-outcome associations is more plausible. Thus, when using summarized data in a one-sample setting, the MR-Egger intercept test can be used to assess the invalidity of the IVs, but other robust methods such as the median (the second robust method widely reported in this review)- and the mode-based methods should be preferred when estimating the robust causal estimates because they do not depend on the InSIDE assumption (Bowden, 2017). Further, their causal estimates are consistent in one-sample context, unlike Egger's estimates which are biased in the direction of the observational association, as shown in simulations (Minelli et al., 2021). Burgess and Thompson offer a careful discussion of the use of MR-Egger (Burgess and Thompson, 2017), while Burgess et al. (2018) show that the statistical power for the intercept test is low in most settings.

Third, using statistical approaches to detect pleiotropic IVs is challenging because apparent manifestations of pleiotropy may be confused with other phenomena, some of which not invalidating MR or requiring different approaches than pleiotropy. For example, the assessment of heterogeneity in the estimated causal effects across different IVs is based on the principle that if IVs are valid, the variation in their corresponding MR estimates should be due to chance (Greco et

al., 2015). Large variations in MR IV-specific estimates are often considered as indicative of pleiotropy, but can be due to other causes such as the non-collapsibility of odds ratios in case of MR analysis with binary outcome (Vansteelandt et al., 2011; Hemani et al., 2018a), heterogeneity in the distribution of confounders of IV-exposure or IV-outcome associations in two-sample settings (Hemani et al., 2018a; Zhao Q. et al., 2020), or differential complier causal effects, i.e., association between the IVs and the exposure that vary importantly across individuals (Baiocchi et al., 2014; Sainani, 2018). Similarly, reasons other than pleiotropy may explain non-null associations between the IVs and the outcome that may be considered indicative of pleiotropic IVs. For example, population stratification which can be addressed by restricting the sample to homogeneous ancestry or by applying correction methods (e.g., adjustment of MR models for principal components) (Davey Smith and Hemani, 2014). Additional causes of violation of the exclusion restriction that can be confused with pleiotropy have been proposed, including an exposure that varies over time, the presence of gene-environment interactions implicating IVs, and linkage disequilibrium between at least one of the IVs and a SNP that also affects the outcome (VanderWeele et al., 2014).

Our review also allows a few observations pertaining to the use of SNPs in the *FTO* gene as IVs for BMI. Four of the six studies that suspected that *FTO* SNPs used as IVs might be pleiotropic involved mental health phenotypes [e.g., subjective well-being (van den Broek et al., 2018) or phobic anxiety symptoms (Walter et al., 2015a), or common mental disorders (Kivimäki et al., 2011), including depression (Walter et al., 2015b)]. This suggests that *FTO* may be associated with mental health through pathways that do not involve BMI, a hypothesis that is supported by animal studies (Hess et al., 2013; Sun et al., 2019). For example, *FTO* regulates the activity of the dopaminergic signaling pathways related to the regulation of learning, reward behavior, motor functions, and feeding in mice (Hess et al., 2013). Furthermore, other work on *FTO*-deficient mice suggested that *FTO* could influence anxiety- and depression-like behaviours via alterations in gut microbiota (Sun et al., 2019). Caution is required regarding the use of *FTO* as an IV for BMI implicating mental health phenotypes.

Two limitations of the current review should be noted. First, we do not present an exhaustive list of the methods to prevent, detect or control, and discuss biases due to the use of pleiotropic IVs. Rather we focus on the methods reported in the 128 studies that we review. While we captured

most of the methods that are currently available, newer methods such as the Causal Analysis Using Summary Effect estimates (CAUSE) (Morrison et al., 2020) and MR analysis using mixture-model (MRMix) (Qi and Chatterjee, 2019) are not reported. Second, the methods reported in this review include common and validated methods, as well as methods or strategies that may be less efficient/optimal for detecting or accounting bias due to pleiotropy in MR studies. Users must exert caution in selecting the best method(s) for the data at hand.

Pleiotropy is a ubiquitous phenomenon that poses a threat to the validity of MR results and that is difficult to assess. MR-related methodological development is thriving, and users are encouraged to use more than one method to assess pleiotropy, heeding the assumptions required for each.

Author Contributions

FM and M-PS designed the study. FM carried out the literature search and data extraction and analysis. All authors participated to the interpretation of the data. FM drafted the manuscript under the supervision of M-PS and TD translated it in English. M-PS and TD critically reviewed the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by a grant from the Canadian Institutes of Health Research (grant number 364471). FM was supported by a PhD scholarship from the Fonds de recherche du Québec–Santé (FRQS) (grant number 36057). M-PS was supported by a Chercheur-Boursier Junior II career award from the Fonds de recherche du Québec–Santé (FRQS) (grant number 296647).

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Aijala, M., Ronkainen, J., Huusko, T., Malo, E., Savolainen, E.-R., Savolainen, M. J., et al. (2015). The Fat Mass and Obesity-Associated (FTO) Gene Variant Rs9939609 Predicts Long-Term Incidence of Cardiovascular Disease and Related Death Independent of the Traditional Risk Factors. *Ann. Med.* 47, 655–663. doi:10.3109/07853890.2015.1091088
- Bae, S. C., and Lee, Y. H. (2019). Causal Association between Body Mass index and Risk of Rheumatoid Arthritis: A Mendelian Randomization Study. *Eur. J. Clin. Invest.* 49, e13076. doi:10.1111/eci.13076
- Baiocchi, M., Cheng, J., and Small, D. S. (2014). Instrumental Variable Methods for Causal Inference. *Statist. Med.* 33, 2297–2340. doi:10.1002/sim.6128
- Bochud, M., and Rousson, V. (2010). Usefulness of Mendelian Randomization in Observational Epidemiology. *Ijerph* 7, 711–728. doi:10.3390/ijerph7030711
- Boef, A. G. C., Dekkers, O. M., and le Cessie, S. (2015). Mendelian Randomization Studies: a Review of the Approaches Used and the Quality of Reporting. *Int. J. Epidemiol.* 44, 496–511. doi:10.1093/ije/dyv071
- Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian Randomization with Invalid Instruments: Effect Estimation and Bias Detection through Egger Regression. *Int. J. Epidemiol.* 44, 512–525. doi:10.1093/ije/dyv080
- Bowden, J., Del Greco, M. F., Minelli, C., Davey Smith, G., Sheehan, N. A., and Thompson, J. (2016). Assessing the Suitability of Summary data for Two-Sample Mendelian Randomization Analyses Using MR-Egger Regression: The Role of the I^2 Statistic. *Int. J. Epidemiol.* 45, 1961–1974. doi:10.1093/ije/dyw220
- Bowden, J., and Holmes, M. V. (2019). Meta-analysis and Mendelian Randomization: A Review. *Res. Syn Meth* 10, 486–496. doi:10.1002/jrsm.1346
- Bowden, J. (2017). Misconceptions on the Use of MR-Egger Regression and the Evaluation of the InSIDE assumption. *Int. J. Epidemiol.* 46, 2097–2099. doi:10.1093/ije/dyx192
- Boyle, E. A., Li, Y. I., and Pritchard, J. K. (2017). An Expanded View of Complex Traits: from Polygenic to Omnigenic. *Cell* 169, 1177–1186. doi:10.1016/j.cell.2017.05.038

- Brower, M. A., Hai, Y., Jones, M. R., Guo, X., Chen, Y.-D. I., Rotter, J. I., et al. (2019). Bidirectional Mendelian Randomization to Explore the Causal Relationships between Body Mass Index and Polycystic Ovary Syndrome. *Hum. Reprod.* 34, 127–136. doi:10.1093/humrep/dey343
- Brumpton, B., Sanderson, E., Heilbron, K., Hartwig, F. P., Harrison, S., Vie, G. A., et al. (2020). Avoiding Dynastic, Assortative Mating, and Population Stratification Biases in Mendelian Randomization through Within-Family Analyses. *Nat. Commun.* 11, 3519. doi:10.1038/s41467-020-17117-4
- Burgess, S., Bowden, J., Dudbridge, F., and Thompson, S. G. (2018). *Assessing the Effectiveness of Robust Instrumental Variable Methods Using Multiple Candidate Instruments with Application to Mendelian Randomization*. arXiv, 1606.03729v03722.
- Burgess, S., Bowden, J., Fall, T., Ingelsson, E., and Thompson, S. G. (2017). Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology* 28, 30–42. doi:10.1097/EDE.0000000000000559
- Burgess, S., Davey Smith, G., Davies, N. M., Dudbridge, F., Gill, D., Glymour, M. M., et al. (2020a). Guidelines for Performing Mendelian Randomization Investigations. *Wellcome Open Res.* 4, 186. doi:10.12688/wellcomeopenres.15555.2
- Burgess, S., Dudbridge, F., and Thompson, S. G. (2016). Combining Information on Multiple Instrumental Variables in Mendelian Randomization: Comparison of Allele Score and Summarized Data Methods. *Statist. Med.* 35, 1880–1906. doi:10.1002/sim.6835
- Burgess, S., Foley, C. N., Allara, E., Staley, J. R., and Howson, J. M. M. (2020b). A Robust and Efficient Method for Mendelian Randomization with Hundreds of Genetic Variants. *Nat. Commun.* 11, 376. doi:10.1038/s41467-019-14156-4
- Burgess, S., Scott, R. A., Scott, R. A., Timpson, N. J., Davey Smith, G., and Thompson, S. G. Epic-InterAct Consortium (2015). Using Published Data in Mendelian Randomization: a Blueprint for Efficient Identification of Causal Risk Factors. *Eur. J. Epidemiol.* 30, 543–552. doi:10.1007/s10654-015-0011-z

- Burgess, S., and Thompson, S. G. (2017). Interpreting Findings from Mendelian Randomization Using the MR-Egger Method. *Eur. J. Epidemiol.* 32, 377–389. doi:10.1007/s10654-017-0255-x
- Burgess, S., and Thompson, S. G. (2013). Use of Allele Scores as Instrumental Variables for Mendelian Randomization. *Int. J. Epidemiol.* 42, 1134–1144. doi:10.1093/ije/dyt093
- Censin, J. C., Nowak, C., Cooper, N., Bergsten, P., Todd, J. A., and Fall, T. (2017). Childhood Adiposity and Risk of Type 1 Diabetes: A Mendelian Randomization Study. *Plos Med.* 14, e1002362. doi:10.1371/journal.pmed.1002362
- Censin, J. C., Peters, S. A. E., Bovijn, J., Ferreira, T., Pulit, S. L., Mägi, R., et al. (2019). Causal Relationships between Obesity and the Leading Causes of Death in Women and Men. *Plos Genet.* 15, e1008405. doi:10.1371/journal.pgen.1008405
- Chen, Y.-C., Fan, H.-Y., Yang, C., Hsieh, R.-H., Pan, W.-H., and Lee, Y. L. (2019). Assessing Causality between Childhood Adiposity and Early Puberty: A Bidirectional Mendelian Randomization and Longitudinal Study. *Metabolism* 100, 153961. doi:10.1016/j.metabol.2019.153961
- Chesmore, K., Bartlett, J., and Williams, S. M. (2018). The Ubiquity of Pleiotropy in Human Disease. *Hum. Genet.* 137, 39–44. doi:10.1007/s00439-017-1854-z
- Cronin, R. M., Field, J. R., Bradford, Y., Shaffer, C. M., Carroll, R. J., Mosley, J. D., et al. (2014). Phenome-wide Association Studies Demonstrating Pleiotropy of Genetic Variants within FTO with and without Adjustment for Body Mass index. *Front. Genet.* 5, 250. doi:10.3389/fgene.2014.00250
- Dale, C. E., Fatemifar, G., Palmer, T. M., White, J., Prieto-Merino, D., Zabaneh, D., et al. (2017). Causal Associations of Adiposity and Body Fat Distribution with Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus. *Circulation* 135, 2373–2388. doi:10.1161/CIRCULATIONAHA.116.026560
- Danchin, A., and Fang, G. (2016). Unknown Unknowns: Essential Genes in Quest for Function. *Microb. Biotechnol.* 9, 530–540. doi:10.1111/1751-7915.12384

- Davey Smith, G., Davies, N. M., Dimou, N., Egger, M., Gallo, V., Golub, R., et al. (2019). STROBE-MR: Guidelines for Strengthening the Reporting of Mendelian Randomization Studies. *PeerJ Preprints* 7, e27857v27851. doi:10.7287/peerj.preprints.27857v1
- Davey Smith, G., and Hemani, G. (2014). Mendelian Randomization: Genetic Anchors for Causal Inference in Epidemiological Studies. *Hum. Mol. Genet.* 23, R89–R98. doi:10.1093/hmg/ddu328
- Davey Smith, G. (2011). Random allocation in observational data: how small but robust effects could facilitate hypothesis-free causal inference. *Epidemiology* 22, 460–463. doi:10.1097/EDE.0b013e31821d0426
- Davies, N. M., Smith, G. D., Windmeijer, F., and Martin, R. M. (2013). Issues in the Reporting and Conduct of Instrumental Variable Studies. *Epidemiology* 24, 363–369. doi:10.1097/EDE.0b013e31828abafb
- Delahanty, R. J., Beeghly-Fadiel, A., Xiang, Y.-B., Long, J., Cai, Q., Wen, W., et al. (2011). Association of Obesity-Related Genetic Variants with Endometrial Cancer Risk: a Report from the Shanghai Endometrial Cancer Genetics Study. *Am. J. Epidemiol.* 174, 1115–1126. doi:10.1093/aje/kwr233
- Dixon, P., Hollingworth, W., Harrison, S., Davies, N. M., and Davey Smith, G. (2020). Mendelian Randomization analysis of the causal effect of adiposity on hospital costs. *J. Health Econ.* 70, 102300. doi:10.1016/j.jhealeco.2020.102300
- Dusingize, J. C., Olsen, C. M., An, J., Pandeya, N., Law, M. H., Thompson, B. S., et al. (2020). Body Mass index and Height and Risk of Cutaneous Melanoma: Mendelian Randomization Analyses. *Int. J. Epidemiol.* 49, 1236–1245. doi:10.1093/ije/dyaa009
- Fan, H.-Y., Huang, Y.-T., Hsieh, R.-H., Chao, J. C.-J., Tung, Y.-C., Lee, Y. L., et al. (2018). Birthweight, Time-Varying Adiposity Growth and Early Menarche in Girls: A Mendelian Randomisation and Mediation Analysis. *Obes. Res. Clin. Pract.* 12, 445–451. doi:10.1016/j.orcp.2018.07.008
- Frayling, T. M., and Stoneman, C. E. (2018). Mendelian Randomisation in Type 2 Diabetes and Coronary Artery Disease. *Curr. Opin. Genet. Development* 50, 111–120. doi:10.1016/j.gde.2018.05.010

- Gao, C., Patel, C. J., Michailidou, K., Peters, U., Gong, J., Schildkraut, J., et al. (2016). Mendelian Randomization Study of Adiposity-Related Traits and Risk of Breast, Ovarian, Prostate, Lung and Colorectal Cancer. *Int. J. Epidemiol.* 45, 896–908. doi:10.1093/ije/dyw129
- Gharahkhani, P., Ong, J.-S., An, J., Law, M. H., Whiteman, D. C., Neale, R. E., et al. (2019). Effect of Increased Body Mass index on Risk of Diagnosis or Death from Cancer. *Br. J. Cancer* 120, 565–570. doi:10.1038/s41416-019-0386-9
- Gianfrancesco, M. A., Glymour, M. M., Walter, S., Rhead, B., Shao, X., Shen, L., et al. (2017). Causal Effect of Genetic Variants Associated with Body Mass Index on Multiple Sclerosis Susceptibility. *Am. J. Epidemiol.* 185, 162–171. doi:10.1093/aje/kww120
- Gill, R., Stratigopoulos, G., Lee, J. H., and Leibel, R. L. (2019). Functional Genomic Characterization of the FTO Locus in African Americans. *Physiol. Genomics* 51, 517–528. doi:10.1152/physiolgenomics.00057.2019
- Glymour, M. M., Tchetgen Tchetgen, E. J., and Robins, J. M. (2012). Credible Mendelian Randomization Studies: Approaches for Evaluating the Instrumental Variable Assumptions. *Am. J. Epidemiol.* 175, 332–339. doi:10.1093/aje/kwr323
- Goodarzi, M. O. (2018). Genetics of Obesity: what Genetic Association Studies Have Taught Us about the Biology of Obesity and its Complications. *Lancet Diabetes Endocrinol.* 6, 223–236. doi:10.1016/S2213-8587(17)30200-0
- Greco M, F. D., Minelli, C., Sheehan, N. A., and Thompson, J. R. (2015). Detecting Pleiotropy in Mendelian Randomisation Studies with Summary Data and a Continuous Outcome. *Statist. Med.* 34, 2926–2940. doi:10.1002/sim.6522
- Guo, J.-Z., Xiao, Q., Gao, S., Li, X.-Q., Wu, Q.-J., and Gong, T.-T. (2021). Review of Mendelian Randomization Studies on Ovarian Cancer. *Front. Oncol.* 11, 681396. doi:10.3389/fonc.2021.681396
- Guo, Y., Warren Andersen, S., Shu, X. O., Michailidou, K., Bolla, M. K., Wang, Q., et al. (2016). Genetically Predicted Body Mass Index and Breast Cancer Risk: Mendelian Randomization Analyses of Data from 145,000 Women of European Descent. *Plos Med.* 13, e1002105. doi:10.1371/journal.pmed.1002105

- Hartwig, F. P., Bowden, J., Loret de Mola, C., Tovo-Rodrigues, L., Davey Smith, G., and Horta, B. L. (2016). Body Mass index and Psychiatric Disorders: a Mendelian Randomization Study. *Sci. Rep.* 6, 32730. doi:10.1038/srep32730
- Hartwig, F. P., Davey Smith, G., and Bowden, J. (2017). Robust Inference in Summary Data Mendelian Randomization via the Zero Modal Pleiotropy assumption. *Int. J. Epidemiol.* 46, 1985–1998. doi:10.1093/ije/dyx102
- Hartwig, F. P., and Davies, N. M. (2016). Why Internal Weights Should Be Avoided (Not Only) in MR-Egger Regression. *Int. J. Epidemiol.* 45, 1676–1678. doi:10.1093/ije/dyw240
- Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., et al. (2018b). The MR-Base Platform Supports Systematic Causal Inference across the Human Phenome. *ELife* 7, e34408. doi:10.7554/eLife.34408.001
- Hemani, G., Bowden, J., and Davey Smith, G. (2018a). Evaluating the Potential Role of Pleiotropy in Mendelian Randomization Studies. *Hum. Mol. Genet.* 27, R195–R208. doi:10.1093/hmg/ddy163
- Hertel, J. K., Johansson, S., Sonestedt, E., Jonsson, A., Lie, R. T., Platou, C. G. P., et al. (2011). FTO, Type 2 Diabetes, and Weight Gain throughout Adult Life. *Diabetes* 60, 1637–1644. doi:10.2337/db10-1340
- Hess, M. E., Hess, S., Meyer, K. D., Verhagen, L. A. W., Koch, L., Brönneke, H. S., et al. (2013). The Fat Mass and Obesity Associated Gene (Fto) Regulates Activity of the Dopaminergic Midbrain Circuitry. *Nat. Neurosci.* 16, 1042–1048. doi:10.1038/nn.3449
- Howe, L. D., Kanayalal, R., Harrison, S., Beaumont, R. N., Davies, A. R., Frayling, T. M., et al. (2019). Effects of Body Mass index on Relationship Status, Social Contact and Socio-Economic Position: Mendelian Randomization and Within-Sibling Study in UK Biobank. *Int. J. Epidemiol.* 49, 1173–1184. doi:10.1093/ije/dyz240
- Iles, M. M., Law, M. H., Stacey, S. N., Han, J., Fang, S., Pfeiffer, R., et al. (2013). A Variant in FTO Shows Association with Melanoma Risk Not Due to BMI. *Nat. Genet.* 45, 428–432. doi:10.1038/ng.2571

- Jordan, D. M., Verbanck, M., and Do, R. (2019). HOPS: a Quantitative Score Reveals Pervasive Horizontal Pleiotropy in Human Genetic Variation Is Driven by Extreme Polygenicity of Human Traits and Diseases. *Genome Biol.* 20, 222. doi:10.1186/s13059-019-1844-7
- Kemp, J. P., Sayers, A., Smith, G. D., Tobias, J. H., and Evans, D. M. (2016). Using Mendelian Randomization to Investigate a Possible Causal Relationship between Adiposity and Increased Bone mineral Density at Different Skeletal Sites in Children. *Int. J. Epidemiol.* 45, 1560–1572. doi:10.1093/ije/dyw079
- Kivimäki, M., Jokela, M., Hamer, M., Geddes, J., Ebmeier, K., Kumari, M., et al. (2011). Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis: The Whitehall II Study, 1985-2004. *Am. J. Epidemiol.* 173, 421–429. doi:10.1093/aje/kwq444
- Kurz, C. F., and Laxy, M. (2020). Application of Mendelian Randomization to Investigate the Association of Body Mass Index with Health Care Costs. *Med. Decis. Making* 40, 156–169. doi:10.1177/0272989X20905809
- Labrecque, J., and Swanson, S. A. (2018). Understanding the Assumptions Underlying Instrumental Variable Analyses: a Brief Review of Falsification Strategies and Related Tools. *Curr. Epidemiol. Rep.* 5, 214–220. doi:10.1007/s40471-018-0152-1
- Larsson, S. C., Bäck, M., Rees, J. M. B., Mason, A. M., and Burgess, S. (2020). Body Mass index and Body Composition in Relation to 14 Cardiovascular Conditions in UK Biobank: a Mendelian Randomization Study. *Eur. Heart J.* 41, 221–226. doi:10.1093/eurheartj/ehz388
- Larsson, S. C., Burgess, S., and Michaëlsson, K. (2018). Genetic Association between Adiposity and Gout: a Mendelian Randomization Study. *Rheumatology* 57, 2145–2148. doi:10.1093/rheumatology/key229
- Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N., and Davey Smith, G. (2008). Mendelian Randomization: Using Genes as Instruments for Making Causal Inferences in Epidemiology. *Statist. Med.* 27, 1133–1163. doi:10.1002/sim.3034
- Lawlor, D. A., Tilling, K., and Davey Smith, G. (2016). Triangulation in Aetiological Epidemiology. *Int. J. Epidemiol.* 45, dyw314–1886. doi:10.1093/ije/dyw314

- Li, H., Kilpeläinen, T. O., Liu, C., Zhu, J., Liu, Y., Hu, C., et al. (2012). Association of Genetic Variation in FTO with Risk of Obesity and Type 2 Diabetes with Data from 96,551 East and South Asians. *Diabetologia* 55, 981–995. doi:10.1007/s00125-011-2370-7
- Lindström, S., Germain, M., Germain, M., Crous-Bou, M., Smith, E. N., Morange, P.-E., et al. (2017). Assessing the Causal Relationship between Obesity and Venous Thromboembolism through a Mendelian Randomization Study. *Hum. Genet.* 136, 897–902. doi:10.1007/s00439-017-1811-x
- Liu, C., Mou, S., and Pan, C. (2013). The FTO Gene Rs9939609 Polymorphism Predicts Risk of Cardiovascular Disease: a Systematic Review and Meta-Analysis. *PLoS One* 8, e71901. doi:10.1371/journal.pone.0071901
- Liu, Q., Pan, J., Berzuini, C., Rutter, M. K., and Guo, H. (2020). Integrative Analysis of Mendelian Randomization and Bayesian Colocalization Highlights Four Genes with Putative BMI-Mediated Causal Pathways to Diabetes. *Sci. Rep.* 10, 7476. doi:10.1038/s41598-020-64493-4
- Lor, G. C. Y., Risch, H. A., Fung, W. T., Au Yeung, S. L., Wong, I. O. L., Zheng, W., et al. (2019). Reporting and Guidelines for Mendelian Randomization Analysis: A Systematic Review of Oncological Studies. *Cancer Epidemiol.* 62, 101577. doi:10.1016/j.canep.2019.101577
- Lyall, D. M., Celis-Morales, C., Ward, J., Iliodromiti, S., Anderson, J. J., Gill, J. M. R., et al. (2017). Association of Body Mass Index with Cardiometabolic Disease in the UK Biobank. *JAMA Cardiol.* 2, 882–889. doi:10.1001/jamacardio.2016.5804
- Meng, X., Li, X., Timofeeva, M. N., He, Y., Spiliopoulou, A., Wei, W.-Q., et al. (2019). Phenome-wide Mendelian-Randomization Study of Genetically Determined Vitamin D on Multiple Health Outcomes Using the UK Biobank Study. *Int. J. Epidemiol.* 48, 1425–1434. doi:10.1093/ije/dyz182
- Minelli, C., Del Greco M., F., van der Plaats, D. A., Bowden, J., Sheehan, N. A., and Thompson, J. (2021). The Use of Two-Sample Methods for Mendelian Randomization Analyses on Single Large Datasets. *Int. J. Epidemiol.* 50, 1651–1659. doi:10.1093/ije/dyab084

- Mokry, L. E., Ross, S., Timpson, N. J., Sawcer, S., Davey Smith, G., and Richards, J. B. (2016). Obesity and Multiple Sclerosis: A Mendelian Randomization Study. *Plos Med.* 13, e1002053. doi:10.1371/journal.pmed.1002053
- Morrison, J., Knoblauch, N., Marcus, J. H., Stephens, M., and He, X. (2020). Mendelian Randomization Accounting for Correlated and Uncorrelated Pleiotropic Effects Using Genome-wide Summary Statistics. *Nat. Genet.* 52, 740–747. doi:10.1038/s41588-020-0631-4
- Noyce, A. J., Kia, D. A., Hemani, G., Nicolas, A., Price, T. R., De Pablo-Fernandez, E., et al. (2017). Estimating the Causal Influence of Body Mass index on Risk of Parkinson Disease: A Mendelian Randomisation Study. *Plos Med.* 14, e1002314. doi:10.1371/journal.pmed.1002314
- Palmer, T. M., Lawlor, D. A., Harbord, R. M., Sheehan, N. A., Tobias, J. H., Timpson, N. J., et al. (2012). Using Multiple Genetic Variants as Instrumental Variables for Modifiable Risk Factors. *Stat. Methods Med. Res.* 21, 223–242. doi:10.1177/0962280210394459
- Pausova, Z., Syme, C., Abrahamowicz, M., Xiao, Y., Leonard, G. T., Perron, M., et al. (2009). A Common Variant of the FTO Gene Is Associated with Not Only Increased Adiposity but Also Elevated Blood Pressure in French Canadians. *Circ. Cardiovasc. Genet.* 2, 260–269. doi:10.1161/CIRCGENETICS.109.857359
- Pingault, J.-B., Cecil, C. A. M., Murray, J., Munafò, M. R., and Viding, E. (2016). Causal Inference in Psychopathology: A Systematic Review of Mendelian Randomisation Studies Aiming to Identify Environmental Risk Factors for Psychopathology. *Psychopathology Rev.* a4, 4–25. doi:10.5127/pr.038115
- Qi, G., and Chatterjee, N. (2019). Mendelian Randomization Analysis Using Mixture Models for Robust and Efficient Estimation of Causal Effects. *Nat. Commun.* 10, 1941. doi:10.1038/s41467-019-09432-2
- Reactive Protein Coronary Heart Disease Genetics Collaboration (Ccg), C., Wensley, F., Gao, P., Burgess, S., Kaptoge, S., Di Angelantonio, E., et al. (2011). Association between C Reactive Protein and Coronary Heart Disease: Mendelian Randomisation Analysis Based on Individual Participant Data. *BMJ* 342, d548. doi:10.1136/bmj.d548

- Rees, J. M. B., Wood, A. M., Dudbridge, F., and Burgess, S. (2019). Robust Methods in Mendelian Randomization via Penalization of Heterogeneous Causal Estimates. *PLoS One* 14, e0222362. doi:10.1371/journal.pone.0222362
- Sainani, K. L. (2018). Instrumental Variables: Uses and Limitations. *PM&R* 10, 303–308. doi:10.1016/j.pmrj.2018.02.002
- Sekula, P., Del Greco M, F., Pattaro, C., and Köttgen, A. (2016). Mendelian Randomization as an Approach to Assess Causality Using Observational Data. *Jasn* 27, 3253–3265. doi:10.1681/ASN.2016010098
- Skaaby, T., Taylor, A. E., Thuesen, B. H., Jacobsen, R. K., Friedrich, N., Møllehave, L. T., et al. (2018). Estimating the Causal Effect of Body Mass index on hay Fever, Asthma and Lung Function Using Mendelian Randomization. *Allergy* 73, 153–164. doi:10.1111/all.13242
- Slob, E. A., Groenen, P. J., Thurik, A. R., and Rietveld, C. A. (2017). A Note on the Use of Egger Regression in Mendelian Randomization Studies. *Int. J. Epidemiol.* 46, 2094–2097. doi:10.1093/ije/dyx191
- Slob, E. A. W., and Burgess, S. (2020). A Comparison of Robust Mendelian Randomization Methods Using Summary Data. *Genet. Epidemiol.* 44, 313–329. doi:10.1002/gepi.22295
- Sun, L., Ma, L., Zhang, H., Cao, Y., Wang, C., Hou, N., et al. (2019). Fto Deficiency Reduces Anxiety- and Depression-like Behaviors in Mice via Alterations in Gut Microbiota. *Theranostics* 9, 721–733. doi:10.7150/thno.31562
- Sun, Y.-Q., Brumpton, B. M., Langhammer, A., Chen, Y., Kvaløy, K., and Mai, X.-M. (2020). Adiposity and Asthma in Adults: a Bidirectional Mendelian Randomisation Analysis of the HUNT Study. *Thorax* 75, 202–208. doi:10.1136/thoraxjnl-2019-213678
- Swerdlow, D. I., Kuchenbaecker, K. B., Shah, S., Sofat, R., Holmes, M. V., White, J., et al. (2016). Selecting Instruments for Mendelian Randomization in the Wake of Genome-wide Association Studies. *Int. J. Epidemiol.* 45, 1600–1616. doi:10.1093/ije/dyw088
- Taylor, A. E., Richmond, R. C., Palviainen, T., Loukola, A., Wootton, R. E., Kaprio, J., et al. (2019). The Effect of Body Mass index on Smoking Behaviour and Nicotine Metabolism: a Mendelian Randomization Study. *Hum. Mol. Genet.* 28, 1322–1330. doi:10.1093/hmg/ddy434

- Thompson, J. R., Minelli, C., Bowden, J., Del Greco, F. M., Gill, D., Jones, E. M., et al. (2017). Mendelian Randomization Incorporating Uncertainty about Pleiotropy. *Stat. Med.* 36, 4627–4645. doi:10.1002/sim.7442
- Tyrrell, J., Jones, S. E., Beaumont, R., Astley, C. M., Lovell, R., Yaghootkar, H., et al. (2016). Height, Body Mass index, and Socioeconomic Status: Mendelian Randomisation Study in UK Biobank. *Bmj* 352, i582. doi:10.1136/bmj.i582
- Tyrrell, J., Mulugeta, A., Wood, A. R., Zhou, A., Beaumont, R. N., Tuke, M. A., et al. (2019). Using Genetics to Understand the Causal Influence of Higher BMI on Depression. *Int. J. Epidemiol.* 48, 834–848. doi:10.1093/ije/dyy223
- van den Broek, N., Treur, J. L., Larsen, J. K., Verhagen, M., Verweij, K. J. H., and Vink, J. M. (2018). Causal Associations between Body Mass index and Mental Health: a Mendelian Randomisation Study. *J. Epidemiol. Community Health* 72, 708–710. doi:10.1136/jech-2017-210000
- VanderWeele, T. J., Tchetgen Tchetgen, E. J., Cornelis, M., and Kraft, P. (2014). Methodological Challenges in Mendelian Randomization. *Epidemiology* 25, 427–435. doi:10.1097/ede.0000000000000081
- Vansteelandt, S., Bowden, J., Babanezhad, M., and Goetghebeur, E. (2011). On Instrumental Variables Estimation of Causal Odds Ratios. *Statist. Sci.* 26, 403–422. doi:10.1214/11-STS360
- Verbanck, M., Chen, C.-Y., Neale, B., and Do, R. (2018). Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred from Mendelian Randomization between Complex Traits and Diseases. *Nat. Genet.* 50, 693–698. doi:10.1038/s41588-018-0099-7
- Walter, S., Glymour, M. M., Koenen, K., Liang, L., Tchetgen Tchetgen, E. J., Cornelis, M., et al. (2015a). Do genetic Risk Scores for Body Mass index Predict Risk of Phobic Anxiety? Evidence for a Shared Genetic Risk Factor. *Psychol. Med.* 45, 181–191. doi:10.1017/S0033291714001226
- Walter, S., Kubzansky, L. D., Koenen, K. C., Liang, L., Tchetgen Tchetgen, E. J., Cornelis, M. C., et al. (2015b). Revisiting Mendelian Randomization Studies of the Effect of Body Mass index on Depression. *Am. J. Med. Genet.* 168, 108–115. doi:10.1002/ajmg.b.32286

- Wang, N., Cheng, J., Ning, Z., Chen, Y., Han, B., Li, Q., et al. (2018). Type 2 Diabetes and Adiposity Induce Different Lipid Profile Disorders: A Mendelian Randomization Analysis. *J. Clin. Endocrinol. Metab.* 103, 2016–2025. doi:10.1210/jc.2017-02789
- Wang, T., Zhang, R., Ma, X., Wang, S., He, Z., Huang, Y., et al. (2018). Causal Association of Overall Obesity and Abdominal Obesity with Type 2 Diabetes: A Mendelian Randomization Analysis. *Obesity* 26, 934–942. doi:10.1002/oby.22167
- Winter-Jensen, M., Afzal, S., Jess, T., Nordestgaard, B. G., and Allin, K. H. (2020). Body Mass index and Risk of Infections: a Mendelian Randomization Study of 101,447 Individuals. *Eur. J. Epidemiol.* 35, 347–354. doi:10.1007/s10654-020-00630-7
- Xu, L., Borges, M. C., Hemani, G., and Lawlor, D. A. (2017). The Role of Glycaemic and Lipid Risk Factors in Mediating the Effect of BMI on Coronary Heart Disease: a Two-step, Two-Sample Mendelian Randomisation Study. *Diabetologia* 60, 2210–2220. doi:10.1007/s00125-017-4396-y
- Xu, S., Gilliland, F. D., and Conti, D. V. (2019). Elucidation of Causal Direction between Asthma and Obesity: a Bi-directional Mendelian Randomization Study. *Int. J. Epidemiol.* 48, 899–907. doi:10.1093/ije/dyz070
- Zhang, L., Tang, L., Huang, T., and Fan, D. (2020). Life Course Adiposity and Amyotrophic Lateral Sclerosis: A Mendelian Randomization Study. *Ann. Neurol.* 87, 434–441. doi:10.1002/ana.25671
- Zhao, Q., Wang, J., Hemani, G., Bowden, J., and Small, D. S. (2020a). Statistical Inference in Two-Sample Summary-Data Mendelian Randomization Using Robust Adjusted Profile Score. *Ann. Statist.* 48, 1742–1769. doi:10.1214/19-AOS1866
- Zhao, Y., Xu, Y., Wang, X., Xu, L., Chen, J., Gao, C., et al. (2020b). Body Mass Index and Polycystic Ovary Syndrome: A 2-Sample Bidirectional Mendelian Randomization Study. *J. Clin. Endocrinol. Metab.* 105, 1778–1784. doi:10.1210/clinem/dgaa125

Table 1*Characteristics of included studies.*

Characteristics of studies	<i>n</i>	%
Types of data used	<i>n</i> = 128	
One-sample only	82	64.1
Two-sample only	30	23.4
Both one and two samples	16	12.5
Types of instruments used in the main analysis	<i>n</i> = 128	
GRS only	63	49.2
Multiple IVs only	40	31.2
Single IVs only	13	10.2
Both GRS and multiple IVs	6	4.7
Both GRS and single IVs	5	3.9
Both single and multiple IVs	1	0.8
Types of GRS used among studies using GRS as IV in the main analysis	<i>n</i> = 74	
Weighted GRS only	48	64.9
Unweighted GRS only	19	25.7
Both weighted and unweighted GRS	7	9.4

Abbreviations: GRS, genetic risk score; IVs, instrumental variables; SNPs, single-nucleotide polymorphisms; *FTO*, fat mass and obesity-associated.

Table 2

Approaches cited for preventing pleiotropy in the selection of IVs.

Approaches	n/33	%
Exclude selected SNPs known or suspected to be pleiotropic	12	36.4
Use of multiple (independent) SNPs combined in a GRS as IVs	10	30.3
Use of multiple (independent) SNPs as IVs	10 ^a	30.3
Use a single SNP as IV	1	3.0

Abbreviations: GRS, genetic risk score; SNP, single-nucleotide polymorphisms; IV, instrumental variable.

^aThese 10 articles discussed the use of multiple independent SNPs without explicitly mentioning GRS even if six of the 10 in fact used a GRS as IV in the main analysis.

Table 3

Methods cited for detection or accounting the violation of the independence and/or restriction exclusion assumptions.

Methods^a	<i>n</i>/128	Studies specifically referred to the method for “pleiotropy” assessment / control / robustness (n=95)	Studies not specifically referred to the method for “pleiotropy” assessment / control / robustness (n=96)
Robust and other IV estimation methods^b	70	69	23
Robust methods	69	68	20
MR-Egger	69 ^c	68	1
Median-based methods	44	26	18
Mode-based methods	9	5	4
MR-PRESSO	6	5	1
MR-RAPS	1	0	1
Other IV estimation methods	13	6	8
Multivariable MR	6	3	3
IVW methods	4	3	1
Likelihood-based methods	4	0	4
Methods to detect heterogeneity of estimated causal effects across IVs	72	46	46
Graphical assessment: scatter plots, forest plots, funnel plots, leave-one-out plots, and histogram	34	20	19
Statistical criteria and tests: I^2 , r^2 , H statistic, Cochran's Q, Rucker's Q, over-identification tests ^d	30	7	23
Comparisons of estimated MR causal effects across IVs (GRS or multiple IVs)	39	25	14
Before and after exclusion of SNPs suspected of pleiotropy	33	20	13
GRS or multiple IVs vs. single SNP(s)	5	5	0
Two subsets of SNPs grouping SNPs with the same biological pathway on the exposure	1	0	1
Detection of outlier/influential SNP(s): Cook's distance, Studentized residuals, <i>HEIDI-outlier</i> , leave-one-out analyses	15	4	13
Methods to detect associations between IVs and the outcome outside of the pathway through the exposure	85	30	61
Estimating the associations between the IVs and measured risk factors for the outcome	75	19	56
Documenting the associations between the IV and risk factors for the outcome in the literature	7	5	2
Adjusting IV-outcome or IV-confounders associations for exposure	18	7	11

Adjusting IV-outcome association or MR analyses for covariates potentially involved in pleiotropic pathways ^e	5	2	3
Comparison of the exposure-outcome conventional vs. IV estimated effects	4	1	3
Comparison of the IV-outcome vs. IV-exposure associations	2	2	0
Estimating the association between the IV and the outcome	2	1	1
Mediation analysis estimating the direct effect of the IV on the outcome	3	0	3
Other methods	5	0	5
The use of positive or negative control outcomes	3	0	3
Colocalization	1	0	1
Verifying the concordance of MR results with those from other studies (MR, clinical trial)	1	0	1
No method reported for assessment of independence or exclusion restriction assumptions	8		

Abbreviations: IV, instrumental variable; MR, Mendelian randomization; PRESSO, pleiotropy residual sum and outlier; SNP, single-nucleotide polymorphism; RAPS, Robust Adjusted Prole Score; IVW, inverse-variance weighted; *FTO*, fat mass and obesity-associated; GRS, genetic risk score; HEIDI, Heterogeneity in Dependent Instruments.

^a Many studies cited more than one method and thus the number of methods reported in the Table exceed the total number of studies (n = 128).

^b See (Burgess et al., 2020a) and (Burgess et al., 2015) for a summary of the listed methods.

^c Of the 69 studies that reported using MR-Egger, 66 used the intercept test p-value to infer whether or not pleiotropy was present, two studies (Tyrrell et al., 2016; Fan et al., 2018) compared the MR-Egger slope and the conventional MR causal effect estimate, while the last study (Chen et al., 2019) did not specify how the MR-Egger results were used.

^d One study (Censin et al., 2017) did not specify the heterogeneity test used.

^e Three articles (Guo et al., 2016; Censin et al., 2019; Sun et al., 2020) mentioned adjustment of MR analyses for covariates potentially involved in pleiotropic pathways without specifying as multivariable MR.

Table 4*Discussion of the independence and/or exclusion restriction assumptions.*

Discussion of the independence and/or exclusion restriction assumptions	<i>n</i>	%
Plausibility of the independence and/or exclusion restriction assumptions	<i>n</i> = 128	
Discussed with specific reference to pleiotropy	89	69.5
Discussed without specific reference to pleiotropy	19	14.9
IV invalidity ^a	<i>n</i> = 128	
Suspected with specific reference to pleiotropy	16	12.5
Suspected without specific reference to pleiotropy	5	3.9
Impact of IV invalidity on the validity of MR results	<i>n</i> = 21	
May have affected validity of results	8	38.1
No or low impact on validity of results	9	42.9
Impact not (clearly) reported	4	19.0
Suspicion of invalidity/pleiotropy of <i>FTO</i> SNP(s)	<i>n</i> = 128	
Yes	6	4.7
<i>FTO</i> SNP(s) suspected to be invalid/pleiotropic	<i>n</i> = 6	
rs1558902 ^b	3	50.0
rs1421085 ^c	2	33.3
rs17817449 ^d	1	16.7

Abbreviations: IV, instrumental variable; MR, Mendelian randomization; SNP, single-nucleotide polymorphism; *FTO*, fat mass and obesity-associated.

^a Refers to the suspicion of invalidity of one or more body mass index IV(s) for any outcome of interest.

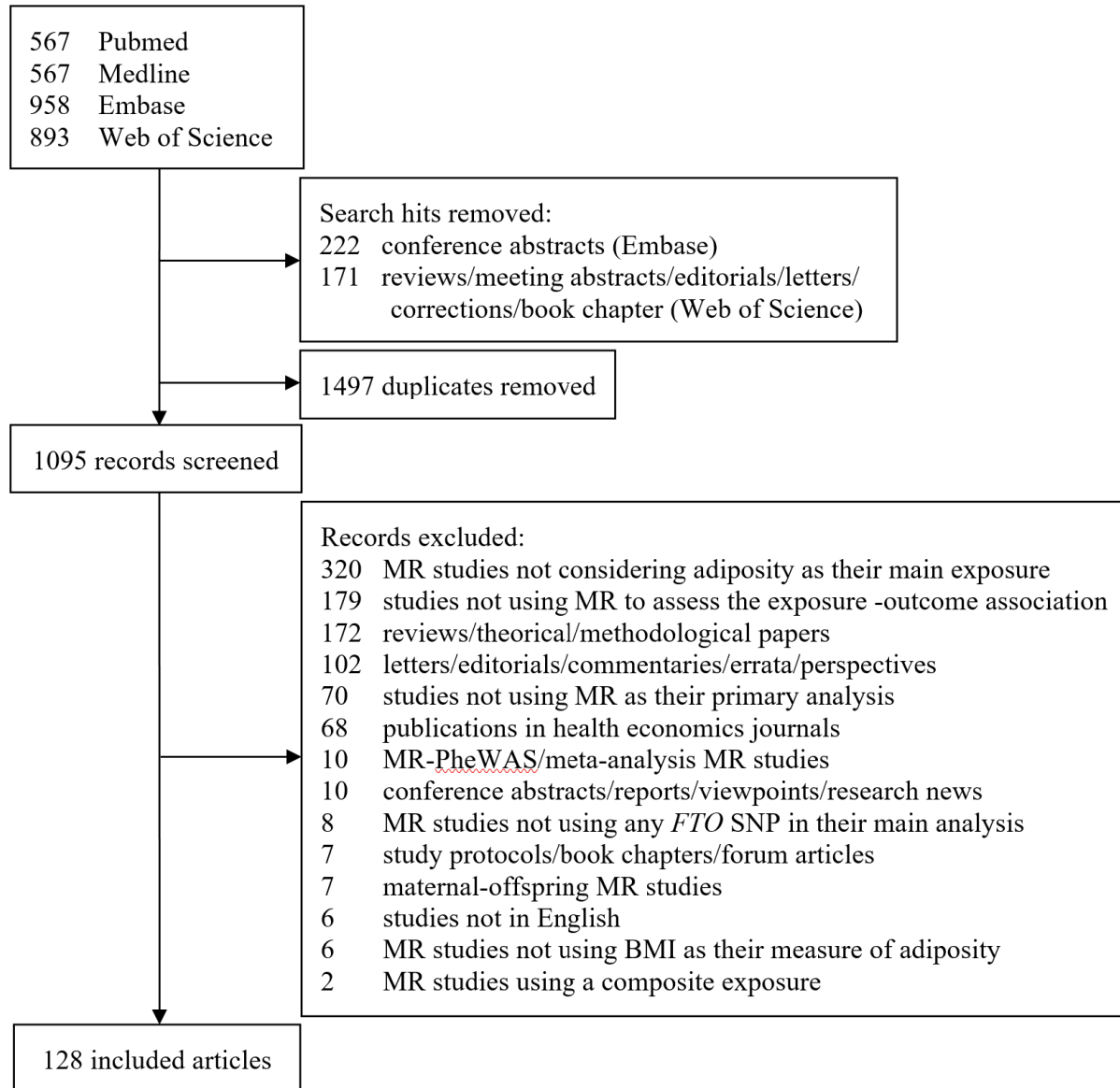
^b The outcomes of interest involved in the suspected invalidity of rs1558902 were multiple sclerosis susceptibility (Gianfrancesco et al., 2017), phobic anxiety symptoms (Walter et al., 2015a), and depression (Walter et al., 2015b).

^c The outcomes of interest involved in the suspected pleiotropy of rs1421085 were common mental disorders (Kivimäki et al., 2011), and subjective well-being (van den Broek et al., 2018).

^d The outcome of interest involved in the suspected pleiotropy of rs17817449 was lipid profiles (Wang N. et al., 2018).

Figure 1

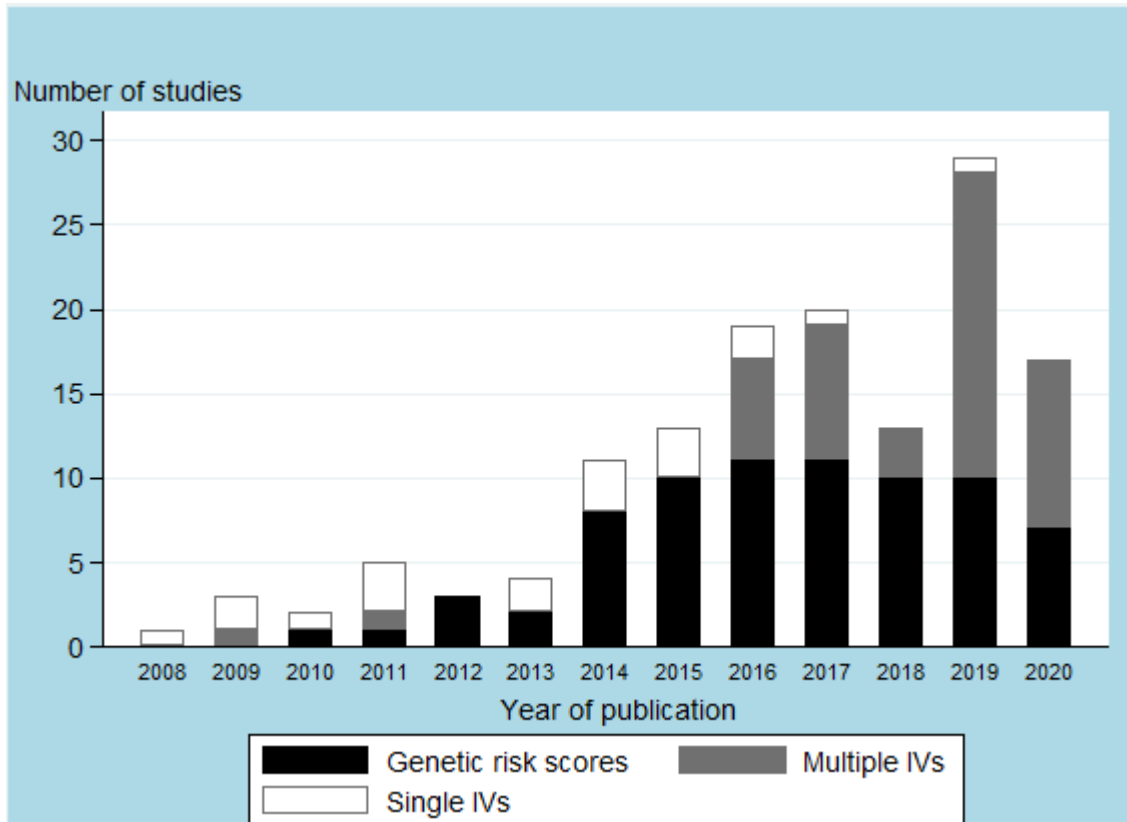
Flowchart of selection of studies includes in the review.



Abbreviations: MR, mendelian randomization; PheWAS, phenome-wide association study; *FTO*, fat mass and obesity-associated; SNP, *single-nucleotide polymorphism*; BMI, body mass index.

Figure 2

Body mass index instrumental variables (IVs) used in the main analysis by article publication year.



Supplementary Material

1. Supplementary Methods

Methods S1. Search strategy

For each of the four databases listed below, the search strategy was operationalized using the keywords "Mendelian randomization", "adiposity" and their related or equivalent terms. Mendelian randomization related or equivalent terms used in this review were borrowed from the literature (Boef et al., 2015).

PubMed

Query:

("Mendelian Randomization Analysis"[Mesh] OR "Mendelian randomisation"[all fields] OR "Mendelian randomization"[all fields] OR (Mendelian[all fields] AND randomi*[all fields]) OR "genetic instrumental variable"[all fields] OR "genetic instrumental variables"[all fields] OR "genetic instrument"[all fields] OR "genetic instruments"[all fields] OR "genes as instruments"[all fields] OR "gene as instrument"[all fields] OR "genes as instrument"[all fields] OR "gene as instruments"[all fields] OR (instrument*[ti] AND (gene[ti] OR genes[ti] OR genetic*[ti] OR mendel*[ti])) OR (("instrumental variable"[all fields] OR "instrumental variables"[all fields] OR "instrumented analysis"[all fields] OR "instrumented analyses"[all fields] OR "instrumental variable analysis"[all fields] OR "instrumental variable analyses"[all fields] OR "instrumental variables analysis"[all fields] OR "instrumental variables analyses"[all fields]) AND (gene OR genes OR genetics OR mendel OR mendelian)) OR ("mendelian"[all fields] AND ("randomisation"[all fields] OR "randomization"[all fields] OR "randomising"[all fields] OR "randomizing"[all fields]))) AND ("adiposity"[Mesh] OR "adiposity"[all fields] OR "obesity"[all fields] OR "body mass index"[Mesh] OR "body mass index"[all fields] OR "BMI"[all fields] OR "fat mass"[all fields] OR "body fat"[all fields] OR "body fat percentage"[all fields] OR "obesity, abdominal"[Mesh] OR "waist-hip ratio"[Mesh] OR "waist-hip ratio"[all fields] OR "waist circumference"[Mesh] OR "waist circumference"[all fields])

Medline

Query:

(Mendelian Randomization Analysis/ OR "Mendelian randomisation".af OR "Mendelian randomization".af OR "genetic instrumental variable".af OR "genetic instrumental variables".af OR "genetic instrument".af OR "genetic instruments".af OR "mendel randomise" OR "mendel randomize" OR "mendel randomization" OR "mendel randomisation" OR "random Mendelian" OR "genes as instruments".af OR "gene as instrument".af OR "genes as instrument".af OR "gene as instruments".af) OR (instrument*.ti AND (gene.ti OR genes.ti OR genetic*.ti OR mendel*.ti)) OR (("instrumental variable".af OR "instrumental variables".af OR "instrumented analysis".af OR "instrumented analyses".af OR "instrumental variable analysis".af OR "instrumental variable analyses".af OR "instrumental variables analysis".af OR "instrumental variables analyses".af) AND (gene OR genes OR genetics OR mendel OR mendelian).af) OR ("mendelian".af AND ("randomisation".af OR "randomization".af OR "randomising".af OR "randomizing".af)) AND (adiposity/ OR "adiposity".af OR "obesity".af OR body mass index/ OR "body mass index".af OR "BMI".af OR "fat mass".af OR "body fat".af OR "body fat percentage".af OR obesity, abdominal/ OR waist-hip ratio/ OR "waist-hip ratio".af OR waist circumference/ OR "waist circumference".af)

Embase

Restriction: no conference abstracts.

Query:

(Mendelian Randomization Analysis/ OR "Mendelian randomisation".af OR "Mendelian randomization".af OR "genetic instrumental variable".af OR "genetic instrumental variables".af OR "genetic instrument".af OR "genetic instruments".af OR "mendel randomise" OR "mendel randomize" OR "mendel randomization" OR "mendel randomisation" OR "random Mendelian" OR "genes as instruments".af OR "gene as instrument".af OR "genes as instrument".af OR "gene as instruments".af) OR (instrument*.ti AND (gene.ti OR genes.ti OR genetic*.ti OR mendel*.ti)) OR (("instrumental variable".af OR "instrumental variables".af OR "instrumented analysis".af OR "instrumented analyses".af OR "instrumental variable analysis".af OR "instrumental variable analyses".af OR "instrumental variables analysis".af OR "instrumental variables analyses".af) AND (gene OR genes OR genetics OR mendel OR mendelian).af) OR ("mendelian".af AND ("randomisation".af OR "randomization".af OR "randomising".af OR "randomizing".af)) AND (adiposity/ OR "adiposity".af OR "obesity".af OR body mass index/ OR "body mass index".af OR "BMI".af OR "fat mass".af OR "body fat".af OR "body fat percentage".af OR obesity, abdominal/ OR waist-hip ratio/ OR "waist-hip ratio".af OR waist circumference/ OR "waist circumference".af)

Web of Science

Restrictions: no review, meeting abstract, editorial material, correction, letter, or book chapter.

Query:

ALL FIELDS: (("Mendelian randomisation" OR "Mendelian randomization" OR "genetic instrumental variable" OR "genetic instrumental variables" OR "genetic instrument" OR "genetic instruments" OR "mendel randomise" OR "mendel randomize" OR "mendel randomization" OR "mendel randomisation" OR "random Mendelian" OR "gene as instrument" OR "gene as instruments" OR "instrumental genetic variable" OR "instrumental genetic variables" OR "instrumental variable" OR "instrumental variable analysis") AND ("adiposity" OR "obesity" OR "body mass index" OR "BMI" OR "fat mass" OR "body fat" OR "body fat percentage" OR "abdominal obesity" OR "waist-hip ratio" OR "waist circumference"))

Methods S2. Specific exclusion criteria

Specific exclusion criteria included studies not in English, conference abstracts, letters, reports, commentaries, corrections, viewpoints, research news, perspectives, study protocols, errata, forum articles, editorials, book chapters, reviews, meta-analysis mendelian randomization (MR) studies, publications in health economics journals, theoretical or methodological papers with MR analyses examples, MR phenome-wide association studies (MR-PheWAS), MR studies with a composite exposure combining body mass index and other risk factor(s), and those involving mother-child pairs (in which the instrument and exposure are assessed in the mother and outcome in the child). In bidirectional MR studies, we only included the analysis that treated body mass index as the exposure.

2. Supplementary Tables

Table S1

List of the 128 studies included.

Studies	Outcomes
(Afzal et al., 2014)	Type 2 diabetes
(Bae and Lee, 2019)	Rheumatoid arthritis
(Barning and Abarin, 2016)	Physical activity
(Benn et al., 2016)	Any non-skin, non melanoma skin, lung, other smoking related (oral, larynx and bladder), colon, kidney, breast, prostate, and other cancers
(Bonfond et al., 2017) ^a	Plasma enzymatic activity of salivary or pancreatic amylase
(Brennan et al., 2009)	Lung, aero-digestive, and kidney cancers
(Brower et al., 2019) ^a	Polycystic ovary syndrome
(Budu-Aggrey et al., 2019) ^a	Psoriasis
(Busch et al., 2019)	Timing of puberty
(Carreras-Torres et al., 2016)	Different histological types of lung cancer
(Carreras-Torres et al., 2018)	Smoking behaviour
(Censin et al., 2017)	Type 1 diabetes
(Censin et al., 2019)	Several non-communicable diseases and cardiometabolic risk factors
(Chatterjee et al., 2017)	Atrial fibrillation
(Chen, Fan, Huang, et al., 2019) ^a	Asthma
(Chen, Fan, Yang, et al., 2019) ^a	Early puberty
(Colak et al., 2016)	Wheezing and asthma
(Cole et al., 2016)	Coronary artery disease
(Corbin et al., 2016)	Type 2 diabetes
(Dale et al., 2017)	Coronary artery disease, type 2 diabetes mellitus, and major stroke subtypes (primary outcomes), and 18 cardiometabolic traits (secondary outcomes)
(Davies et al., 2015)	Prostate cancer risk and mortality
(den Hollander et al., 2017) ^a	Helicobacter pylori colonization
(Dixon et al., 2016)	Subtypes of ovarian cancer
(Dusingize et al., 2020)	Melanoma
(Eriksson et al., 2017) ^a	Serum testosterone
(Fall et al., 2013)	24 cardiometabolic phenotypes
(Fall et al., 2015)	14 cardiovascular risk factors
(Fan et al., 2018)	Early age at menarche
(Fussey et al., 2020)	Benign nodular thyroid disease and differentiated thyroid cancer
(Gao et al., 2016)	Breast, ovarian, prostate, colorectal, and lung cancers

(Geng et al., 2018)	Type 2 diabetes, coronary artery disease, and chronic kidney disease (primary outcomes); several cardiometabolic traits (secondary outcomes)
(Gharahkhani et al., 2019)	Overall cancer risk and mortality, and specific type cancer risk
(Gianfrancesco, Glymour, et al., 2017)	Multiple sclerosis susceptibility
(Gianfrancesco, Stridh, et al., 2017)	Pediatric-onset multiple sclerosis
(Granell et al., 2014)	Current asthma and asthma subtypes
(Guo et al., 2017)	Breast cancer survival
(Guo et al., 2016)	Breast cancer
(Hagg et al., 2015)	Coronary heart disease, heart failure, and ischaemic stroke
(Hartwig et al., 2016)	Bipolar and major depressive disorders, and schizophrenia
(Holmes et al., 2014)	Several cardiometabolic traits and events
(Howe et al., 2019)	Six measures of socio-economic position and three measures of social contact
(Howe et al., 2017)	Body dissatisfaction and smoking behaviour
(Huang et al., 2016)	Peripheral arterial disease
(Hung et al., 2014)	Major depression
(Jacobs et al., 2020) ^a	Multiple sclerosis
(Jarvis et al., 2016)	Colorectal cancer
(Jokela et al., 2012)	Depressive symptoms
(Kaltoft et al., 2020)	Aortic valve stenosis and replacement
(Kemp et al., 2016) ^a	Bone mineral density
(Kivimäki et al., 2011)	Common mental disorders
(Kivimaki et al., 2008)	Adult carotid intima-media thickness and various atherosclerotic risk factors
(Klovaite et al., 2015)	First events of deep venous thrombosis with or without pulmonary embolism
(Kurz and Laxy, 2020)	Total annual health care costs
(Larsson et al., 2020)	14 cardiovascular conditions
(Larsson et al., 2018)	Gout risk and serum urate concentrations
(Lawlor et al., 2011)	Psychological distress
(Lee et al., 2018)	Hypertension
(Lee et al., 2019)	Bone health measures
(Lewis et al., 2010)	Prostate cancer
(Lim et al., 2009)	Cataract outcomes
(Lindstrom et al., 2017)	Venous thromboembolism
(Liu et al., 2020)	Diabetes
(Lyll et al., 2017)	Stroke, coronary heart disease, hypertension, type 2 diabetes, systolic and diastolic blood pressure, and pulse rate.
(Lyngdoh et al., 2012) ^a	Serum uric acid
(Mao et al., 2017)	Gastric cancer
(Marini et al., 2020) ^a	Multiple cerebrovascular disease phenotypes

(Martins-Silva et al., 2019) ^a	Attention-deficit hyperactivity disorder
(Mokry et al., 2016)	Multiple sclerosis
(Mukherjee et al., 2015)	Alzheimer's disease and dementia
(Mumby et al., 2011)	Early menarche
(Nordestgaard et al., 2012)	Ischemic heart disease
(Nordestgaard et al., 2017)	Alzheimer's disease
(Noyce et al., 2017)	Parkinson disease
(Ooi et al., 2019) ^a	Breast size and breast cancer risk
(Painter et al., 2016)	Endometrial cancer
(Panoutsopoulou et al., 2014)	Knee and hip osteoarthritis
(Qian, Rookus, et al., 2019)	Ovarian cancer risk
(Qian, Wang, et al., 2019)	Breast cancer risk
(Reed et al., 2017) ^a	Disordered eating patterns
(Reed et al., 2020) ^a	DNA methylation
(Richardson et al., 2020)	Coronary artery disease, type 2 diabetes, breast cancer, and prostate cancer
(Richmond et al., 2014) ^a	Physical activity level
(Robinson et al., 2020)	Postoperative complications
(Rode et al., 2014)	Short telomere length
(Schnurr et al., 2018)	Sedentary time and physical activity
(Shu et al., 2019)	Breast cancer risk
(Shungin et al., 2015)	Periodontitis
(Skaaby et al., 2018)	Asthma, hay fever, allergic sensitization, serum total immunoglobulin E, forced expiratory volume in one-second and forced vital capacity
(Speed et al., 2019) ^a	Depression
(Stender et al., 2013)	Symptomatic gallstone disease
(Sun et al., 2020) ^a	Asthma
(Sun et al., 2019)	All cause and cause specific (cardiovascular, cancer, and non-cardiovascular non-cancer) and incident diseases (cardiovascular and cancer)
(Takahashi et al., 2019)	Meningioma
(Tan et al., 2019)	Cataract
(Taylor et al., 2019) ^a	Smoking behaviours, DNA methylation, cotinine levels, and nicotine metabolite ratio
(Taylor et al., 2016)	Free T3 and T4 levels
(Thakkinstian et al., 2014) ^a	Circulating fetuin-A level
(Thrift et al., 2015)	Colorectal cancer
(Thrift et al., 2014)	Esophageal adenocarcinoma and Barrett's esophagus
(Timpson et al., 2009)	Systolic and diastolic blood pressure
(Timpson et al., 2011) ^a	Circulating C-reactive protein

(Todd et al., 2015)	Diabetic kidney disease
(Tyrrell et al., 2016)	Age completed full time education, degree level education, job class, annual household income, and Townsend deprivation index.
(Tyrrell et al., 2019)	Depression
(van den Broek et al., 2018) ^a	Subjective well-being and depressive symptoms
(van 't Hof et al., 2017)	Intracranial and abdominal aortic aneurysms
(Varbo et al., 2015)	Ischemic heart disease
(Vasan et al., 2019)	Circulating plasma succinate and 12,13-dihydroxy-9Z-octadecenoic acid
(Vimaleswaran et al., 2013) ^a	25-hydroxyvitamin D
(Wade, Carslake, et al., 2018)	All-cause and cause-specific mortality
(Wade, Chiesa, et al., 2018)	Blood pressure, heart rate, carotid femoral pulse wave velocity, carotid intima-media thickness, and left ventricular mass index
(Wainberg et al., 2019)	Type 2 diabetes
(Walter, Glymour, et al., 2015)	Phobic anxiety
(Walter, Kubzansky, et al., 2015)	Depression
(Wang, Cheng, et al., 2018)	Total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, and triglycerides
(Wang, Lu, et al., 2018)	Blood lead level
(Wang et al., 2016)	Insulin secretion and sensitivity
(Wang, Zhang, et al., 2018)	Type 2 diabetes development, insulin secretion, and insulin resistant traits
(Warodomwichit et al., 2013)	Bone mineral density
(Welsh et al., 2010) ^a	Circulating levels of C-reactive protein and leptin
(Winter-Jensen et al., 2020)	Hospital contacts related to infections
(Wurtz et al., 2014)	82 metabolic measures
(Xu et al., 2017) ^a	Coronary heart disease
(Xu et al., 2019) ^a	Asthma
(Zeng et al., 2019) ^a	Amyotrophic lateral sclerosis
(Zhang et al., 2020)	Amyotrophic lateral sclerosis
(Zhao, Xu, et al., 2020) ^a	Polycystic ovary syndrome
(Zhou et al., 2019)	Alzheimer's disease

^a Bidirectional mendelian randomization study. We only analysed the analysis that treated body mass index as exposure.

Table S2

Frequency of use of 31 FTO single-nucleotide polymorphisms as instruments for body mass index.

FTO SNPs	n/128	References
rs1558902	64	(Bae and Lee, 2019; Bonnefond et al., 2017; Budu-Aggrey et al., 2019; Busch et al., 2019; Carreras-Torres et al., 2016; Chatterjee et al., 2017; Cole et al., 2016; Corbin et al., 2016; Dale et al., 2017; Davies et al., 2015; Dixon et al., 2016; Dusingize et al., 2020; Eriksson et al., 2017; Fall et al., 2015; Fussey et al., 2020; Gao et al., 2016; Gianfrancesco, Glymour, et al., 2017; Gianfrancesco, Stridh, et al., 2017; Granell et al., 2014; Guo et al., 2017; Guo et al., 2016; Hartwig et al., 2016; Howe et al., 2019; Howe et al., 2017; Jarvis et al., 2016; Jokela et al., 2012; Kurz and Laxy, 2020; Larsson et al., 2020; Larsson et al., 2018; Lindstrom et al., 2017; Lyall et al., 2017; Mao et al., 2017; Mokry et al., 2016; Mukherjee et al., 2015; Noyce et al., 2017; Ooi et al., 2019; Painter et al., 2016; Qian, Rookus, et al., 2019; Qian, Wang, et al., 2019; Reed et al., 2017; Reed et al., 2020; Richmond et al., 2014; Robinson et al., 2020; Shu et al., 2019; Sun et al., 2020; Sun et al., 2019; Taylor et al., 2019; Taylor et al., 2016; Thrift et al., 2015; Todd et al., 2015; Tyrrell et al., 2016; Tyrrell et al., 2019; van 't Hof et al., 2017; Wade, Carslake, et al., 2018; Wade, Chiesa, et al., 2018; Wainberg et al., 2019; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015; Wang et al., 2016; Wang, Zhang, et al., 2018; Wurtz et al., 2014; Xu et al., 2017; Xu et al., 2019; Zhang et al., 2020)
rs9939609	28	(Afzal et al., 2014; Barning and Abarin, 2016; Benn et al., 2016; Brennan et al., 2009; Chen, Fan, Yang, et al., 2019; Colak et al., 2016; Fall et al., 2013; Kalsoft et al., 2020; Kivimäki et al., 2008; Klovaite et al., 2015; Lawlor et al., 2011; Lee et al., 2018; Lewis et al., 2010; Lim et al., 2009; Nordestgaard et al., 2012; Nordestgaard et al., 2017; Rode et al., 2014; Skaaby et al., 2018; Stender et al., 2013; Tan et al., 2019; Thakkinstian et al., 2014; Timpson et al., 2009; Timpson et al., 2011; Varbo et al., 2015; Vimalaswaran et al., 2013; Warodomwicht et al., 2013; Welsh et al., 2010; Winter-Jensen et al., 2020)
rs1421085	20	(Brower et al., 2019; Carreras-Torres et al., 2018; Censin et al., 2017; den Hollander et al., 2017; Dusingize et al., 2020; Gao et al., 2016; Geng et al., 2018; Hagg et al., 2015; Holmes et al., 2014; Jacobs et al., 2020; Jarvis et al., 2016; Kemp et al., 2016; Kivimäki et al., 2011; Lim et al., 2009; Liu et al., 2020; Martins-Silva et al., 2019; Schnurr et al., 2018; van den Broek et al., 2018; Zhang et al., 2020; Zhou et al., 2019)
rs1121980	5	(Barning and Abarin, 2016; Lyngdoh et al., 2012; Mumby et al., 2011; Shungin et al., 2015; Skaaby et al., 2018)
rs17817449	5	(Huang et al., 2016; Lee et al., 2019; Lim et al., 2009; Wang, Cheng, et al., 2018; Wang, Lu, et al., 2018)
rs7193144	3	(Censin et al., 2019; Lim et al., 2009; Lyngdoh et al., 2012)
rs11644943	3	(Chen, Fan, Huang, et al., 2019; Chen, Fan, Yang, et al., 2019; Fan et al., 2018)
rs11642015	3	(Gharahkhani et al., 2019; Speed et al., 2019; Zhao, Xu, et al., 2020)
rs9930506	2	(den Hollander et al., 2017; Thrift et al., 2014)
rs2075205	2	(Gharahkhani et al., 2019; Vasan et al., 2019)
rs8050136	1	(Lim et al., 2009)

rs3751812	1	(Hung et al., 2014)
rs8044769	1	(Panoutsopoulou et al., 2014)
rs62048402	1	(Fan et al., 2018)
rs9939973	1	(Lim et al., 2009)
rs9940128	1	(Lim et al., 2009)
rs1121980	1	(Lim et al., 2009)
rs9926289	1	(Lim et al., 2009)
rs1075440	1	(Marini et al., 2020)
rs1421090	1	(Marini et al., 2020)
rs56094641	1	(Richardson et al., 2020)
rs2072518	1	(Takahashi et al., 2019)
rs8047395	1	(Takahashi et al., 2019)
rs12597712	1	(Zeng et al., 2019)
rs9931164	1	(Jacobs et al., 2020)
rs3751813	1	(Vasan et al., 2019)
rs2540769	1	(Lyngdoh et al., 2012)
rs2665272	1	(Lyngdoh et al., 2012)
rs6499658	1	(Lyngdoh et al., 2012)
rs8060649	1	(Speed et al., 2019)
rs16952465	1	(Gharahkhani et al., 2019)

Abbreviations: FTO, fat mass and obesity-associated; SNPs, single-nucleotide polymorphisms.

Table S3*Characteristics of included studies.*

Characteristics of studies	n	References
Types of data used	n=128	
One-sample only	82	(Afzal et al., 2014; Barning and Abarin, 2016; Benn et al., 2016; Bonnefond et al., 2017; Brennan et al., 2009; Chatterjee et al., 2017; Chen, Fan, Huang, et al., 2019; Colak et al., 2016; Cole et al., 2016; Dale et al., 2017; Davies et al., 2015; den Hollander et al., 2017; Dixon et al., 2016; Eriksson et al., 2017; Fall et al., 2013; Fall et al., 2015; Fan et al., 2018; Gharahkhani et al., 2019; Gianfrancesco, Glymour, et al., 2017; Gianfrancesco, Stridh, et al., 2017; Granell et al., 2014; Guo et al., 2017; Hagg et al., 2015; Holmes et al., 2014; Howe et al., 2017; Huang et al., 2016; Hung et al., 2014; Jarvis et al., 2016; Jokela et al., 2012; Kaltoft et al., 2020; Kemp et al., 2016; Kivimäki et al., 2011; Kivimaki et al., 2008; Klovaitte et al., 2015; Larsson et al., 2020; Lawlor et al., 2011; Lee et al., 2018; Lee et al., 2019; Lewis et al., 2010; Lim et al., 2009; Lyall et al., 2017; Lyngdoh et al., 2012; Mao et al., 2017; Mukherjee et al., 2015; Mumby et al., 2011; Nordestgaard et al., 2012; Painter et al., 2016; Panoutsopoulou et al., 2014; Qian, Rookus, et al., 2019; Qian, Wang, et al., 2019; Reed et al., 2020; Richmond et al., 2014; Robinson et al., 2020; Rode et al., 2014; Schnurr et al., 2018; Shungin et al., 2015; Stender et al., 2013; Sun et al., 2020; Sun et al., 2019; Tan et al., 2019; Taylor et al., 2016; Thakkinstian et al., 2014; Thrift et al., 2015; Thrift et al., 2014; Timpson et al., 2009; Timpson et al., 2011; Todd et al., 2015; Tyrrell et al., 2016; van 't Hof et al., 2017; Varbo et al., 2015; Vimalaswaran et al., 2013; Wainberg et al., 2019; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015; Wang, Cheng, et al., 2018; Wang, Lu, et al., 2018; Wang et al., 2016; Wang, Zhang, et al., 2018; Warodomwicht et al., 2013; Welsh et al., 2010; Winter-Jensen et al., 2020; Wurtz et al., 2014)
Two-sample only	30	(Bae and Lee, 2019; Brower et al., 2019; Busch et al., 2019; Carreras-Torres et al., 2016; Censin et al., 2017; Corbin et al., 2016; Dusingize et al., 2020; Gao et al., 2016; Geng et al., 2018; Hartwig et al., 2016; Jacobs et al., 2020; Larsson et al., 2018; Lindstrom et al., 2017; Liu et al., 2020; Marini et al., 2020; Martins-Silva et al., 2019; Mokry et al., 2016; Noyce et al., 2017; Ooi et al., 2019; Richardson et al., 2020; Speed et al., 2019; Takahashi et al., 2019; Taylor et al., 2019; van den Broek et al., 2018; Vasan et al., 2019; Xu et al., 2017; Zeng et al., 2019; Zhang et al., 2020; Zhao, Xu, et al., 2020; Zhou et al., 2019)
Both one and two samples	16	(Budu-Aggrey et al., 2019; Carreras-Torres et al., 2018; Censin et al., 2019; Chen, Fan, Yang, et al., 2019; Fussey et al., 2020; Guo et al., 2016; Howe et al., 2019; Kurz and Laxy, 2020; Nordestgaard et al., 2017; Reed et al., 2017; Shu et al., 2019; Skaaby et al., 2018; Tyrrell et al., 2019; Wade, Carlslake, et al., 2018; Wade, Chiesa, et al., 2018; Xu et al., 2019)
Types of instruments used in the main analysis	n=128	
GRS only	63	(Afzal et al., 2014; Benn et al., 2016; Carreras-Torres et al., 2018; Chen, Fan, Huang, et al., 2019; Chen, Fan, Yang, et al., 2019; Colak et al., 2016; Cole et al., 2016; Davies et al., 2015; den Hollander et al., 2017; Dixon et al., 2016; Eriksson et al., 2017; Fall et al., 2015; Fan et al., 2018; Gianfrancesco, Glymour, et al., 2017; Gianfrancesco, Stridh, et al., 2017; Granell et al., 2014; Guo et al., 2017; Guo et al., 2016; Hagg et al., 2015; Holmes et al., 2014; Howe et al., 2019; Howe et al., 2017; Huang et al., 2016; Jokela et al., 2012; Kaltoft et al., 2020; Lee et al., 2018; Lee et al., 2019; Lyall et al., 2017; Lyngdoh et al., 2012; Mao et al., 2017; Mukherjee et al., 2015; Mumby et al., 2011; Nordestgaard et al., 2012; Nordestgaard et al., 2017; Painter et al., 2016; Qian, Rookus, et al., 2019; Qian, Wang, et al., 2019; Reed et al., 2020; Richmond et al., 2014; Robinson et al., 2020; Rode et al., 2014; Schnurr et al., 2018; Shungin et al., 2015; Skaaby et al., 2018; Stender et al., 2013; Sun et

		al., 2020; Sun et al., 2019; Taylor et al., 2016; Thrift et al., 2015; Thrift et al., 2014; Todd et al., 2015; Tyrrell et al., 2016; Varbo et al., 2015; Vimalaewaran et al., 2013; Wade, Carslake, et al., 2018; Wade, Chiesa, et al., 2018; Wang, Cheng, et al., 2018; Wang, Lu, et al., 2018; Wang et al., 2016; Wang, Zhang, et al., 2018; Welsh et al., 2010; Winter-Jensen et al., 2020; Wurtz et al., 2014)
Multiple IVs only	40	(Bae and Lee, 2019; Bonnefond et al., 2017; Brower et al., 2019; Budu-Aggrey et al., 2019; Busch et al., 2019; Carreras-Torres et al., 2016; Censin et al., 2017; Corbin et al., 2016; Dale et al., 2017; Dusingize et al., 2020; Gao et al., 2016; Geng et al., 2018; Gharahkhani et al., 2019; Hartwig et al., 2016; Jacobs et al., 2020; Jarvis et al., 2016; Larsson et al., 2020; Larsson et al., 2018; Lindstrom et al., 2017; Liu et al., 2020; Marini et al., 2020; Martins-Silva et al., 2019; Mokry et al., 2016; Noyce et al., 2017; Ooi et al., 2019; Richardson et al., 2020; Speed et al., 2019; Takahashi et al., 2019; Taylor et al., 2019; Timpson et al., 2009; van 't Hof et al., 2017; van den Broek et al., 2018; Vasana et al., 2019; Wainberg et al., 2019; Xu et al., 2017; Xu et al., 2019; Zeng et al., 2019; Zhang et al., 2020; Zhao, Xu, et al., 2020; Zhou et al., 2019)
Single IVs only	13	(Barning and Abarin, 2016; Brennan et al., 2009; Fall et al., 2013; Kivimäki et al., 2011; Kivimaki et al., 2008; Klovaite et al., 2015; Lewis et al., 2010; Lim et al., 2009; Panoutsopoulou et al., 2014; Tan et al., 2019; Thakkinstian et al., 2014; Timpson et al., 2011; Warodomwicht et al., 2013)
Both GRS and multiple IVs	6	(Censin et al., 2019; Fussey et al., 2020; Kurz and Laxy, 2020; Reed et al., 2017; Shu et al., 2019; Tyrrell et al., 2019)
Both GRS and single IVs	5	(Chatterjee et al., 2017; Hung et al., 2014; Kemp et al., 2016; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015)
Both single and multiple IVs	1	(Lawlor et al., 2011)
Types of GRS used among studies using GRS as IV in the main analysis	74	
Weighted GRS only	48	(Carreras-Torres et al., 2018; Censin et al., 2019; Chatterjee et al., 2017; Chen, Fan, Huang, et al., 2019; Chen, Fan, Yang, et al., 2019; Davies et al., 2015; Dixon et al., 2016; Eriksson et al., 2017; Fan et al., 2018; Fussey et al., 2020; Gianfrancesco, Glymour, et al., 2017; Granell et al., 2014; Guo et al., 2017; Guo et al., 2016; Holmes et al., 2014; Howe et al., 2019; Howe et al., 2017; Huang et al., 2016; Hung et al., 2014; Lee et al., 2019; Lyall et al., 2017; Mao et al., 2017; Mukherjee et al., 2015; Painter et al., 2016; Qian, Rookus, et al., 2019; Qian, Wang, et al., 2019; Reed et al., 2017; Reed et al., 2020; Richmond et al., 2014; Robinson et al., 2020; Schnurr et al., 2018; Shu et al., 2019; Sun et al., 2020; Sun et al., 2019; Taylor et al., 2016; Thrift et al., 2015; Todd et al., 2015; Tyrrell et al., 2016; Tyrrell et al., 2019; Vimalaewaran et al., 2013; Wade, Carslake, et al., 2018; Wade, Chiesa, et al., 2018; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015; Wang, Cheng, et al., 2018; Wang, Lu, et al., 2018; Wang, Zhang, et al., 2018; Wurtz et al., 2014)
Unweighted GRS only	19	(Afzal et al., 2014; Benn et al., 2016; Colak et al., 2016; Cole et al., 2016; den Hollander et al., 2017; Fall et al., 2015; Hagg et al., 2015; Jokela et al., 2012; Kemp et al., 2016; Kurz and Laxy, 2020; Lyngdoh et al., 2012; Mumby et al., 2011; Nordestgaard et al., 2012; Rode et al., 2014; Shungin et al., 2015; Skaaby et al., 2018; Stender et al., 2013; Varbo et al., 2015; Welsh et al., 2010)
Both weighted and unweighted GRS	7	(Gianfrancesco, Stridh, et al., 2017; Kalsoft et al., 2020; Lee et al., 2018; Nordestgaard et al., 2017; Thrift et al., 2014; Wang et al., 2016; Winter-Jensen et al., 2020)

Abbreviations: GRS, genetic risk score; IVs, instrumental variables; SNPs, single-nucleotide polymorphisms; *FTO*, fat mass and obesity-associated.

Table S4*Approaches cited for preventing pleiotropy in the selection of IVs.*

Approaches	n/33	References
Exclude selected SNPs known or suspected to be pleiotropic	12	(Bonnetfond et al., 2017; Censin et al., 2017; Chen, Fan, Huang, et al., 2019; Chen, Fan, Yang, et al., 2019; Fan et al., 2018; Howe et al., 2019; Sun et al., 2020; Sun et al., 2019; Tyrrell et al., 2016; Tyrrell et al., 2019; Xu et al., 2019; Zhou et al., 2019)
Use of multiple (independent) SNPs combined in a GRS as IVs	10	(Benn et al., 2016; Fall et al., 2015; Holmes et al., 2014; Kurz and Laxy, 2020; Shungin et al., 2015; Taylor et al., 2016; Todd et al., 2015; Varbo et al., 2015; Vimalleswaran et al., 2013; Winter-Jensen et al., 2020)
Use of multiple (independent) SNPs as IVs	10 ¹	(Carreras-Torres et al., 2016; Colak et al., 2016; Dale et al., 2017; Dixon et al., 2016; Eriksson et al., 2017; Lawlor et al., 2011; Lewis et al., 2010; Nordestgaard et al., 2012; Stender et al., 2013; Wurtz et al., 2014)
Use a single SNP as IV	1	(Tan et al., 2019)

Abbreviations: GRS, genetic risk score; SNP, single-nucleotide polymorphisms; IV, instrumental variable.

¹These 10 articles discussed the use of multiple independent SNPs without explicitly mentioning GRS even if six (Colak et al., 2016; Dixon et al., 2016; Eriksson et al., 2017; Nordestgaard et al., 2012; Stender et al., 2013; Wurtz et al., 2014) of the 10 in fact used a GRS as IV in the main analysis.

Table S5

Methods cited for detection or accounting the violation of the independence and/or restriction exclusion assumptions.

Methods ¹	n/128	Studies specifically referred to the method for 'pleiotropy' assessment / control / robustness (n=95)		Studies not specifically referred to the method for 'pleiotropy' assessment / control / robustness (n=96)	
		n	References	n	References
Robust and other IV estimation methods²	70	69		23	
Robust methods	69	68		20	
MR-Egger	69 ³	68	(Bae and Lee, 2019; Bonnefond et al., 2017; Brower et al., 2019; Budu-Aggrey et al., 2019; Busch et al., 2019; Carreras-Torres et al., 2016; Carreras-Torres et al., 2018; Censin et al., 2017; Censin et al., 2019; Chen, Fan, Huang, et al., 2019; Chen, Fan, Yang, et al., 2019; Corbin et al., 2016; Dale et al., 2017; Dixon et al., 2016; Dusingize et al., 2020; Fussey et al., 2020; Gao et al., 2016; Geng et al., 2018; Gharahkhani et al., 2019; Guo et al., 2017; Guo et al., 2016; Hartwig et al., 2016; Howe et al., 2019; Jacobs et al., 2020; Jarvis et al., 2016; Kemp et al., 2016; Kurz and Laxy, 2020; Larsson et al., 2020; Larsson et al., 2018; Lee et al., 2019; Lindstrom et al., 2017; Liu et al., 2020; Lyall et al., 2017; Mao et al., 2017; Marini et al., 2020; Martins-Silva et al., 2019; Mokry et al., 2016; Noyce et al., 2017; Ooi et al., 2019; Painter et al., 2016; Qian, Rookus, et al., 2019; Qian, Wang, et al., 2019; Reed et al., 2017; Richardson et al., 2020; Schnurr et al., 2018; Shu et al., 2019; Skaaby et al., 2018; Speed et al., 2019; Sun et al., 2020; Sun et al., 2019; Takahashi et al., 2019; Taylor et al., 2019; Tyrrell et al., 2016; Tyrrell et al., 2019; van 't Hof et al., 2017; van den Broek et al., 2018; Vasan et al., 2019; Wade, Carslake, et al., 2018; Wade, Chiesa, et al., 2018; Wainberg et al., 2019; Wang, Zhang, et al., 2018; Winter-Jensen et al., 2020; Xu et al., 2017; Xu et al., 2019; Zeng et al., 2019; Zhang et al., 2020; Zhao, Xu, et al., 2020; Zhou et al., 2019)	1	(Fan et al., 2018)

Median-based methods	44	26	(Bae and Lee, 2019; Budu-Aggrey et al., 2019; Busch et al., 2019; Carreras-Torres et al., 2016; Carreras-Torres et al., 2018; Censin et al., 2017; Censin et al., 2019; Chen, Fan, Huang, et al., 2019; Chen, Fan, Yang, et al., 2019; Dale et al., 2017; Fussey et al., 2020; Geng et al., 2018; Hartwig et al., 2016; Howe et al., 2019; Marini et al., 2020; Martins-Silva et al., 2019; Mokry et al., 2016; Reed et al., 2017; Skaaby et al., 2018; Speed et al., 2019; Taylor et al., 2019; Tyrrell et al., 2019; van den Broek et al., 2018; Vasan et al., 2019; Wade, Chiesa, et al., 2018; Xu et al., 2019)	18	(Corbin et al., 2016; Dusingize et al., 2020; Gharahkhani et al., 2019; Jacobs et al., 2020; Kurz and Laxy, 2020; Larsson et al., 2020; Larsson et al., 2018; Liu et al., 2020; Ooi et al., 2019; Sun et al., 2019; van 't Hof et al., 2017; Wade, Carslake, et al., 2018; Wang, Zhang, et al., 2018; Xu et al., 2017; Zeng et al., 2019; Zhang et al., 2020; Zhao, Xu, et al., 2020; Zhou et al., 2019)
Mode-based methods	9	5	(Budu-Aggrey et al., 2019; Carreras-Torres et al., 2018; Martins-Silva et al., 2019; Taylor et al., 2019; Zhou et al., 2019)	4	Jacobs et al., 2020; Marini et al., 2020; Ooi et al., 2019; Wade, Carslake, et al., 2018)
MR-PRESSO	6	5	(Marini et al., 2020; Ooi et al., 2019; Xu et al., 2019; Zeng et al., 2019; Zhou et al., 2019)	1	(Larsson et al., 2020)
MR-RAPS	1	0		1	(Liu et al., 2020)
Other IV estimation methods	13	6		8	
Multivariable MR	6	3	(Censin et al., 2017; Huang et al., 2016; Kemp et al., 2016)	3	(Dale et al., 2017; Jacobs et al., 2020; Marini et al., 2020)
IVW methods	4	3	(Chen, Fan, Yang, et al., 2019; Howe et al., 2019; Jarvis et al., 2016)	1	(Fan et al., 2018)
Likelihood-based methods	4	0		4	(Brower et al., 2019; Censin et al., 2017; Gharahkhani et al., 2019; Jacobs et al., 2020)
Methods to detect heterogeneity of estimated causal effects across IVs	72	46		46	
Graphical assessment: scatter plots, forest plots, funnel plots, leave-one-out plots, and histogram	34	20	(Bae and Lee, 2019; Censin et al., 2017; Chen, Fan, Huang, et al., 2019; Corbin et al., 2016; Dusingize et al., 2020; Fan et al., 2018; Gharahkhani et al., 2019; Granell et al., 2014; Hagg et al., 2015; Hartwig et al., 2016; Kemp et al., 2016; Lyall et al., 2017; Mokry et al., 2016; Noyce et al., 2017; Painter et al., 2016; Sun et al., 2019; Wade, Carslake, et al., 2018; Winter-Jensen et al., 2020; Zhao, Xu, et al., 2020; Zhou et al., 2019)	19	(Brower et al., 2019; Busch et al., 2019; Carreras-Torres et al., 2018; Chen, Fan, Huang, et al., 2019; Corbin et al., 2016; Davies et al., 2015; Dusingize et al., 2020; Gao et al., 2016; Guo et al., 2017; Guo et al., 2016; Larsson et al., 2018; Liu et al., 2020; Noyce et al., 2017; Ooi et al., 2019; Qian, Wang, et al., 2019; Takahashi et al., 2019; Taylor et al., 2019; Zeng et al., 2019; Zhou et al., 2019)

Statistical criteria and tests: I^2 , r^2 , H statistic, Cochran's Q, Rucker's Q, over-identification tests ⁴	30	7	(Censin et al., 2017; Granell et al., 2014; Hartwig et al., 2016; Kurz and Laxy, 2020; Painter et al., 2016; Qian, Rookus, et al., 2019; Reed et al., 2017)	23	(Bae and Lee, 2019; Carreras-Torres et al., 2018; Chen, Fan, Huang, et al., 2019; Dusingize et al., 2020; Fan et al., 2018; Gianfrancesco, Glymour, et al., 2017; Holmes et al., 2014; Larsson et al., 2018; Lawlor et al., 2011; Mao et al., 2017; Marini et al., 2020; Mukherjee et al., 2015; Nordestgaard et al., 2012; Noyce et al., 2017; Qian, Wang, et al., 2019; Richardson et al., 2020; Sun et al., 2020; Tyrrell et al., 2019; van den Broek et al., 2018; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015; Zeng et al., 2019; Zhou et al., 2019)
Comparisons of estimated MR causal effects across IVs (GRS or multiple IVs)	39	25		14	
Before and after exclusion of SNPs suspected of pleiotropy	33	20	(Benn et al., 2016; Carreras-Torres et al., 2016; Carreras-Torres et al., 2018; Davies et al., 2015; Gao et al., 2016; Gharahkhani et al., 2019; Hagg et al., 2015; Hartwig et al., 2016; Howe et al., 2017; Huang et al., 2016; Jacobs et al., 2020; Lyall et al., 2017; Shu et al., 2019; Speed et al., 2019; Taylor et al., 2019; Taylor et al., 2016; Todd et al., 2015; Wade, Chiesa, et al., 2018; Wang, Cheng, et al., 2018; Zeng et al., 2019)	13	(Censin et al., 2017; Chen, Fan, Huang, et al., 2019; Dale et al., 2017; Dusingize et al., 2020; Gianfrancesco, Glymour, et al., 2017; Guo et al., 2017; Holmes et al., 2014; Jarvis et al., 2016; Liu et al., 2020; Ooi et al., 2019; Sun et al., 2019; Wade, Carslake, et al., 2018; Wang, Lu, et al., 2018)
GRS/multiple IVs vs. single SNP(s)	5	5	(Budu-Aggrey et al., 2019; Colak et al., 2016; Eriksson et al., 2017; Fall et al., 2015; Richmond et al., 2014)	0	
Two subsets of SNPs grouping SNPs with the same biological pathway on the exposure	1	0		1	(Martins-Silva et al., 2019)
Detection of outlier/influential SNP(s): Cook's distance, Studentized residuals, HEIDI-outlier, and leave-one-out analyses	15	4	(Corbin et al., 2016; Takahashi et al., 2019; Wurtz et al., 2014; Zhou et al., 2019)	13	(Bae and Lee, 2019; Corbin et al., 2016; Hartwig et al., 2016; Jacobs et al., 2020; Liu et al., 2020; Martins-Silva et al., 2019; Noyce et al., 2017; Ooi et al., 2019; Takahashi et al., 2019; van den Broek et al., 2018; Zeng et al., 2019; Zhang et al., 2020; Zhao, Xu, et al., 2020)
Methods to detect associations between IVs and the outcome outside of the pathway through the exposure	85	30		61	
Estimating the associations between the IVs and measured risk factors for the outcome	75	19	(Benn et al., 2016; Budu-Aggrey et al., 2019; Carreras-Torres et al., 2016; Carreras-Torres et al., 2018; Davies et al., 2015; Gharahkhani et al., 2019; Howe et al., 2017; Kaltoft et al., 2020;	56	(Afzal et al., 2014; Barning and Abarin, 2016; Brennan et al., 2009; Censin et al., 2017; Censin et al., 2019; Chen, Fan, Huang, et al., 2019; Chen, Fan, Yang, et al., 2019; Colak et al., 2016; Dale et al.,

			Kemp et al., 2016; Kurz and Laxy, 2020; Lyall et al., 2017; Lyngdoh et al., 2012; Shu et al., 2019; Todd et al., 2015; Wade, Carslake, et al., 2018; Wang, Cheng, et al., 2018; Winter-Jensen et al., 2020; Zeng et al., 2019; Zhou et al., 2019)		2017; Dixon et al., 2016; Eriksson et al., 2017; Fan et al., 2018; Geng et al., 2018; Gianfrancesco, Glymour, et al., 2017; Granell et al., 2014; Guo et al., 2017; Guo et al., 2016; Hagg et al., 2015; Holmes et al., 2014; Howe et al., 2019; Huang et al., 2016; Hung et al., 2014; Jokela et al., 2012; Kivimäki et al., 2011; Kivimaki et al., 2008; Klovaite et al., 2015; Lawlor et al., 2011; Lee et al., 2018; Lee et al., 2019; Lewis et al., 2010; Mao et al., 2017; Nordestgaard et al., 2012; Nordestgaard et al., 2017; Qian, Wang, et al., 2019; Reed et al., 2017; Richmond et al., 2014; Rode et al., 2014; Schnurr et al., 2018; Shungin et al., 2015; Skaaby et al., 2018; Stender et al., 2013; Sun et al., 2020; Sun et al., 2019; Tan et al., 2019; Thakkestian et al., 2014; Thrift et al., 2015; Thrift et al., 2014; Timpson et al., 2009; Timpson et al., 2011; Tyrrell et al., 2019; Varbo et al., 2015; Wade, Chiesa, et al., 2018; Walter, Kubzansky, et al., 2015; Wang, Lu, et al., 2018; Warodomwichit et al., 2013; Wurtz et al., 2014)
Documenting the associations between the IV and risk factors for the outcome in the literature	7	5	(Carreras-Torres et al., 2016; Fall et al., 2013; Gao et al., 2016; Kaltoft et al., 2020; Wade, Carslake, et al., 2018)	2	(Chatterjee et al., 2017; Jarvis et al., 2016)
Adjusting IV-outcome or IV-confounders associations for exposure	18	7	(Chatterjee et al., 2017; Kivimäki et al., 2011; Painter et al., 2016; Robinson et al., 2020; Timpson et al., 2009; Wang, Cheng, et al., 2018; Wurtz et al., 2014)	11	(Cole et al., 2016; Eriksson et al., 2017; Gianfrancesco, Glymour, et al., 2017; Huang et al., 2016; Jokela et al., 2012; Kivimaki et al., 2008; Lawlor et al., 2011; Panoutsopoulou et al., 2014; Tyrrell et al., 2019; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015)
Adjusting IV-outcome association or MR analyses for covariates potentially involved in pleiotropic pathways ⁵	5	2	(Chatterjee et al., 2017; Kaltoft et al., 2020)	3	(Censin et al., 2019; Guo et al., 2016; Sun et al., 2020)
Comparison of the exposure-outcome conventional vs. IV estimated effects	4	1	(Painter et al., 2016)	3	(Kivimäki et al., 2011; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015)
Comparison of the IV-outcome vs. IV-exposure associations	2	2	(Barning and Abarin, 2016; Tyrrell et al., 2019)	0	
Estimating the association between the IV and the outcome	2	1	(Speed et al., 2019)	1	(Wang, Lu, et al., 2018)
Mediation analysis estimating the direct effect of the IV on the outcome	3	0		3	(Afzal et al., 2014; Fan et al., 2018; Gianfrancesco, Glymour, et al., 2017)

Other methods	5	0		5	
The use of positive or negative control outcomes	3	0		3	(Nordestgaard et al., 2017; Richardson et al., 2020; Tyrrell et al., 2019)
Colocalization	1	0		1	(Liu et al., 2020)
Verifying the concordance of MR results with those from other studies (MR, clinical trial)	1	0		1	(Holmes et al., 2014)
No method reported for assessment of independence or exclusion restriction assumptions	8		(den Hollander et al., 2017; Gianfrancesco, Stridh, et al., 2017; Lim et al., 2009; Mumby et al., 2011; Reed et al., 2020; Vimalaswaran et al., 2013; Wang et al., 2016; Welsh et al., 2010)		

Abbreviations: IV, instrumental variable; MR, mendelian randomization; PRESSO, pleiotropy residual sum and outlier; SNP, single-nucleotide polymorphism; RAPS, Robust Adjusted Prole Score; IVW, inverse-variance weighted; *FTO*, fat mass and obesity-associated; GRS, genetic risk score; HEIDI, Heterogeneity in Dependent Instruments.

¹Many studies cited more than one method and thus the number of methods reported in the Table exceed the total number of studies (n=128).

²See (Burgess et al., 2020) and (Burgess et al., 2015) for a summary of the listed methods.

³Of the 69 studies that reported using MR-Egger, 66 used the intercept test p-value to infer whether or not pleiotropy was present, two studies (Fan et al., 2018; Tyrrell et al., 2016) compared the MR-Egger slope and the conventional MR causal effect estimate, while the last study (Chen, Fan, Yang, et al., 2019) did not specify how the MR-Egger results were used.

⁴One study (Censin et al., 2017) did not specify the heterogeneity test used.

⁵Three articles (Censin et al., 2019; Guo et al., 2016; Sun et al., 2020) mentioned adjustment of MR analyses for covariates potentially involved in pleiotropic pathways without specifying as multivariable MR.

Table S6

Summary description of methods cited for detection or accounting the violation of the independence and/or restriction exclusion assumptions in the review.

Methods	Description	Main assumptions and/or limitations ¹	References
Robust and other IV estimation methods			
Robust methods			
MR-Egger	Regression of SNPs-outcome on SNPs-exposure effects weighted by the inverse-variance of the SNPs-outcome associations, allowing a non-zero intercept. The intercept of the regression may be interpreted as the estimated average pleiotropic effect, and the slope as the pleiotropy-adjusted causal effect estimates.	Assumes the InSIDE assumption (independence between the SNPs-exposure and the direct SNPs-outcome effects) ² , estimates are sensitive to outliers and influential datapoints, low precision of the intercept and causal estimates compared to methods like IVW, increased type 1 error if InSIDE is not satisfied.	(Bowden et al., 2015; Burgess et al., 2017; Burgess et al., 2020)
Median-based methods	Estimates the causal effect as the median of the of the SNPs-specific causal effect estimates.	Assumes the majority valid assumption (at least 50% of the SNPs are valid), likely to be inefficient when the individual SNP estimates receive equal weight in the analysis, may be affected by addition or removal of SNPs, increased type 1 error if there exist an asymmetric distribution of the SNP pleiotropic effects around zero.	(Bowden, Davey Smith, et al., 2016; Burgess et al., 2017; Burgess et al., 2020)
Mode-based methods	Estimates the causal effect as the mode of the smoothed empirical density function of the SNPs-specific causal effect estimates.	Assumes the plurality valid assumption (more SNPs estimate the true causal effect than estimate any other quantity), sensitive to bandwidth parameter required for density estimation, may be affected by addition or removal of SNPs, low power.	(Burgess et al., 2020; Carreras-Torres et al., 2018; Hartwig et al., 2017)
MR-PRESSO	Performs the IVW method using all SNPs and omitting each SNP in turn, and then excludes the SNP(s) whose omission showed a significant decrease in the residual sum of squares (used as a measure of heterogeneity of SNP ratio estimates).	Higher false positive rate if there are many invalid SNPs.	(Slob and Burgess, 2020; Verbanck et al., 2018)

MR-RAPS	Robust Adjusted Profile Score: Models the distribution of the pleiotropic effects of SNPs using a random-effect distribution.	Assumes the pleiotropic effects of SNPs to be symmetric around zero, may be affected by violations of balanced pleiotropy assumption.	(Slob and Burgess, 2020; Zhao, Wang, et al., 2020)
Other IV estimation methods			
Multivariable MR	Extension of MR analysis that estimates the direct causal effect of several risk factors on a single outcome using multiple SNPs.	Requires that each SNP be associated with at least one exposure and not affect the outcome other than through these exposures, does not address unmeasured pleiotropy, weak instrument bias may be substantial if large numbers of SNPs used.	(Burgess et al., 2020; Burgess and Thompson, 2015; Sanderson et al., 2019)
Inverse-variance weighted	Regression of SNPs-outcome on SNPs-exposure associations, weighted by the inverse of variance of SNPs-outcome associations, allowing a null intercept.	Assumes all SNPs to be valid, biased if average pleiotropic effect differs from zero, vulnerable to weak instrument bias, more efficient than MR-Egger.	(Bowden et al., 2015; Bowden, Del Greco M, et al., 2016; Burgess et al., 2013)
Likelihood-based methods	SNP effects on exposure and outcome are modelled as a bivariate normal distribution and estimated using maximum likelihood estimation or Bayesian methods.	Assumes linear relationship between the exposure and the outcome, and a bivariate normal distribution for the SNPs-exposure and SNPs-outcome associations, computationally complex.	(Burgess et al., 2013; Burgess et al., 2016; Burgess et al., 2015)
Methods to detect heterogeneity of estimated causal effects across IVs			
Graphical assessment: scatter plots, forest plots, funnel plots, leave-one-out plots, and histograms	Visual examination of whether the causal effect estimates from different IVs are similar. Scatter plots show the SNPs-outcome vs. SNPs-exposure effects, with the linear estimation of the overall causal effect. Forest plots show the causal effect of each SNP as well as the overall causal estimate. Leave-one-out plots show the overall polygenic causal estimate as well as the re-estimated causal effect after sequential removing of one SNP at a time. Funnel plots show the SNP strength vs. the causal estimate from each SNP. Histograms show the distribution of estimates from the leave-one out analysis.	Subjective interpretation of heterogeneity.	(Burgess et al., 2017; Burgess et al., 2020; Carreras-Torres et al., 2018; Hemani et al., 2018; Zhou et al., 2019)
Statistical criteria and tests: I^2 , r^2 , H statistic, Cochran's Q, Rucker's Q, over-identification tests	Statistical tests of heterogeneity of IV-specific causal effects.	Low power for detecting heterogeneity when the number of IVs is small, overestimation of heterogeneity in presence of all valid IVs with different magnitudes of causal effect.	(Bowden et al., 2017; Burgess et al., 2017; Glymour et al., 2012; Hartwig et al., 2016; Hemani et al., 2018; Zhou et al., 2019)

Comparisons of estimated MR causal effects across IVs (GRS or multiple IVs)			
Before and after exclusion of SNPs suspected of pleiotropy	Evaluate whether the causal effects estimated from polygenic MR analyses before and after exclusion of SNP(s) suspected of pleiotropy are consistent.	Subjective interpretation of consistency.	(Carreras-Torres et al., 2016; Hemani et al., 2018)
GRS or multiple IVs vs. single SNP(s)	Evaluate whether the causal effect estimated from a polygenic MR analysis is consistent with that using single SNP(s).	Subjective interpretation of consistency.	(Colak et al., 2016; Fall et al., 2015; Hemani et al., 2018)
Two subsets of SNPs grouping SNPs with the same biological pathway on the exposure	Evaluate whether the causal effects estimated from MR analyses using two subsets of SNPs, each grouping SNPs with the same biological pathway on the exposure, are consistent.	Subjective interpretation of consistency.	(Hemani et al., 2018; Martins-Silva et al., 2019)
Detection of outlier/influential SNP(s): Cook's distance, Studentized residuals, <i>HEIDI-outlier</i> , leave-one-out analyses	Detect outlier or influential SNP(s). Cook's distance is a measure of leverage, allowing to detect SNPs that exerted a disproportionately large influence on the overall causal effect. A Studentized residual is the ratio between a residual from the regression model and an estimate of its standard error, indicating the goodness-of-fit in the model for that point. The leave-one-out analysis evaluates whether the overall polygenic causal effect estimate is similar with the re-estimated causal effect after sequential removing of one SNP at a time. <i>HEIDI-outlier</i> uses linkage disequilibrium information to test whether there exist heterogeneity in the causal effects estimated from near-independent SNPs and exclude from analysis SNP(s) that have apparent pleiotropic effects on both exposure and outcome.	Cut-off values to determine influential SNPs are sensitive to sample size and may require subjective interpretation, <i>HEIDI-outlier</i> may not detect certainly pleiotropic outlier SNPs (e.g. those with very small effects).	(Burgess and Thompson, 2017; Hemani et al., 2018; Verbanck et al., 2018; Zhou et al., 2019; Zhu et al., 2016)
Methods to detect associations between IVs and the outcome outside of the pathway through the exposure			
Estimating the associations between the IVs and measured risk factors for the outcome	Statistically significant associations between IVs and measured covariates may be indicative of violation of MR assumptions.	Does not account for unmeasured covariates, significant association with a covariate may be due to other causes than pleiotropy (e.g., mediation, population stratification, selection/collider bias), non-significant association may be due to low power,	(Burgess et al., 2017; Burgess et al., 2020; Wade, Carslake, et al., 2018)

		statistical significance threshold may be difficult to set in case of large number of IVs.	
Documenting the associations between the IV and risk factors for the outcome in the literature	Documents reported associations between the IVs and covariates that may affect the outcome of interest which may be indicative of violation of MR assumptions.	Does not account for unmeasured covariates, significant association with a covariate may be due to other causes than pleiotropy (e.g. mediation, population stratification, selection/collider bias).	(Fall et al., 2013; Wade, Carslake, et al., 2018)
Adjusting IV-outcome or IV-confounders associations for exposure	Consists of assessing whether the association IV-outcome or IV-confounders is attenuated by adjustment for the exposure as would be expected if the effect of the IV on the outcome or on a confounder is mediated through the exposure.	Subjective interpretation of effect attenuation, does not account for unmeasured confounders.	(Walter, Glymour, et al., 2015)
Adjusting IV-outcome association or MR analyses for covariates potentially involved in pleiotropic pathways	Consists in assessing whether the association IV-outcome is attenuated by adjustment for a covariate associated with the outcome, as would be expected if the effect of the IV on the outcome is mediated through this covariate.	Subjective interpretation of effect attenuation, does not account for unmeasured confounders, may be difficult to distinguish vertical to horizontal pleiotropy.	(Chatterjee et al., 2017)
	Account pleiotropy by adjusting MR analyses for covariates potentially involved in pleiotropic pathways.	Requires individual-level data, does not deal with unmeasured confounders.	(Burgess et al., 2017)
Comparison of the exposure-outcome conventional vs. IV estimated effects	The IV is suspected to be invalid if the effect of the exposure on the outcome estimated by ordinary least squares is greater than that estimated by MR.	Assumes positive exposure-outcome unmeasured confounding, performs better in case of pleiotropy by direct effect of the IV on the outcome than by indirect effect via confounders, may detect valid IVs as invalids (false-positive rate).	(Burgess, 2012; Glymour et al., 2012)
Comparison of the IV-outcome vs. IV-exposure associations	Consists in justifying the non-violation of exclusion restriction given that the IVs are associated with exposure and not with the outcome.	A null association between the IV and the outcome may exist despite violation of exclusion restriction (e.g. when the direct and indirect effects of the IV on the outcome are opposite and symmetrical to zero, the total effect estimated by regressing the outcome on the IV may be non-significant).	(Barning and Abarin, 2016)
	Consists in suspecting the presence of pleiotropy given that the SNPs have larger effects on the outcome than their effects on the exposure.	SNP-exposure and SNP-outcome effects may be difficult to compare if evaluate on different scales.	(Tyrrell et al., 2019)

Estimating the association between the IV and the outcome	Suspects the violation of the exclusion restriction assumption by detecting significant associations between the SNPs and the outcome of interest.	A significant association between the SNPs and the outcome may be due to vertical pleiotropy.	(Speed et al., 2019; Wang, Lu, et al., 2018)
Mediation analysis estimating the direct effect of the IV on the outcome	Decomposes the total effect of the IV on the outcome into an indirect effect via the exposure and a direct effect (independent of exposure or pleiotropic).	Assumes no IV-outcome, exposure-outcome, and IV-exposure confounding; no exposure-outcome confounders affected by the IV, may be biased if unmanaged IV-exposure interaction, does not apply to unmeasured confounders.	(VanderWeele, 2016)
Other methods			
The use of positive or negative control outcomes	Assesses whether there is a genetic association between the exposure and an outcome for which one is certain that the exposure is causal (positive control), or for which one is not thought to be affected by the exposure (negative control).	Adequate negative and positive controls may be difficult to find, a significant genetic association may not be indicative of pleiotropy.	(Burgess and Davey Smith, 2017; Burgess et al., 2020)
Colocalization	Aims to detect whether a shared causal SNP between the two traits (exposure and outcome) is plausible.	Low statistical power, may be inexact if there exist multiple conditionally independent causal variants in the cis region, or if the exposure and outcome trait effects were estimated in populations with different linkage disequilibrium patterns.	(Burgess et al., 2020; Giambartolomei et al., 2014; Hemani et al., 2018; Liu et al., 2020)
Verifying the concordance of MR results with those from other studies (MR, clinical trial)	Justify the whether the exclusion restriction assumption is not violated based on the concordance between its results and those of previous MR studies using different genetic variants and of a clinical trial.	Concordance may be due to chance, different populations, different methods, depends on the quality of the study compared to.	(Holmes et al., 2014)

Note. Entries from this table are drawn from several publications (Burgess et al., 2017; Burgess et al., 2020; Burgess and Thompson, 2017; Hemani et al., 2018; Slob and Burgess, 2020). References are provided for a more detailed descriptions of each method.

Abbreviations: MR, mendelian randomization; IV, instrumental variable; SNP, single-nucleotide polymorphism; InSIDE, Instrument Strength Independent of Direct Effect; IVW, inverse-variance weighted; HEIDI, heterogeneity in dependent instruments; PRESSO, Pleiotropy Residual Sum and Outlier; RAPS, Robust Adjusted Profile Score.

¹All the methods are vulnerable to measurement errors.

²The MR-egger intercept test does not rely to the InSIDE assumption, but the estimation of the slope as well as the interpretation of the intercept coefficient as the average pleiotropic effect require that InSIDE be satisfied (Burgess and Thompson, 2017).

Table S7*Discussion of the independence and/or exclusion restriction assumptions.*

Discussion of the independence and/or exclusion restriction assumptions	n	References
<p>Plausibility of the independence and/or exclusion restriction assumptions</p> <p>Discussed with specific reference to pleiotropy</p> <p>Discussed without specific reference to pleiotropy</p>	<p>n=128</p> <p>89</p> <p>19</p>	<p>(Bae and Lee, 2019; Benn et al., 2016; Brower et al., 2019; Budu-Aggrey et al., 2019; Carreras-Torres et al., 2016; Censin et al., 2017; Chatterjee et al., 2017; Chen, Fan, Huang, et al., 2019; Chen, Fan, Yang, et al., 2019; Colak et al., 2016; Cole et al., 2016; Corbin et al., 2016; Dale et al., 2017; Davies et al., 2015; Dixon et al., 2016; Dusingize et al., 2020; Eriksson et al., 2017; Fall et al., 2013; Fall et al., 2015; Gao et al., 2016; Geng et al., 2018; Granell et al., 2014; Guo et al., 2016; Hagg et al., 2015; Hartwig et al., 2016; Holmes et al., 2014; Howe et al., 2019; Huang et al., 2016; Jacobs et al., 2020; Jarvis et al., 2016; Jokela et al., 2012; Kaltoft et al., 2020; Kemp et al., 2016; Kivimäki et al., 2011; Larsson et al., 2020; Larsson et al., 2018; Lee et al., 2018; Lee et al., 2019; Lewis et al., 2010; Lindstrom et al., 2017; Liu et al., 2020; Lyall et al., 2017; Lyngdoh et al., 2012; Mao et al., 2017; Marini et al., 2020; Martins-Silva et al., 2019; Mokry et al., 2016; Mukherjee et al., 2015; Mumby et al., 2011; Nordestgaard et al., 2012; Noyce et al., 2017; Ooi et al., 2019; Painter et al., 2016; Reed et al., 2017; Richardson et al., 2020; Richmond et al., 2014; Robinson et al., 2020; Schnurr et al., 2018; Shu et al., 2019; Shungin et al., 2015; Skaaby et al., 2018; Speed et al., 2019; Stender et al., 2013; Sun et al., 2020; Sun et al., 2019; Takahashi et al., 2019; Tan et al., 2019; Taylor et al., 2019; Taylor et al., 2016; Thrift et al., 2015; Thrift et al., 2014; Timpson et al., 2009; Timpson et al., 2011; Todd et al., 2015; Tyrrell et al., 2016; Varbo et al., 2015; Vimalleswaran et al., 2013; Wade, Carslake, et al., 2018; Wade, Chiesa, et al., 2018; Wang, Cheng, et al., 2018; Welsh et al., 2010; Winter-Jensen et al., 2020; Wurtz et al., 2014; Xu et al., 2017; Xu et al., 2019; Zeng et al., 2019; Zhang et al., 2020; Zhao, Xu, et al., 2020; Zhou et al., 2019)</p> <p>(Afzal et al., 2014; Busch et al., 2019; Carreras-Torres et al., 2018; Censin et al., 2019; Gianfrancesco, Glymour, et al., 2017; Hung et al., 2014; Kivimaki et al., 2008; Klovaite et al., 2015; Kurz and Laxy, 2020; Lawlor et al., 2011; Panoutsopoulou et al., 2014; Qian, Wang, et al., 2019; Rode et al., 2014; Thakkinstian et al., 2014; Tyrrell et al., 2019; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015; Wang, Lu, et al., 2018; Warodomwichit et al., 2013)</p>
<p>IV invalidity¹</p> <p>Suspected with specific reference to pleiotropy</p>	<p>n=128</p> <p>16</p>	<p>(Censin et al., 2017; Corbin et al., 2016; Gharahkhani et al., 2019; Howe et al., 2019; Kivimäki et al., 2011; Larsson et al., 2020; Painter et al., 2016; Shu et al., 2019; Speed et al., 2019; Sun et al., 2020; Takahashi et al., 2019; van den Broek et al., 2018; Wade, Carslake, et al., 2018; Wang, Cheng, et al., 2018; Winter-Jensen et al., 2020; Xu et al., 2019)</p>

Suspected without specific reference to pleiotropy	5	(Afzal et al., 2014; Gianfrancesco, Glymour, et al., 2017; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015; Wang, Lu, et al., 2018)
Impact of IV invalidity on the validity of MR results	n=21	
May have affected validity of results	8	(Censin et al., 2017; Corbin et al., 2016; Howe et al., 2019; Kivimäki et al., 2011; Painter et al., 2016; van den Broek et al., 2018; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015)
No or low impact on validity of results	9	(Gharahkhani et al., 2019; Gianfrancesco, Glymour, et al., 2017; Shu et al., 2019; Speed et al., 2019; Takahashi et al., 2019; Wade, Carslake, et al., 2018; Wang, Lu, et al., 2018; Winter-Jensen et al., 2020; Xu et al., 2019)
Impact not (clearly) reported	4	(Afzal et al., 2014; Larsson et al., 2020; Sun et al., 2020; Wang, Cheng, et al., 2018)
Suspicion of invalidity/pleiotropy of FTO SNP(s)	n=128	
Yes	6	(Gianfrancesco, Glymour, et al., 2017; Kivimäki et al., 2011; van den Broek et al., 2018; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015; Wang, Cheng, et al., 2018)
FTO SNP(s) suspected to be invalid/pleiotropic	n=6	
rs1558902 ²	3	(Gianfrancesco, Glymour, et al., 2017; Walter, Glymour, et al., 2015; Walter, Kubzansky, et al., 2015)
rs1421085 ³	2	(Kivimäki et al., 2011; van den Broek et al., 2018)
rs17817449 ⁴	1	(Wang, Cheng, et al., 2018)

Abbreviations: IV, instrumental variable; MR, mendelian randomization; SNP, single-nucleotide polymorphism; *FTO*, fat mass and obesity-associated.

¹Refers to the suspicion of invalidity of one or more body mass index IV(s) for any outcome of interest.

²The outcomes of interest involved in the suspected invalidity of rs1558902 were multiple sclerosis susceptibility (Gianfrancesco, Glymour, et al., 2017), phobic anxiety symptoms (Walter, Glymour, et al., 2015), and depression (Walter, Kubzansky, et al., 2015).

³The outcomes of interest involved in the suspected pleiotropy of rs1421085 were common mental disorders (Kivimäki et al., 2011), and subjective well-being (van den Broek et al., 2018).

⁴The outcome of interest involved in the suspected pleiotropy of rs17817449 was lipid profiles (Wang, Cheng, et al., 2018).

Table S8

List of 21 studies that suspected IV pleiotropy/invalidity.¹

Studies	SNP suspected of pleiotropy /invalidity	Chr ²	Nearest gene ³	Variable suspected in a pathway from the IV to the outcome outside of the exposure pathway, if any ⁴	Outcome(s) ⁵
Afzal et al., 2014	Not reported	-	-	NA	Type 2 diabetes
Censin et al., 2017	rs13130484	4	<i>GNPDA2</i>	Education	Type 1 diabetes
	rs13107325	4	<i>SLC39A8</i>	Education	
Corbin et al., 2016	rs7903146	10	<i>TCF7L2</i>	NA	Type 2 diabetes
Gharahkhani et al., 2019	Not reported	-	-	Smoking, alcohol intake, coffee/tea consumption, height	Overall cancer risk and mortality, and specific type cancer risk
Gianfrancesco, Glymour, et al., 2017	rs11126666	2	<i>KCNK3</i>	NA	Multiple sclerosis
	rs2112347	5	<i>POC5</i>		
	rs1558902	16	<i>FTO</i>		
	rs7243357	18	<i>GRP</i>		
	rs7599312	2	<i>ERBB4</i>		
Howe et al., 2019	Not reported	-	-	NA	Education
Kivimäki et al., 2011	rs1421085	16	<i>FTO</i>	NA	Common mental disorders
Larsson et al., 2020	Not reported	-	-	NA	Coronary artery disease, atrial fibrillation, and transient ischaemic attack
Painter et al., 2016	rs2075650	19	<i>TOMM40</i>	NA	Endometrial cancer
Shu et al., 2019	rs1800437	19	<i>GIPR</i>	2-h glucose	Breast cancer risk
	rs1801282	3	<i>PPARG</i>	Fasting insulin	
	rs2176040	2	<i>NYAP2, MIR5702</i>	Fasting insulin	
	rs7903146	10	<i>TCF7L2</i>	Fasting glucose, fasting insulin	
Speed et al., 2019	Not reported	-	-	NA	Depression
Sun et al., 2020	Not reported	-	-	NA	Doctor-diagnosed active asthma
Takahashi et al., 2019	rs1814170	7	<i>ZC3HAV1</i>	NA	Meningioma
	rs3803286	14	<i>TRAF3</i>		
	rs7084454	10	<i>MLLT10</i>		
van den Broek et al., 2018	rs1421085	16	<i>FTO</i>	NA	Subjective well-being
	rs943005	6	<i>FTH1P5, RPS17P5</i>		

	rs13021737	2	<i>TMEM18,</i> <i>LINC01875</i>		
Wade, Carslake, et al., 2018	rs17024393	1	<i>GNAT2</i>	NA	Mortality from other cancers
	rs13107325	4	<i>SLC39A8</i>		
	rs17001654	4	<i>SCARB2</i>		
	rs1167827	7	<i>HIP1</i>		
	rs11030104	11	<i>BDNF</i>		
	rs3888190	16	<i>ATP2A1</i>		
	rs2075650	19	<i>TOMM40</i>		
Walter, Glymour, et al., 2015	rs1558902	16	<i>FTO</i>	NA	Phobic anxiety symptoms
Walter, Kubzansky, et al., 2015	rs1558902	16	<i>FTO</i>	NA	Depression
Wang, Cheng, et al., 2018	rs4776970	15	<i>MAP2K5</i>	Blood pressure	Lipid profiles
	rs12597579	16	<i>SNRPEP3, GP2</i>	Blood pressure	
	rs17817449	16	<i>FTO</i>	Blood pressure, age	
	rs261967	5	<i>PCSK1, CAST</i>	Blood pressure	
	rs4715210	6	<i>FTHIP5</i>	Blood pressure, fasting plasma glucose, age	
	rs6567160	18	<i>MC4R</i>	Blood pressure	
	rs574367	1	<i>SEC16B</i>	Blood pressure	
	rs6545814	2	<i>ADCY3</i>	Fasting plasma glucose	
Wang, Lu, et al., 2018 ⁶	Not reported	-	-	Lipid profile, hypertension	Blood lead level
Winter-Jensen et al., 2020	Not reported	-	-	Smoking status	Hospital contacts related to infections
Xu et al., 2019	rs2287019	19	<i>QPCTL</i>	NA	Asthma

Abbreviations: IV, instrumental variable; SNPs, single nucleotide polymorphisms; Chr, chromosome; NA, not applicable.

¹ Corresponds to studies that suspected invalidity of the IV using MR assumption assessment methods, regardless of whether there was an impact on the validity of the results.

² When not explicitly reported in the manuscript, the chromosome was identified by searching the SNP in GWAS catalog (<https://www.ebi.ac.uk/gwas/>).

³ When not explicitly reported in the manuscript, the nearest gene was identified by searching in GWAS catalog (<https://www.ebi.ac.uk/gwas/>). In instances where a SNP appeared close to two genes, we reported both.

⁴ Applicable to studies that reported associations between the IVs and measured covariates.

⁵ If several outcomes of interest were considered in the study, we reported only that (those) suspected of being affected by the SNP irrespective of the pathway via exposure.

⁶ When discussing the validity of the exclusion restriction assumption, the study mentioned 3 SNPs (rs11671664, rs261967, and rs3932549) associated with the outcome of interest, but whose exclusion did not change the results.

3. References

- Afzal, S., Brøndum-Jacobsen, P., Bojesen, S. E., and Nordestgaard, B. G. (2014). Vitamin D concentration, obesity, and risk of diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol.* 2, 298-306. doi:10.1016/S2213-8587(13)70200-6
- Bae, S.-C., and Lee, Y. H. (2019). Causal association between body mass index and risk of rheumatoid arthritis: A Mendelian randomization study. *Eur. J. Clin. Invest.* 49, e13076. doi:10.1111/eci.13076
- Barning, F., and Abarin, T. (2016). Assessing the Causality Factors in the Association between (Abdominal) Obesity and Physical Activity among the Newfoundland Population-A Mendelian Randomization Analysis. *Genet. Epigenet.* 8, 15-24. doi:10.4137/GEG.S38289
- Benn, M., Tybjærg-Hansen, A., Davey Smith, G., and Nordestgaard, B. G. (2016). High body mass index and cancer risk-a Mendelian randomisation study. *Eur. J. Epidemiol.* 31, 879-892. doi:10.1007/s10654-016-0147-5
- Boef, A. G. C., Dekkers, O. M., and le Cessie, S. (2015). Mendelian randomization studies: a review of the approaches used and the quality of reporting. *Int. J. Epidemiol.* 44, 496-511. doi:10.1093/ije/dyv071
- Bonnefond, A., Yengo, L., Dechaume, A., Canouil, M., Castelain, M., Roger, E., et al. (2017). Relationship between salivary/pancreatic amylase and body mass index: a systems biology approach. *BMC Med.* 15, 37. doi:10.1186/s12916-017-0784-x
- Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* 44, 512-525. doi:10.1093/ije/dyv080
- Bowden, J., Davey Smith, G., Haycock, P. C., and Burgess, S. (2016). Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* 40, 304-314. doi:10.1002/gepi.21965
- Bowden, J., Del Greco M, F., Minelli, C., Davey Smith, G., Sheehan, N., and Thompson, J. (2017). A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat. Med.* 36, 1783-1802. doi:10.1002/sim.7221

- Bowden, J., Del Greco M, F., Minelli, C., Davey Smith, G., Sheehan, N. A., and Thompson, J. R. (2016). Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I² statistic. *Int. J. Epidemiol.* 45, 1961-1974. doi:10.1093/ije/dyw220
- Brennan, P., McKay, J., Moore, L., Zaridze, D., Mukeria, A., Szeszenia-Dabrowska, N., et al. (2009). Obesity and cancer: Mendelian randomization approach utilizing the FTO genotype. *Int. J. Epidemiol.* 38, 971-975. doi:10.1093/ije/dyp162
- Brower, M. A., Hai, Y., Jones, M. R., Guo, X., Chen, Y. D. I., Rotter, J. I., et al. (2019). Bidirectional Mendelian randomization to explore the causal relationships between body mass index and polycystic ovary syndrome. *Hum. Reprod.* 34, 127-136. doi:10.1093/humrep/dey343
- Budu-Aggrey, A., Brumpton, B., Tyrrell, J., Watkins, S., Modalsli, E. H., Celis-Morales, C., et al. (2019). Evidence of a causal relationship between body mass index and psoriasis: A mendelian randomization study. *PLoS Med.* 16, e1002739. doi:10.1371/journal.pmed.1002739
- Burgess, S. (2012). Re: "credible mendelian randomization studies: Approaches for evaluating the instrumental variable assumptions". *Am. J. Epidemiol.* 176, 456-457. doi:10.1093/aje/kws249
- Burgess, S., Bowden, J., Fall, T., Ingelsson, E., and Thompson, S. G. (2017). Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiology* 28, 30-42. doi:10.1097/EDE.0000000000000559
- Burgess, S., Butterworth, A., and Thompson, S. G. (2013). Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* 37, 658-665. doi:10.1002/gepi.21758
- Burgess, S., and Davey Smith, G. (2017). Mendelian randomization implicates high-density lipoprotein cholesterol-associated mechanisms in etiology of age-related macular degeneration. *Ophthalmology* 124, 1165-1174. doi:10.1016/j.ophtha.2017.03.042

- Burgess, S., Davey Smith, G., Davies, N. M., Dudbridge, F., Gill, D., Glymour, M. M., et al. (2020). Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res.* 4, 186. doi:10.12688/wellcomeopenres.15555.2
- Burgess, S., Dudbridge, F., and Thompson, S. G. (2016). Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Stat. Med.* 35, 1880-1906. doi:10.1002/sim.6835
- Burgess, S., Scott, R. A., Timpson, N. J., Davey Smith, G., Thompson, S. G., and Epic- InterAct Consortium (2015). Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur. J. Epidemiol.* 30, 543-552. doi:10.1007/s10654-015-0011-z
- Burgess, S., and Thompson, S. G. (2015). Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am. J. Epidemiol.* 181, 251-260. doi:10.1093/aje/kwu283
- Burgess, S., and Thompson, S. G. (2017). Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* 32, 377-389. doi:10.1007/s10654-017-0255-x
- Busch, A. S., Hollis, B., Day, F. R., Sorensen, K., Aksglaede, L., Perry, J. R. B., et al. (2019). Voice break in boys-temporal relations with other pubertal milestones and likely causal effects of BMI. *Hum. Reprod.* 34, 1514-1522. doi:10.1093/humrep/dez118
- Carreras-Torres, R., Haycock, P. C., Relton, C. L., Martin, R. M., Smith, G. D., Kraft, P., et al. (2016). The causal relevance of body mass index in different histological types of lung cancer: A Mendelian randomization study. *Sci. Rep.* 6, 31121. doi:10.1038/srep31121
- Carreras-Torres, R., Johansson, M., Haycock, P. C., Relton, C. L., Smith, G. D., Brennan, P., et al. (2018). Role of obesity in smoking behaviour: Mendelian randomisation study in UK Biobank. *Br. Med. J.* 361, k1767. doi:10.1136/bmj.k1767
- Censin, J. C., Nowak, C., Cooper, N., Bergsten, P., Todd, J. A., and Fall, T. (2017). Childhood adiposity and risk of type 1 diabetes: A Mendelian randomization study. *PLoS Med.* 14, e1002362. doi:10.1371/journal.pmed.1002362

- Censin, J. C., Peters, S. A. E., Bovijn, J., Ferreira, T., Pulit, S. L., Magi, R., et al. (2019). Causal relationships between obesity and the leading causes of death in women and men. *PLoS Genet.* 15, e1008405. doi:10.1371/journal.pgen.1008405
- Chatterjee, N. A., Giulianini, F., Geelhoed, B., Lunetta, K. L., Misialek, J. R., Niemeijer, M. N., et al. (2017). Genetic Obesity and the Risk of Atrial Fibrillation Causal Estimates from Mendelian Randomization. *Circulation* 135, 741-754. doi:10.1161/circulationaha.116.024921
- Chen, Y.-C., Fan, H.-Y., Huang, Y.-T., Huang, S.-Y., Liou, T.-H., and Lee, Y. L. (2019). Causal relationships between adiposity and childhood asthma: bi-directional Mendelian Randomization analysis. *Int. J. Obes. (Lond.)* 43, 73-81. doi:10.1038/s41366-018-0160-8
- Chen, Y.-C., Fan, H.-Y., Yang, C., Hsieh, R.-H., Pan, W.-H., and Lee, Y. L. (2019). Assessing causality between childhood adiposity and early puberty: A bidirectional Mendelian randomization and longitudinal study. *Metabolism* 100, 153961. doi:10.1016/j.metabol.2019.153961
- Colak, Y., Afzal, S., Lange, P., and Nordestgaard, B. G. (2016). Obese individuals experience wheezing without asthma but not asthma without wheezing: a Mendelian randomisation study of 85 437 adults from the Copenhagen General Population Study. *Thorax* 71, 247-254. doi:10.1136/thoraxjnl-2015-207379
- Cole, C. B., Nikpay, M., Stewart, A. F., and McPherson, R. (2016). Increased genetic risk for obesity in premature coronary artery disease. *Eur. J. Hum. Genet.* 24, 587-591. doi:10.1038/ejhg.2015.162
- Corbin, L. J., Richmond, R. C., Wade, K. H., Burgess, S., Bowden, J., Smith, G. D., et al. (2016). BMI as a Modifiable Risk Factor for Type 2 Diabetes: Refining and Understanding Causal Estimates Using Mendelian Randomization. *Diabetes* 65, 3002-3007. doi:10.2337/db16-0418
- Dale, C. E., Fatemifar, G., Palmer, T. M., White, J., Prieto-Merino, D., Zabaneh, D., et al. (2017). Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. *Circulation* 135, 2373-2388. doi:10.1161/CIRCULATIONAHA.116.026560

- Davies, N. M., Gaunt, T. R., Lewis, S. J., Holly, J., Donovan, J. L., Hamdy, F. C., et al. (2015). The effects of height and BMI on prostate cancer incidence and mortality: a Mendelian randomization study in 20,848 cases and 20,214 controls from the PRACTICAL consortium. *Cancer Causes Control* 26, 1603-1616. doi:10.1007/s10552-015-0654-9
- den Hollander, W. J., Broer, L., Schurmann, C., Meyre, D., den Hoed, C. M., Mayerle, J., et al. (2017). Helicobacter pylori colonization and obesity - a Mendelian randomization study. *Sci. Rep.* 7, 14467. doi:10.1038/s41598-017-14106-4
- Dixon, S. C., Nagle, C. M., Thrift, A. P., Pharoah, P. D., Pearce, C. L., Zheng, W., et al. (2016). Adult body mass index and risk of ovarian cancer by subtype: a Mendelian randomization study. *Int. J. Epidemiol.* 45, 884-895. doi:10.1093/ije/dyw158
- Dusingize, J. C., Olsen, C. M., An, J., Pandeya, N., Law, M. H., Thompson, B. S., et al. (2020). Body mass index and height and risk of cutaneous melanoma: Mendelian randomization analyses. *Int. J. Epidemiol.* 49, 1236-1245. doi:10.1093/ije/dyaa009
- Eriksson, J., Haring, R., Grarup, N., Vandenput, L., Wallaschofski, H., Lorentzen, E., et al. (2017). Causal relationship between obesity and serum testosterone status in men: A bidirectional mendelian randomization analysis. *PLoS One* 12, e0176277. doi:10.1371/journal.pone.0176277
- Fall, T., Hagg, S., Magi, R., Ploner, A., Fischer, K., Horikoshi, M., et al. (2013). The Role of Adiposity in Cardiometabolic Traits: A Mendelian Randomization Analysis. *PLoS Med.* 10, e1001474. doi:10.1371/journal.pmed.1001474
- Fall, T., Hagg, S., Ploner, A., Magi, R., Fischer, K., Draisma, H. H., et al. (2015). Age- and sex-specific causal effects of adiposity on cardiovascular risk factors. *Diabetes* 64, 1841-1852. doi:10.2337/db14-0988
- Fan, H. Y., Huang, Y. T., Hsieh, R. H., Chao, J. C. J., Tung, Y. C., Lee, Y. L., et al. (2018). Birthweight, time-varying adiposity growth and early menarche in girls: A Mendelian randomisation and mediation analysis. *Obes. Res. Clin. Pract.* 12, 445-451. doi:10.1016/j.orcp.2018.07.008

- Fussey, J. M., Beaumont, R. N., Wood, A. R., Vaidya, B., Smith, J., and Tyrrell, J. (2020). Does obesity cause thyroid cancer? A Mendelian randomization study. *J. Clin. Endocrinol. Metab.* 105, e2398-e2407. doi:10.1210/clinem/dgaa250
- Gao, C., Patel, C. J., Michailidou, K., Peters, U., Gong, J., Schildkraut, J., et al. (2016). Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Int. J. Epidemiol.* 45, 896-908. doi:10.1093/ije/dyw129
- Geng, T., Smith, C. E., Li, C., and Huang, T. (2018). Childhood BMI and adult type 2 diabetes, coronary artery diseases, chronic kidney disease, and cardiometabolic traits: A Mendelian randomization analysis. *Diabetes Care* 41, 1089-1096. doi:10.2337/dc17-2141/-/DC1
- Gharahkhani, P., Ong, J. S., An, J. Y., Law, M. H., Whiteman, D. C., Neale, R. E., et al. (2019). Effect of increased body mass index on risk of diagnosis or death from cancer. *Br. J. Cancer* 120, 565-570. doi:10.1038/s41416-019-0386-9
- Giambartolomei, C., Vukcevic, D., Schadt, E. E., Franke, L., Hingorani, A. D., Wallace, C., et al. (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* 10, e1004383. doi:10.1371/journal.pgen.1004383
- Gianfrancesco, M. A., Glymour, M. M., Walter, S., Rhead, B., Shao, X., Shen, L., et al. (2017). Causal Effect of Genetic Variants Associated With Body Mass Index on Multiple Sclerosis Susceptibility. *Am. J. Epidemiol.* 185, 162-171. doi:10.1093/aje/kww120
- Gianfrancesco, M. A., Stridh, P., Rhead, B., Shao, X. R., Xu, E., Graves, J. S., et al. (2017). Evidence for a causal relationship between low vitamin D, high BMI, and pediatric-onset MS. *Neurology* 88, 1623-1629. doi:10.1212/WNL.0000000000003849
- Glymour, M. M., Tchetgen Tchetgen, E. J., and Robins, J. M. (2012). Credible mendelian randomization studies: Approaches for evaluating the instrumental variable assumptions. *Am. J. Epidemiol.* 175, 332-339. doi:10.1093/aje/kwr323
- Granel, R., Henderson, A. J., Evans, D. M., Smith, G. D., Ness, A. R., Lewis, S., et al. (2014). Effects of BMI, fat mass, and lean mass on asthma in childhood: a Mendelian randomization study. *PLoS Med.* 11, e1001669. doi:10.1371/journal.pmed.1001669

- Guo, Q., Burgess, S., Turman, C., Bolla, M. K., Wang, Q., Lush, M., et al. (2017). Body mass index and breast cancer survival: a Mendelian randomization analysis. *Int. J. Epidemiol.* 46, 1814-1822. doi:10.1093/ije/dyx131
- Guo, Y., Andersen, S. W., Shu, X. O., Michailidou, K., Bolla, M. K., Wang, Q., et al. (2016). Genetically Predicted Body Mass Index and Breast Cancer Risk: Mendelian Randomization Analyses of Data from 145,000 Women of European Descent. *PLoS Med.* 13, e1002105. doi:10.1371/journal.pmed.1002105
- Hagg, S., Fall, T., Ploner, A., Magi, R., Fischer, K., Draisma, H. H. M., et al. (2015). Adiposity as a cause of cardiovascular disease: A Mendelian randomization study. *Int. J. Epidemiol.* 44, 578-586. doi:10.1093/ije/dyv094
- Hartwig, F. P., Bowden, J., de Mola, C. L., Tovo-Rodrigues, L., Smith, G. D., and Horta, B. L. (2016). Body mass index and psychiatric disorders: a Mendelian randomization study. *Sci. Rep.* 6, 32730. doi:10.1038/srep32730
- Hartwig, F. P., Davey Smith, G., and Bowden, J. (2017). Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* 46, 1985-1998. doi:10.1093/ije/dyx102
- Hemani, G., Bowden, J., and Davey Smith, G. (2018). Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* 27, R195-R208. doi:10.1093/hmg/ddy163
- Holmes, M. V., Lange, L. A., Palmer, T., Lanktree, M. B., North, K. E., Almqvera, B., et al. (2014). Causal effects of body mass index on cardiometabolic traits and events: a Mendelian randomization analysis. *Am. J. Hum. Genet.* 94, 198-208. doi:10.1016/j.ajhg.2013.12.014
- Howe, L. D., Kanayalal, R., Harrison, S., Beaumont, R. N., Davies, A. R., Frayling, T. M., et al. (2019). Effects of body mass index on relationship status, social contact and socio-economic position: Mendelian randomization and within-sibling study in UK Biobank. *Int. J. Epidemiol.* 49, 1173-1184. doi:10.1093/ije/dyz240

- Howe, L. J., Trela-Larsen, L., Taylor, M., Heron, J., Munafò, M. R., and Taylor, A. E. (2017). Body mass index, body dissatisfaction and adolescent smoking initiation. *Drug Alcohol Depend.* 178, 143-149. doi:10.1016/j.drugalcdep.2017.04.008
- Huang, Y., Xu, M., Xie, L., Wang, T., Huang, X., Lv, X., et al. (2016). Obesity and peripheral arterial disease: A Mendelian Randomization analysis. *Atherosclerosis* 247, 218-224. doi:10.1016/j.atherosclerosis.2015.12.034
- Hung, C. F., Rivera, M., Craddock, N., Owen, M. J., Gill, M., Korszun, A., et al. (2014). Relationship between obesity and the risk of clinically significant depression: Mendelian randomisation study. *Br. J. Psychiatry* 205, 24-28. doi:10.1192/bjp.bp.113.130419
- Jacobs, B. M., Noyce, A. J., Giovannoni, G., and Dobson, R. (2020). BMI and low vitamin D are causal factors for multiple sclerosis: A Mendelian Randomization study. *Neurol. Neuroimmunol. Neuroinflamm.* 7, e662. doi:10.1212/NXI.0000000000000662
- Jarvis, D., Mitchell, J. S., Law, P. J., Palin, K., Tuupainen, S., Gylfe, A., et al. (2016). Mendelian randomisation analysis strongly implicates adiposity with risk of developing colorectal cancer. *Br. J. Cancer* 115, 266-272. doi:10.1038/bjc.2016.188
- Jokela, M., Elovainio, M., Keltikangas-Järvinen, L., Batty, G. D., Hintsanen, M., Seppälä, I., et al. (2012). Body mass index and depressive symptoms: instrumental-variables regression with genetic risk score. *Genes Brain Behav.* 11, 942-948. doi:10.1111/j.1601-183X.2012.00846.x
- Kaltoft, M., Langsted, A., and Nordestgaard, B. G. (2020). Obesity as a Causal Risk Factor for Aortic Valve Stenosis. *J. Am. Coll. Cardiol.* 75, 163-176. doi:10.1016/j.jacc.2019.10.050
- Kemp, J. P., Sayers, A., Davey Smith, G., Tobias, J. H., and Evans, D. M. (2016). Using Mendelian randomization to investigate a possible causal relationship between adiposity and increased bone mineral density at different skeletal sites in children. *Int. J. Epidemiol.* 45, 1560-1572. doi:10.1093/ije/dyw079
- Kivimäki, M., Jokela, M., Hamer, M., Geddes, J., Ebmeier, K., Kumari, M., et al. (2011). Examining overweight and obesity as risk factors for common mental disorders using fat mass and obesity-associated (FTO) genotype-instrumented analysis: The Whitehall II Study, 1985-2004. *Am. J. Epidemiol.* 173, 421-429. doi:10.1093/aje/kwq444

- Kivimaki, M., Smith, G. D., Timpson, N. J., Lawlor, D. A., Batty, G. D., Kahonen, M., et al. (2008). Lifetime body mass index and later atherosclerosis risk in young adults: examining causal links using Mendelian randomization in the Cardiovascular Risk in Young Finns study. *Eur. Heart J.* 29, 2552-2560. doi:10.1093/eurheartj/ehn252
- Klovaite, J., Benn, M., and Nordestgaard, B. G. (2015). Obesity as a causal risk factor for deep venous thrombosis: a Mendelian randomization study. *J. Intern. Med.* 277, 573-584. doi:10.1111/joim.12299
- Kurz, C. F., and Laxy, M. (2020). Application of Mendelian Randomization to Investigate the Association of Body Mass Index with Health Care Costs. *Med. Decis. Making* 40, 156-169. doi:10.1177/0272989X20905809
- Larsson, S. C., Back, M., Rees, J. M. B., Mason, A. M., and Burgess, S. (2020). Body mass index and body composition in relation to 14 cardiovascular conditions in UK Biobank: a Mendelian randomization study. *Eur. Heart J.* 41, 221-226. doi:10.1093/eurheartj/ehz388
- Larsson, S. C., Burgess, S., and Michaëlsson, K. (2018). Genetic association between adiposity and gout: a Mendelian randomization study. *Rheumatology* 57, 2145-2148. doi:10.1093/rheumatology/key229
- Lawlor, D. A., Harbord, R. M., Tybjaerg-Hansen, A., Palmer, T. M., Zacho, J., Benn, M., et al. (2011). Using genetic loci to understand the relationship between adiposity and psychological distress: A Mendelian Randomization study in the Copenhagen General Population Study of 53221 adults. *J. Intern. Med.* 269, 525-537. doi:10.1111/j.1365-2796.2011.02343.x
- Lee, M. R., Lim, Y. H., and Hong, Y. C. (2018). Causal association of body mass index with hypertension using a Mendelian randomization design. *Medicine (Baltimore)* 97, e11252. doi:10.1097/MD.00000000000011252
- Lee, S. J., Lee, J. Y., and Sung, J. (2019). Obesity and Bone Health Revisited: A Mendelian Randomization Study for Koreans. *J. Bone Miner. Res.* 34, 1058-1067. doi:10.1002/jbmr.3678

- Lewis, S. J., Murad, A., Chen, L. N., Davey Smith, G., Donovan, J., Palmer, T., et al. (2010). Associations between an Obesity Related Genetic Variant (FTO rs9939609) and Prostate Cancer Risk. *PLoS One* 5, e13485. doi:10.1371/journal.pone.0013485
- Lim, L. S., Tai, E. S., Aung, T., Tay, W. T., Saw, S. M., Seielstad, M., et al. (2009). Relation of age-related cataract with obesity and obesity genes in an Asian population. *Am. J. Epidemiol.* 169, 1267-1274. doi:10.1093/aje/kwp045
- Lindstrom, S., Germain, M., Crous-Bou, M., Smith, E. N., Morange, P. E., Vlieg, A. V., et al. (2017). Assessing the causal relationship between obesity and venous thromboembolism through a Mendelian Randomization study. *Hum. Genet.* 136, 897-902. doi:10.1007/s00439-017-1811-x
- Liu, Q., Pan, J., Berzuini, C., Rutter, M. K., and Guo, H. (2020). Integrative analysis of Mendelian randomization and Bayesian colocalization highlights four genes with putative BMI-mediated causal pathways to diabetes. *Sci. Rep.* 10, 7476. doi:10.1038/s41598-020-64493-4
- Lyall, D. M., Celis-Morales, C., Ward, J., Iliodromiti, S., Anderson, J. J., Gill, J. M. R., et al. (2017). Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank A Mendelian Randomization Study. *JAMA Cardiol.* 2, 882-889. doi:10.1001/jamacardio.2016.5804
- Lyngdoh, T., Vuistiner, P., Marques-Vidal, P., Rousson, V., Waeber, G., Vollenweider, P., et al. (2012). Serum uric acid and adiposity: deciphering causality using a bidirectional Mendelian randomization approach. *PLoS One* 7, e39321. doi:10.1371/journal.pone.0039321
- Mao, Y., Yan, C., Lu, Q., Zhu, M., Yu, F., Wang, C., et al. (2017). Genetically predicted high body mass index is associated with increased gastric cancer risk. *Eur. J. Hum. Genet.* 25, 1061-1066. doi:10.1038/ejhg.2017.103
- Marini, S., Merino, J., Montgomery, B. E., Malik, R., Sudlow, C. L., Dichgans, M., et al. (2020). Mendelian Randomization Study of Obesity and Cerebrovascular Disease. *Ann. Neurol.* 87, 516-524. doi:10.1002/ana.25686

- Martins-Silva, T., Vaz, J. S., Hutz, M. H., Salatino-Oliveira, A., Genro, J. P., Hartwig, F. P., et al. (2019). Assessing causality in the association between attention-deficit/hyperactivity disorder and obesity: a Mendelian randomization study. *Int. J. Obes. (Lond.)* 43, 2500-2508. doi:10.1038/s41366-019-0346-8
- Mokry, L. E., Ross, S., Timpson, N. J., Sawcer, S., Smith, G. D., and Richards, J. B. (2016). Obesity and Multiple Sclerosis: A Mendelian Randomization Study. *PLoS Med.* 13, e1002053. doi:10.1371/journal.pmed.1002053
- Mukherjee, S., Walter, S., Kauwe, J. S. K., Saykin, A. J., Bennett, D. A., Larson, E. B., et al. (2015). Genetically predicted body mass index and Alzheimer's disease-related phenotypes in three large samples: Mendelian randomization analyses. *Alzheimers Dement.* 11, 1439-1451. doi:10.1016/j.jalz.2015.05.015
- Mumby, H. S., Elks, C. E., Li, S., Sharp, S. J., Khaw, K. T., Luben, R. N., et al. (2011). Mendelian Randomisation Study of Childhood BMI and Early Menarche. *J. Obes.* 2011, 180729. doi:10.1155/2011/180729
- Nordestgaard, B. G., Palmer, T. M., Benn, M., Zacho, J., Tybaerg-Hansen, A., Smith, G. D., et al. (2012). The Effect of Elevated Body Mass Index on Ischemic Heart Disease Risk: Causal Estimates from a Mendelian Randomisation Approach. *PLoS Med.* 9, e1001212. doi:10.1371/journal.pmed.1001212
- Nordestgaard, L. T., Tybjaerg-Hansen, A., Nordestgaard, B. G., and Frikke-Schmidt, R. (2017). Body Mass Index and Risk of Alzheimer's Disease: A Mendelian Randomization Study of 399,536 Individuals. *J. Clin. Endocrinol. Metab.* 102, 2310-2320. doi:10.1210/jc.2017-00195
- Noyce, A. J., Kia, D. A., Hemani, G., Nicolas, A., Price, T. R., De Pablo-Fernandez, E., et al. (2017). Estimating the causal influence of body mass index on risk of Parkinson disease: A Mendelian randomisation study. *PLoS Med.* 14, e1002314. doi:10.1371/journal.pmed.1002314
- Ooi, B. N. S., Loh, H., Ho, P. J., Milne, R. L., Giles, G., Gao, C., et al. (2019). The genetic interplay between body mass index, breast size and breast cancer risk: a Mendelian randomization analysis. *Int. J. Epidemiol.* 48, 781-794. doi:10.1093/ije/dyz124

- Painter, J. N., O'Mara, T. A., Marquart, L., Webb, P. M., Attia, J., Medland, S. E., et al. (2016). Genetic Risk Score Mendelian Randomization Shows that Obesity Measured as Body Mass Index, but not Waist:Hip Ratio, Is Causal for Endometrial Cancer. *Cancer Epidemiol. Biomarkers Prev.* 25, 1503-1510. doi:10.1158/1055-9965.EPI-16-0147
- Panoutsopoulou, K., Metrustry, S., Doherty, S. A., Laslett, L. L., Maciewicz, R. A., Hart, D. J., et al. (2014). The effect of FTO variation on increased osteoarthritis risk is mediated through body mass index: a Mendelian randomisation study. *Ann. Rheum. Dis.* 73, 2082-2086. doi:10.1136/annrheumdis-2013-203772
- Qian, F., Rookus, M. A., Leslie, G., Risch, H. A., Greene, M. H., Aalfs, C. M., et al. (2019). Mendelian randomisation study of height and body mass index as modifiers of ovarian cancer risk in 22,588 BRCA1 and BRCA2 mutation carriers. *Br. J. Cancer* 121, 180-192. doi:10.1038/s41416-019-0492-8
- Qian, F., Wang, S. F., Mitchell, J., McGuffog, L., Barrowdale, D., Leslie, G., et al. (2019). Height and Body Mass Index as Modifiers of Breast Cancer Risk in BRCA1/2 Mutation Carriers: A Mendelian Randomization Study. *J. Natl. Cancer Inst.* 111, 350-364. doi:10.1093/jnci/djy132
- Reed, Z. E., Micali, N., Bulik, C. M., Davey Smith, G., and Wade, K. H. (2017). Assessing the causal role of adiposity on disordered eating in childhood, adolescence, and adulthood: a Mendelian randomization analysis. *Am. J. Clin. Nutr.* 106, 764-772. doi:10.3945/ajcn.117.154104
- Reed, Z. E., Suderman, M. J., Relton, C. L., Davis, O. S. P., and Hemani, G. (2020). The association of DNA methylation with body mass index: distinguishing between predictors and biomarkers. *Clin. Epigenetics* 12, 50. doi:10.1186/s13148-020-00841-5
- Richardson, T. G., Sanderson, E., Elsworth, B., Tilling, K., and Davey Smith, G. (2020). Use of genetic variation to separate the effects of early and later life adiposity on disease risk: mendelian randomisation study. *Br. Med. J.* 369, m1203. doi:10.1136/bmj.m1203
- Richmond, R. C., Davey Smith, G., Ness, A. R., den Hoed, M., McMahon, G., and Timpson, N. J. (2014). Assessing Causality in the Association between Child Adiposity and Physical

- Activity Levels: A Mendelian Randomization Analysis. *PLoS Med.* 11, e1001618. doi:10.1371/journal.pmed.1001618
- Robinson, J. R., Carroll, R. J., Bastarache, L., Chen, Q., Mou, Z., Wei, W. Q., et al. (2020). Association of Genetic Risk of Obesity with Postoperative Complications Using Mendelian Randomization. *World J. Surg.* 44, 84-94. doi:10.1007/s00268-019-05202-9
- Rode, L., Nordestgaard, B. G., Weischer, M., and Bojesen, S. E. (2014). Increased Body Mass Index, Elevated C-reactive Protein, and Short Telomere Length. *J. Clin. Endocrinol. Metab.* 99, E1671-E1675. doi:10.1210/jc.2014-1161
- Sanderson, E., Davey Smith, G., Windmeijer, F., and Bowden, J. (2019). An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int. J. Epidemiol.* 48, 713-727. doi:10.1093/ije/dyy262
- Schnurr, T. M., Viitasalo, A., Eloranta, A. M., Damsgaard, C. T., Mahendran, Y., Have, C. T., et al. (2018). Genetic predisposition to adiposity is associated with increased objectively assessed sedentary time in young children. *Int. J. Obes. (Lond.)* 42, 111-114. doi:10.1038/ijo.2017.235
- Shu, X., Wu, L., Khankari, N. K., Shu, X. O., Wang, T. J., Michailidou, K., et al. (2019). Associations of obesity and circulating insulin and glucose with breast cancer risk: a Mendelian randomization analysis. *Int. J. Epidemiol.* 48, 795-806. doi:10.1093/ije/dyy201
- Shungin, D., Cornelis, M. C., Divaris, K., Holtfreter, B., Shaffer, J. R., Yu, Y. H., et al. (2015). Using genetics to test the causal relationship of total adiposity and periodontitis: Mendelian randomization analyses in the Gene-Lifestyle Interactions and Dental Endpoints (GLIDE) Consortium. *Int. J. Epidemiol.* 44, 638-650. doi:10.1093/ije/dyv075
- Skaaby, T., Taylor, A. E., Thuesen, B. H., Jacobsen, R. K., Friedrich, N., Møllehave, L. T., et al. (2018). Estimating the causal effect of body mass index on hay fever, asthma and lung function using Mendelian randomization. *Allergy* 73, 153-164. doi:10.1111/all.13242
- Slob, E. A., and Burgess, S. (2020). A comparison of robust Mendelian randomization methods using summary data. *Genet. Epidemiol.* 44, 313-329. doi:10.1002/gepi.22295

- Speed, M. S., Jepsen, O. H., Borglum, A. D., Speed, D., and Ostergaard, S. D. (2019). Investigating the association between body fat and depression via Mendelian randomization. *Transl. Psychiatry* 9, 184. doi:10.1038/s41398-019-0516-4
- Stender, S., Nordestgaard, B. G., and Tybjaerg-Hansen, A. (2013). Elevated body mass index as a causal risk factor for symptomatic gallstone disease: a Mendelian randomization study. *Hepatology* 58, 2133-2141. doi:10.1002/hep.26563
- Sun, Y. Q., Brumpton, B., Langhammer, A., Chen, Y., Kvaloy, K., and Mai, X. M. (2020). Adiposity and asthma in adults: a bidirectional Mendelian randomisation analysis of The HUNT Study. *Thorax* 75, 202-208. doi:10.1136/thoraxjnl-2019-213678
- Sun, Y. Q., Burgess, S., Staley, J. R., Wood, A. M., Bell, S., Kaptoge, S. K., et al. (2019). Body mass index and all cause mortality in HUNT and UK Biobank studies: linear and non-linear mendelian randomisation analyses. *Br. Med. J.* 364, 11042. doi:10.1136/bmj.11042
- Takahashi, H., Cornish, A. J., Sud, A., Law, P. J., Disney-Hogg, L., Calvocoressi, L., et al. (2019). Mendelian randomization provides support for obesity as a risk factor for meningioma. *Sci. Rep.* 9, 309. doi:10.1038/s41598-018-36186-6
- Tan, A. G., Kifley, A., Flood, V. M., Holliday, E. G., Scott, R. J., Cumming, R. G., et al. (2019). Evaluating the associations between obesity and age-related cataract: a Mendelian randomization study. *Am. J. Clin. Nutr.* 110, 969-976. doi:10.1093/ajcn/nqz167
- Taylor, A. E., Richmond, R. C., Palviainen, T., Loukola, A., Wootton, R. E., Kaprio, J., et al. (2019). The effect of body mass index on smoking behaviour and nicotine metabolism: a Mendelian randomization study. *Hum. Mol. Genet.* 28, 1322-1330. doi:10.1093/hmg/ddy434
- Taylor, P. N., Richmond, R., Davies, N., Sayers, A., Stevenson, K., Woltersdorf, W., et al. (2016). Paradoxical Relationship Between Body Mass Index and Thyroid Hormone Levels: A Study Using Mendelian Randomization. *J. Clin. Endocrinol. Metab.* 101, 730-738. doi:10.1210/jc.2015-3505
- Thakkinstian, A., Chailurkit, L., Warodomwichit, D., Ratanachaiwong, W., Yamwong, S., Chanprasertyothin, S., et al. (2014). Causal relationship between body mass index and

- fetuin-A level in the asian population: a bidirectional mendelian randomization study. *Clin. Endocrinol. (Oxf.)* 81, 197-203. doi:10.1111/cen.12303
- Thrift, A. P., Gong, J., Peters, U., Chang-Claude, J., Rudolph, A., Slattery, M. L., et al. (2015). Mendelian Randomization Study of Body Mass Index and Colorectal Cancer Risk. *Cancer Epidemiol. Biomarkers Prev.* 24, 1024-1031. doi:10.1158/1055-9965.EPI-14-1309
- Thrift, A. P., Shaheen, N. J., Gammon, M. D., Bernstein, L., Reid, B. J., Onstad, L., et al. (2014). Obesity and Risk of Esophageal Adenocarcinoma and Barrett's Esophagus: A Mendelian Randomization Study. *J. Natl. Cancer Inst.* 106, dju252. doi:10.1093/jnci/dju252
- Timpson, N. J., Harbord, R., Davey Smith, G., Zacho, J., Tybjaerg-Hansen, A., and Nordestgaard, B. G. (2009). Does greater adiposity increase blood pressure and hypertension risk?: Mendelian randomization using the FTO/MC4R genotype. *Hypertension* 54, 84-90. doi:10.1161/HYPERTENSIONAHA.109.130005
- Timpson, N. J., Nordestgaard, B. G., Harbord, R. M., Zacho, J., Frayling, T. M., Tybjaerg-Hansen, A., et al. (2011). C-reactive protein levels and body mass index: elucidating direction of causation through reciprocal Mendelian randomization. *Int. J. Obes. (Lond.)* 35, 300-308. doi:10.1038/ijo.2010.137
- Todd, J. N., Dahlström, E. H., Salem, R. M., Sandholm, N., Forsblom, C., McKnight, A. J., et al. (2015). Genetic Evidence for a Causal Role of Obesity in Diabetic Kidney Disease. *Diabetes* 64, 4238-4246. doi:10.2337/db15-0254
- Tyrrell, J., Jones, S. E., Beaumont, R., Astley, C. M., Lovell, R., Yaghootkar, H., et al. (2016). Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. *Br. Med. J.* 352, i582. doi:10.1136/bmj.i582
- Tyrrell, J., Mulugeta, A., Wood, A. R., Zhou, A., Beaumont, R. N., Tuke, M. A., et al. (2019). Using genetics to understand the causal influence of higher BMI on depression. *Int. J. Epidemiol.* 48, 834-848. doi:10.1093/ije/dyy223
- van 't Hof, F. N. G., Vaucher, J., Holmes, M. V., de Wilde, A., Baas, A. F., Blankensteijn, J. D., et al. (2017). Genetic variants associated with type 2 diabetes and adiposity and risk of intracranial and abdominal aortic aneurysms. *Eur. J. Hum. Genet.* 25, 758-762. doi:10.1038/ejhg.2017.48

- van den Broek, N., Treur, J. L., Larsen, J. K., Verhagen, M., Verweij, K. J. H., and Vink, J. M. (2018). Causal associations between body mass index and mental health: a Mendelian randomisation study. *J. Epidemiol. Community Health* 72, 708-710. doi:10.1136/jech-2017-210000
- VanderWeele, T. J. (2016). Mediation analysis: a practitioner's guide. *Annu. Rev. Public Health* 37, 17-32. doi:10.1146/annurev-publhealth-032315-021402
- Varbo, A., Benn, M., Smith, G. D., Timpson, N. J., Tybjaerg-Hansen, A., and Nordestgaard, B. G. (2015). Remnant cholesterol, low-density lipoprotein cholesterol, and blood pressure as mediators from obesity to ischemic heart disease. *Circ. Res.* 116, 665-673. doi:10.1161/CIRCRESAHA.116.304846
- Vasan, S. K., Noordam, R., Gowri, M. S., Neville, M. J., Karpe, F., and Christodoulides, C. (2019). The proposed systemic thermogenic metabolites succinate and 12,13-diHOME are inversely associated with adiposity and related metabolic traits: evidence from a large human cross-sectional study. *Diabetologia* 62, 2079-2087. doi:10.1007/s00125-019-4947-5
- Verbanck, M., Chen, C.-Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* 50, 693-698. doi:10.1038/s41588-018-0099-7
- Vimalaswaran, K. S., Berry, D. J., Lu, C., Tikkanen, E., Pilz, S., Hiraki, L. T., et al. (2013). Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts. *PLoS Med.* 10, e1001383. doi:10.1371/journal.pmed.1001383
- Wade, K. H., Carlslake, D., Sattar, N., Davey Smith, G., and Timpson, N. J. (2018). BMI and Mortality in UK Biobank: Revised Estimates Using Mendelian Randomization. *Obesity (Silver Spring, Md.)* 26, 1796-1806. doi:10.1002/oby.22313
- Wade, K. H., Chiesa, S. T., Hughes, A. D., Chaturvedi, N., Charakida, M., Rapala, A., et al. (2018). Assessing the causal role of body mass index on cardiovascular health in young adults: Mendelian randomization and recall-by-genotype analyses. *Circulation* 138, 2187-2201. doi:10.1161/CIRCULATIONAHA.117.033278

- Wainberg, M., Mahajan, A., Kundaje, A., McCarthy, M. I., Ingelsson, E., Sinnott-Armstrong, N., et al. (2019). Homogeneity in the association of body mass index with type 2 diabetes across the UK Biobank: A Mendelian randomization study. *PLoS Med.* 16, e1002982. doi:10.1371/journal.pmed.1002982
- Walter, S., Glymour, M. M., Koenen, K., Liang, L., Tchetgen Tchetgen, E. J., Cornelis, M., et al. (2015). Do genetic risk scores for body mass index predict risk of phobic anxiety? Evidence for a shared genetic risk factor. *Psychol. Med.* 45, 181-191. doi:10.1017/S0033291714001226
- Walter, S., Kubzansky, L. D., Koenen, K. C., Liang, L., Tchetgen Tchetgen, E. J., Cornelis, M. C., et al. (2015). Revisiting Mendelian randomization studies of the effect of body mass index on depression. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 168B, 108-115. doi:10.1002/ajmg.b.32286
- Wang, N., Cheng, J., Ning, Z., Chen, Y., Han, B., Li, Q., et al. (2018). Type 2 Diabetes and Adiposity Induce Different Lipid Profile Disorders: A Mendelian Randomization Analysis. *J. Clin. Endocrinol. Metab.* 103, 2016-2025. doi:10.1210/jc.2017-02789
- Wang, N., Lu, M., Chen, C., Xia, F., Han, B., Li, Q., et al. (2018). Adiposity genetic risk score modifies the association between blood lead level and body mass index. *J. Clin. Endocrinol. Metab.* 103, 4005-4013. doi:10.1210/jc.2018-00472
- Wang, T., Ma, X., Tang, T., Jin, L., Peng, D., Zhang, R., et al. (2016). Overall and central obesity with insulin sensitivity and secretion in a Han Chinese population: a Mendelian randomization analysis. *Int. J. Obes. (Lond.)* 40, 1736-1741. doi:10.1038/ijo.2016.155
- Wang, T., Zhang, R., Ma, X. J., Wang, S. Y., He, Z., Huang, Y. P., et al. (2018). Causal Association of Overall Obesity and Abdominal Obesity with Type 2 Diabetes: A Mendelian Randomization Analysis. *Obesity (Silver Spring, Md.)* 26, 934-942. doi:10.1002/oby.22167
- Warodomwicht, D., Sritara, C., Thakkinstian, A., Chailurkit, L. O., Yamwong, S., Ratanachaiwong, W., et al. (2013). Causal inference of the effect of adiposity on bone mineral density in adults. *Clin. Endocrinol. (Oxf.)* 78, 694-699. doi:10.1111/cen.12061
- Welsh, P., Polisecki, E., Robertson, M., Jahn, S., Buckley, B. M., de Craen, A. J., et al. (2010). Unraveling the directional link between adiposity and inflammation: a bidirectional

- Mendelian randomization approach. *J. Clin. Endocrinol. Metab.* 95, 93-99. doi:10.1210/jc.2009-1064
- Winter-Jensen, M., Afzal, S., Jess, T., Nordestgaard, B. G., and Allin, K. H. (2020). Body mass index and risk of infections: a Mendelian randomization study of 101,447 individuals. *Eur. J. Epidemiol.* 35, 347-354. doi:10.1007/s10654-020-00630-7
- Wurtz, P., Wang, Q., Kangas, A. J., Richmond, R. C., Skarp, J., Tiainen, M., et al. (2014). Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. *PLoS Med.* 11, e1001765. doi:10.1371/journal.pmed.1001765
- Xu, L., Borges, M. C., Hemani, G., and Lawlor, D. A. (2017). The role of glycaemic and lipid risk factors in mediating the effect of BMI on coronary heart disease: a two-step, two-sample Mendelian randomisation study. *Diabetologia* 60, 2210-2220. doi:10.1007/s00125-017-4396-y
- Xu, S., Gilliland, F. D., and Conti, D. V. (2019). Elucidation of causal direction between asthma and obesity: a bi-directional Mendelian randomization study. *Int. J. Epidemiol.* 48, 899-907. doi:10.1093/ije/dyz070
- Zeng, P., Yu, X., and Xu, H. (2019). Association Between Premorbid Body Mass Index and Amyotrophic Lateral Sclerosis: Causal Inference Through Genetic Approaches. *Front. Neurol.* 10, 543. doi:10.3389/fneur.2019.00543
- Zhang, L., Tang, L., Huang, T., and Fan, D. (2020). Life Course Adiposity and Amyotrophic Lateral Sclerosis: A Mendelian Randomization Study. *Ann. Neurol.* 87, 434-441. doi:10.1002/ana.25671
- Zhao, Q., Wang, J., Hemani, G., Bowden, J., and Small, D. S. (2020). Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *Ann. Stat.* 48, 1742-1769. doi:10.1214/19-AOS1866
- Zhao, Y., Xu, Y., Wang, X., Xu, L., Chen, J., Gao, C., et al. (2020). Body Mass Index and Polycystic Ovary Syndrome: A 2-Sample Bidirectional Mendelian Randomization Study. *J. Clin. Endocrinol. Metab.* 105, dgaa125. doi:10.1210/clinem/dgaa125

Zhou, Y. C., Sun, X. B., and Zhou, M. G. (2019). Body Shape and Alzheimer's Disease: A Mendelian Randomization Analysis. *Front. Neurosci.* 13, 1084. doi:10.3389/fnins.2019.01084

Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M. R., Powell, J. E., et al. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* 48, 481-487. doi:10.1038/ng.3538

4.2. Article 2

Detecting invalid instruments in Mendelian randomization studies: Evaluating the performance of the ‘Leveraging positive confounding’ method for falsification of instrumental variable assumptions

Fiston Ikwa Ndol Mbutiwi^{1,2}

David Boucher-Roy¹

Christian Léger³

Marie-Pierre Sylvestre^{1,4}

¹University of Montreal Hospital Research Centre (CRCHUM), Montréal, Québec, Canada

²Faculty of Medicine, University of Kikwit, Kikwit, Democratic Republic of the Congo

³Department of Mathematics and Statistics, Faculty of Arts and Science, University of Montreal, Montréal, Québec, Canada

⁴Department of Social and Preventive Medicine, University of Montreal Public Health School (ESPUM), Montréal, Québec, Canada

Ce manuscrit est en préparation et sera soumis pour évaluation.

Abstract

Background: Assessing the plausibility of the exclusion restriction assumption is critical in Mendelian randomization (MR) studies. The *leveraging positive confounding* (LPC) method attempts to identify invalid instrumental variables (IVs) by comparing MR and conventional regression point estimates obtained in a given study. This paper evaluates the performance of the LPC method in realistic simulation scenarios and illustrates its application in published MR studies.

Methods: MR studies were conducted on simulated data with a continuous or binary outcome, using either a single-nucleotide polymorphism (SNP) or a genetic risk score (GRS) as IV. Sample sizes ranged from 2,000 to 30,000, and both direct and indirect pleiotropy involving the IVs were considered. In scenarios considering a GRS as the IV, proportions of pleiotropic SNPs of 10%, 30%, and 50% were considered. The performance of the LPC method was assessed as the percentage of simulated datasets in which the LPC method detected an invalid IV. The LPC method was applied to 28 MR analyses from 15 published MR studies and the conclusion of the LPC method compared to that reported in the articles.

Results: When using a single SNP as IV, the LPC method performed better in scenarios with direct than indirect pleiotropy, yet, as the sample size increased, the detection of indirect pleiotropy improved, and the observed false-positive findings declined to zero. Compared with cases with binary outcomes, cases with continuous outcomes were associated with a larger percentage of correctly identified invalid IVs and with a lower percentage of incorrectly identified valid IVs. When using a GRS as IV, the performance of the LPC method increased with the proportion of pleiotropic SNPs in the GRS. The illustration of the use of the LPC method in published articles suggested a moderate consistency with the discussion of pleiotropy in the articles and highlighted the challenges associated with interpreting the results of the LPC method when MR estimates are not sufficiently precise.

Conclusions: The LPC method is helpful in identifying invalid IVs in large MR studies that use a single SNP as the IV. The performance of the LPC method in MR studies that use a GRS as the IV is limited to situations in which many SNPs are invalid IVs.

Keywords: Mendelian Randomization; instrumental variable; IV assumptions; pleiotropy; exclusion restriction, falsification method, net positive confounding.

Introduction

Mendelian randomization (MR) is an approach that uses a genetic variant associated with an exposure (MR assumption 1, *relevance*) as instrumental variable (IV) to estimate the causal effect of this exposure on an outcome, assuming no confounding between the IV and the outcome (MR assumption 2, *independence*) and no effect of the IV on the outcome outside of the exposition (MR assumption 3, *exclusion restriction*).¹⁻⁴ Although the MR assumptions 2 and 3 cannot be formally verified with data,^{1,5} various methods are proposed as sensitivity analyses to assess violations of the MR assumptions and ensuing bias in MR estimates.^{4,6-8}

In 2012, Glymour, Tchetgen Tchetgen and Robins¹ proposed a method aimed at “falsifying IV assumptions by leveraging prior causal assumptions” (referred to as the *leveraging positive confounding* (LPC) method⁴ hereafter) to detect violations of MR assumptions 2 and 3. The method assumes net positive confounding of the association between an exposure and a continuous outcome (i.e. the exposure and the outcome share one common cause that affects each one positively, or several common causes with which the net confounding effect on the association between the exposure and the outcome is positive) to suggest that an IV is invalid if the MR estimate of the association between the exposure and the outcome is larger than the corresponding ordinary least square (OLS) estimate.¹ The intuition provided by Glymour et al.¹ relies on the representation of the MR parameter in the population as the ratio of the total effect of the IV on the outcome over the total effect of the IV on the exposure. Under net positive confounding, the existence of an association between the IV and the outcome outside of the exposure will inflate the numerator and thus the MR estimate. In contrast, the OLS effect that is obtained by regressing the outcome on the exposure does not include the direct portion of the IV-outcome association. Hence, a MR estimate larger than the OLS estimate suggest an association between the IV and the outcome independent of the exposure.¹

The main advantage of this method is to be easily implementable from published regression coefficients, possibly without having access to data. However, while the method has been used in several MR applications,⁹⁻¹² its performance has not been thoroughly investigated. A simulation study from Burgess¹³ suggested that the LPC method is adequate to detect pleiotropy by direct effect of the IV on the outcome (known as direct pleiotropy⁶ or uncorrelated pleiotropy;¹⁴ with a 99.8% success rate), but less satisfactory in scenarios in which the pathways between the IV and

outcome involve the confounder of the exposure-outcome association (also called indirect pleiotropy⁶ or correlated pleiotropy;¹⁴ with a 65% success rate). However, the sample size in the simulation study was modest for MR standards (n=2,000), and it did not consider binary outcomes or genetic risk scores (GRSs) that are commonly used in MR studies.⁸

In the current study, we use simulations to extend the assessment of the performance of the LPC method in a host of realistic MR applications including large sample sizes, binary outcome, and the use of GRS as IV. Finally, we illustrate the use of the method in published MR studies estimating the causal effect of body mass index (BMI) on different outcomes from a recent review,⁸ and discuss the application of the method in practical settings in which sampling variation and finite sample biases for IV estimation complicate the interpretation of the results.

Methods

The LPC method to identify invalid instruments

Original approach

We first briefly describe the application of the LPC method for a single IV and a continuous outcome.¹ Let X be a continuous exposure, G an IV of X , and Y a continuous outcome. The conventional OLS estimate of the association between X and Y ($\hat{\beta}_{CO}$) is obtained by regressing Y on X . The MR estimate ($\hat{\beta}_{MR}$) of the association between X and Y can be obtained using two-stage least squares (2SLS) estimation. The first stage corresponds to the linear regression of X on G , while the second stage regresses Y on the fitted values obtained from the first stage regression. Under the assumption of net positive confounding between X and Y , situations for which $\hat{\beta}_{MR} > \hat{\beta}_{CO}$ suggest that G is an invalid instrument for X in its association with Y .¹

The LPC method does not account for sampling variability, whether for the estimated coefficients or for their difference, and it also does not account for the finite sample bias pertaining to the IV estimation.¹⁵ Consequently, situations for which $\hat{\beta}_{MR}$ and $\hat{\beta}_{CO}$ are close should be interpreted with caution because they could be indicative of imprecise rather than invalid instruments.

Extension of the LPC method to binary outcome

While the LPC method was designed for continuous outcome, it has been applied with binary outcomes using a generalized linear model (GLM) instead of OLS for the conventional analysis and the second stage of the two-stage method in MR.¹⁰ Two alternative IV estimators may be used, including the two-stage predictor substitution (2SPS) which corresponds to the classical 2SLS method by replacing the regression of the second stage by a GLM, and the two-stage residual inclusion (2SRI) or control function that is similar to 2SPS but includes the residuals from the first stage regression as an additional independent variable in the second stage regression.¹⁶⁻¹⁸

Our simulations focus on rare binary outcomes that were modeled using logistic regression for conventional analysis and the 2SRI estimator with logistic regression in the second stage for MR. However, in situations where the outcome is common, log-linear models (e.g., robust Poisson regression) are preferred to logistic models because the odds ratios from logistic regression overestimate the risk ratios.¹⁹⁻²¹

Simulations

We conducted a simulation study to investigate the performance of the LPC method in cases involving (i) continuous vs. binary outcomes (Y), and (ii) a single SNP (G) vs. unweighted GRS aggregating 25 SNPs into a single variable as IV. Each of the four cases included scenarios investigating the impact of sample size, the type of pleiotropy (direct vs. indirect) and in cases in which a GRS was used as the IV, the proportion of pleiotropic SNPs. We designed our scenarios based on previously published simulations studies (i.e., Burgess 2012¹³ for MR with one SNP and Burgess and Thompson (2013) for GRS in MR²²), extending them to consider additional parameter values and scenarios.

We reported the mean F statistic from the first stage model in the 2SLS approach across iterations as a measure of the strength of the IV²³ and calculated the average of the estimated standard errors of the MR (SE_{MR}) and of the conventional (SE_{CO}) estimated coefficients across iterations to assess the precision of both estimations. We evaluate the performance of the LPC method by calculating the percentage of simulated data sets in which this method detects invalid IV, i.e., where $\hat{\beta}_{MR}$ is greater than $\hat{\beta}_{CO}$.

Each of the four cases are described in turn below. **Table S1** summarizes the parameters considered in the different simulation cases. All simulations were performed using the R software versions 4.2.1.

Case 1: single SNP as IV and continuous outcome

We first replicate and extend the simulations from Burgess¹³ that evaluated the performance of the LPC method in a single SNP MR study with a continuous outcome.

Data generation

We generated a continuous exposure X as the linear sum of single binary G (dividing the population into two equal subgroups), unmeasured confounder (U), and error term (\mathcal{E}_X). Specifically, for individual i , $X_i = \beta_{XG}G_i + \beta_{XU}U_i + \mathcal{E}_{Xi}$ with β_{XG} and β_{XU} set to 0.3 and 0.5, respectively. We generated a continuous Y depending on X , U , and \mathcal{E}_Y , i.e., $Y_i = \beta_{YX}X_i + \beta_{YU}U_i + \mathcal{E}_{Yi}$ where β_{YX} is the true causal effect of X on Y , and β_{YU} set to 0.5. In each scenario described below, we considered β_{YX} equal to 0 (null true causal effect), 0.4 (positive true causal effect) and -0.4 (negative true causal effect). We initially generated samples of size $n=2,000$ individuals as in Burgess,¹³ and then considered larger sample sizes, i.e., $n=5,000$, $10,000$, $20,000$, and $30,000$. We simulated 10,000 datasets for each of the three scenarios pertaining to the validity of G as an IV, as described below.

Simulated scenarios

Scenario 1: G as valid IV, so $U_i, \mathcal{E}_{Xi}, \mathcal{E}_{Yi} \sim N(0, 1)$ independently.

Scenario 2: G as invalid IV through association with U (indirect pleiotropy): $U_i|G \sim N(\beta_{UG}G_i, 1)$.

Scenario 3: G as invalid IV through direct association with Y (direct pleiotropy): $Y_i = \beta_{YX}X_i + \beta_{YG}G_i + \beta_{YU}U_i + \mathcal{E}_{Yi}$.

The indirect (β_{UG}) and direct (β_{YG}) pleiotropic effects of G were each set to 0.2.

Case 2: single SNP as IV and binary outcome

We modified the simulation setting in case 1 to accommodate a binary Y . Specifically, we simulated Y from a binomial distribution with each participant's probability of disease (P_i) using the following equations: $\text{logit}(P_i) = \beta_0 + \beta_{YX}X_i + \beta_{YU}U_i$ in the scenarios where the instrument G

was valid or pleiotropic through association with U , and $\text{logit}(P_i) = \beta_0 + \beta_{YX}X_i + \beta_{YU}U_i + \beta_{YG}G_i$ in the last scenario where G was pleiotropic through direct association with Y . The intercept term (β_0) was fixed to achieve a prevalence of the outcome of about 10%.

Case 3: GRS as IV and continuous outcome

To evaluate the performance of the LPC method in the context of an MR study using a GRS as IV, we relied on and extended two of the seven scenarios of simulations from Burgess and Thompson²² that evaluated the use of GRS in MR. The two scenarios were (i) valid instruments with equal-sized effects, and (ii) invalid instruments by association with U . We added a new scenario involving pleiotropic SNPs through direct association with Y . In each of these three scenarios, we considered an unweighted GRS of 25 simulated SNPs.

Data generation

We generated data on 25 independent SNPs G (indexed by j , Hardy-Weinberg equilibrium satisfied, effect allele frequency of 0.3) coded 0, 1, 2 expressing the number of effect allele for each SNP. We generated data on continuous X as the linear sum of G_j , U , and \mathcal{E}_X :

$$X_i = \sum_{j=1}^J \beta_{XG} G_{ij} + \beta_{XU} U_i + \mathcal{E}_{Xi} \text{ where } \beta_{XG} \text{ was set to } 0.06 \text{ for each } G_j, \text{ and } \beta_{XU} \text{ to } 1.$$

Continuous outcomes were generated as in cases 1 and we set β_{YU} to 1. Sample sizes varied from $n=2,000$ to $n=30,000$ and 1,000 iterations were considered for each of the three scenarios pertaining to the validity of the GRS as an IV, as described below.

Simulated scenarios

Scenario 1: GRS as valid IV: U_i , \mathcal{E}_{Xi} , and $\mathcal{E}_{Yi} \sim N(0, 1)$ independently.

Scenario 2: GRS as invalid IV through association with U (indirect pleiotropy):

$$X_i = \sum_{j=1}^J \beta_{XG} Z_j G_{ij} + \beta_{XU} U_i + \mathcal{E}_{Xi}; \text{ and } U_i \sim N\left(\sum_{j=1}^J \beta_{UG}(1 - Z_j) G_{ij}, 1\right)$$

Scenario 3: GRS as invalid IV through direct association with Y (direct pleiotropy):

$$Y_i = \beta_{YX} X_i + \beta_{YU} U_i + \left(\sum_{j=1}^J \beta_{YG}(1 - Z_j) G_{ij} \right) + \mathcal{E}_{Yi}$$

In scenarios 2 and 3, Z_j represent binary random dummy variables that take the value 1 if the SNP G_j is valid and 0 if pleiotropic. The probability of having pleiotropic SNPs was set to 0.1, 0.3, and 0.5, and the coefficient of indirect (β_{UG} , scenario 2) or direct (β_{YG} , scenario 3) pleiotropic effect of an invalid G_j to be equal to β_{XG} .

Case 4: GRS as IV and binary outcome

We modified the simulations setting in case 3 to accommodate a binary outcome Y .

Comparison of the LPC method with alternative approaches in assessing the validity of the instruments in selected MR studies

We illustrated the use of the LPC method in our recent review⁸ of 128 MR studies that used at least one SNP in the fat mass and obesity-associated (*FTO*) gene as IV for BMI. We restricted our application to studies that (i) used one-sample MR, (ii) used single IV (single SNP(s) or GRS) in the main analysis, and (iii) reported the MR and corresponding conventional estimates of the association(s) between BMI and the outcome(s) of interest using the same measure of association. To further limit the number of studies considered, we selected studies that either investigated systolic (SBP) or diastolic blood pressure (DBP), hypertension, or mental health conditions as outcomes. We reported the IV(s) used, the sample size, OLS or GLM estimates, as well as MR estimates, and then applied the LPC method to each study. We then compare the conclusion from the LPC method to the discussion of IV violations based on alternative approaches provided by the authors.

Results

Simulations

Several observations pertaining to the four cases are first made to contextualize the results. First, the mean of the $\hat{\beta}_{CO}$ across simulations are indicative of the important bias observed in the conventional estimates due to the unmeasured confounder U . Second, the performance of the LPC method was similar across values of the true causal effect. Due to space limitations, we present the results when the true causal effect is 0, because the mean of the estimated coefficients across

simulations thus represents the bias. Results considering positive as well as negative causal effect are reported in supplementary material (**Tables S2 and S3**).

Third, the generated IVs were strong as evidenced by the mean first stage F statistic values that were much greater than 10 in all scenarios.²⁴ Fourth, as expected, $\hat{\beta}_{MR}$ was close to the true causal effect in the absence of pleiotropy, while it was biased away from the null in presence of pleiotropy. Finally, the MR estimates were less accurate than the conventional estimates, especially for the binary outcomes, as indicated by the mean SE_{MR} and SE_{CO} reported in **Tables 1 and 2**.

Case 1: single SNP as IV and continuous outcome

The left panel of **Table 1** shows the performance of the LPC method in MR using a single SNP as IV and a continuous outcome. When the IV was valid, the LPC method led to false positive findings (i.e., identifying IV as invalid in the absence of pleiotropy) in less than 3% of simulated datasets, except when the sample size was $n=2,000$. The performance of the LPC method in detecting invalid IV in the presence of indirect pleiotropy increased from 65% when the sample size was $n=2,000$ to 94% when $n=30,000$. Regardless of sample size, the LPC method detected direct pleiotropy in more than 99% of the simulated datasets.

Case 2: single SNP as IV and binary outcome

The right panel of **Table 1** shows the results of simulations with single SNP as IV and binary outcome. Compared the scenario with a continuous outcome, the LPC method led to higher percentages of false positive findings when the outcome was binary, ranging from 35% when $n=2,000$ to 7% when $n=30,000$. The performance of the LPC method improved with increasing sample sizes, with the method correctly detecting invalid IV in the presence of indirect pleiotropy in 55% to 70% of simulated datasets, when the sample size ranged from $n=2,000$ to $n=30,000$, respectively. The corresponding percentages of detection of invalid IV ranged from 82% to 100% in presence of direct pleiotropy.

Case 3: GRS as IV and continuous outcome

The left panel of **Table 2** shows the results of simulations in RM using the GRS as the IV and continuous outcome. Contrary to case 1, the LPC method did not detect any invalid IV in scenario

1 in which all SNPs were valid instruments. The performance of the LPC method was similar in scenarios with indirect and direct pleiotropy: it detected 0%, 2 to 14%, and 48 to 52% of datasets with an invalid IV (GRS) when the percentage of pleiotropic SNPs were 10%, 30%, and 50%, respectively. In addition, the performance of the method decreased as sample size increased when 10% or 30% of SNPs were pleiotropic, while it was constant across sample size when the pleiotropy involved half of the SNPs.

Case 4: GRS as IV and binary outcome

The right panel of **Table 2** shows the results for the binary outcome when using the GRS as the IV. Unlike in case 3 (continuous outcome), the LPC method falsely identified the GRS as an invalid IV in more than 10% of simulated datasets when $n=2,000$ to $< 4\%$ when $n \geq 5,000$. In the presence of pleiotropic IVs, the results for case 4 were similar to those for case 3, except that the performance of the LPC method was better for binary than continuous outcomes when the percentages of pleiotropic SNPs were 10% or 30%.

Comparison of the LPC method with alternative approaches in assessing the validity of the instruments in selected MR studies

Of the 128 articles screened, we excluded studies that (i) used two-sample MR ($n=30$), (ii) did not use single SNP or GRS as IV in the primary analysis ($n=10$), and (iii) used only outcome(s) other than related to blood pressure or mental health conditions ($n=71$), leading to 17 eligible articles. We further excluded two studies on which the LPC method was not applicable, either because the study did not report the estimated conventional effect ($n=1$) or reported the conventional and the MR estimates using different measures of association ($n=1$) leading to a total of 15 articles^{9,11,12,25-36} included in this illustration (**Table 3**).

Of the 15 articles, 11^{9,11,12,25-31,33} considered a continuous outcome. In the 11 studies, we identified a total of 18 estimated associations according to the IV used. In three^{9,11,12} of the 18 cases, the authors suspected the IV (single *FTO* SNP) to be invalid, which is consistent with the conclusion from the LPC method. In the remaining 15 estimated associations, the authors did not report invalid IVs unlike the LPC method which identified invalid IVs in six of the cases. Two of the six cases had modest sample sizes (2,204²⁸ and 1,731³³), while two others reported OLS and MR estimates

that were very close (0.047 vs. 0.06¹¹ and 0.66 vs. 0.70²⁷), and thus it is unclear whether their comparison is indicative of a violation of MR assumptions or of sampling error (**Table 3**).

Six^{25,29,32,34-36} of the 15 included articles considered a binary outcome. None of the studies reported any invalid IVs, while the LPC method identified two cases of invalid IVs. In both instances the conventional and MR estimates were very close (1.126 vs. 1.128²⁵ and 1.16 vs. 1.18³²), which suggests that the results of the LPC method are difficult to interpret. In one study,³⁶ six different GRS were used as IVs and the LPC method led to different conclusions depending on the GRS used (**Table 3**).

Discussion

The LPC method proposed by Glymour et al.¹ states that under net positive confounding, an IV is invalid if its corresponding MR point estimate of the exposure-to-outcome association is larger than the point estimate obtained using conventional regression. Investigation of the performance of the LPC method has been limited to small-sample MR studies with a continuous outcome and a single SNP as the IV.¹³ Our investigation extends previous work by (i) investigating binary outcomes in addition to continuous outcomes; (ii) considering GRS constructed from multiple SNPs as IV; (iii) generating several plausible scenarios regarding the sample size; and (iv) using published studies to illustrate the use of the LPC method and to compare its conclusions with that made by the authors based on alternative approaches.

Four main observations may be highlighted: (i) in MR using a single SNP as IV, the LPC method performed better in detecting direct than indirect pleiotropy, yet, as the sample size increased, the detection of indirect pleiotropy improved; (ii) the performance of the LPC method in detecting invalidity of a GRS as IV improved with the number of pleiotropic SNPs included in the GRS; (iii) the LPC method did not perform as well in cases with binary rather than continuous outcomes when the IV was a single SNP, but this was not the case when using the GRS as IV; and (iv) the LPC method was easily used in published articles but its conclusions regarding the presence of invalid IVs may be challenged when the conventional point estimates compared are not sufficiently precise.

Our results simulating MR studies using a SNP as a single IV align with and extend a previous simulation study by Burgess.¹³ Burgess¹³ simulated a small sample size by MR standards ($n=2,000$) and observed that while the LPC method detected invalid IVs in almost all the scenarios with direct pleiotropy, its performance strikingly declined in the presence of indirect pleiotropy (65% of invalid IVs detected). Further, the LPC method incorrectly identifies valid IVs as invalid in 11% of cases in which there was no pleiotropy. Using a similar simulation design, our results suggest that given a sufficiently large sample size that is typical of many MR studies ($n=30,000$), the LPC method detects more than 94% of simulated samples in scenario with indirect pleiotropy. In addition, the percentage of valid IVs falsely identified as invalid drops to 0 when in sample sizes greater than 10,000. We observed that increasing the sample size and thus the precision of both conventional and MR estimates improves the performance of the LPC method. Precise estimates reduce the likelihood of incorrectly identifying an IV as invalid (or valid, in turn) because of sampling variation. Finally, our study suggests that in MR studies with a single SNP used as the IV and a binary outcome, the performance of the LPC method improves with increasing sample size similarly to what was observed with a continuous outcome.

GRS are increasingly used in MR studies to increase the strength of the instrument and thus the precision of MR studies.^{37,38} The inclusion of pleiotropic SNPs in the GRS may lead to bias in the MR estimates.²² Our investigation of the performance of the LPC method in MR studies using a GRS as IV suggests that the LPC method is unlikely to incorrectly identify valid IVs as invalid in the absence of pleiotropy except in case of binary outcome and small sample. Indeed, our simulations showed that the performance of the method improved with increasing proportion of pleiotropic SNPs included in the GRS. In our simulations, the bias in the causal effect estimate increased with the number of pleiotropic SNPs in the GRS and so did the performance of the LPC method to detect invalid instruments. In real settings, caution must be exercised in interpreting the results of the LPC method when the IV is a GRS because the method could fail to identify the IV as invalid in the presence of only a few pleiotropic SNPs that could still bias the MR estimates.

Our illustration of the use of the LPC method in 15 MR studies highlights both a strength and a weakness of the method. First, we observe that the LPC method was applicable in 15 of 17 (88.2%) published MR studies that investigate the association between BMI and SBP, DBP, hypertension, or mental health. If this generalizes to other exposures and outcomes of interest, then it suggests that

the LPC method may be used as useful tool in reading or reviewing published MR studies, without requiring access to the dataset of interest. Second, our application also highlights that the results of the LPC method may be challenging to interpret when the MR and conventional estimates are very close. In finite samples, when the MR and the conventional estimates are close, the LPC method may falsely conclude that an IV is invalid (or not invalid) because of sampling variability, especially when the sample size is low, and the estimates are imprecise. Integrating the sampling variation around the point estimates and their difference in the LPC method may improve its interpretability.

MR studies involving single SNP are still frequently used in estimating causal effects of several exposures³⁹⁻⁴⁴ and our results suggest that the LPC method is worth adding to the toolbox of methods used to assess the validity of IVs in MR studies using a single SNP as IV. Indeed, the LPC method performed well in scenarios with plausible sample sizes. In addition, methods to assess the validity of IVs usually rely on several SNPs and are thus not applicable to studies using a single SNP as the IV (e.g. robust methods, methods based on the analysis of heterogeneity of SNP-specific causal effects, etc.).⁸ When several SNPs are available, the LPC method may still complement existing methods that can assess the validity of a GRS as an IV, e.g. MR-Egger intercept test or heterogeneity test based on Cochran's Q statistic using summary statistics from the SNPs-outcome and SNPs-exposure associations. Nevertheless, all of these methods, including the LPC method, provide tools to identify invalid IVs, not to demonstrate that IVs are valid.

A limitation of the current study is that although we evaluated the performance of the LPC method in several practical applications of MR studies, we did not consider the impact of varying the strength of the IV or the magnitude of their associations with the outcome, whether direct or thorough a confounder of the exposure and the outcome. Indeed, previous analytical work using continuous outcomes and multiple IVs suggest that the performance of the LPC method depends on several parameters, including the magnitude of the IV-exposure, and IV-outcome associations.⁴⁵ Additional studies are required to extend the analytical findings to scenarios including a binary outcome and a GRS as an IV. In addition, future studies will need to compare the performance of the LPC method against alternative approaches, with an emphasis on cases for which the assumptions underlying each of the approaches are differentially met to highlight the relative possible benefit of each method.

In summary, the present study shows that the LPC method proposed by Glymour et al.⁵ can be useful as a sensitivity analysis method to assess pleiotropy in MR using a single SNP as IV provided the sample size is sufficient. In contrast, in MR study using a GRS as IV, the LPC method seems to perform less well in detecting pleiotropy, especially if the proportion of pleiotropic SNPs in the GRS is not large.

Conflict of Interest

The authors declare no conflict of interest in this study.

Author Contributions

FINM and MPS designed the study and planned the simulation scenarios. DBR programmed the simulations with collaboration of MPS. FINM performed the simulations and conducted the literature search and data extraction. All authors participated to the interpretation of the data. FINM draft the manuscript and updated all versions of the manuscript. MPS, DBR, and CL critically reviewed the manuscript. All authors approved the final version of the manuscript.

Funding

This research was supported by the Canadian Institutes of Health Research (grant number CIHR PJT-148866). FINM was supported by a PhD scholarship from the Fonds de recherche du Québec – Santé (FRQS) (file number 36057). MPS is supported by a Chercheur-Boursier Junior II career award from the Fonds de recherche du Québec – Santé (FRQS) (file number 34875).

References

1. Glymour MM, Tchetgen Tchetgen EJ, Robins JM. Credible mendelian randomization studies: Approaches for evaluating the instrumental variable assumptions. *Am J Epidemiol* 2012; **175**: 332-9.
2. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nature Reviews Methods Primers* 2022; **2**: 6.
3. Spiga F, Gibson M, Dawson S, et al. Tools for assessing quality and risk of bias in Mendelian randomization studies: a systematic review. *Int J Epidemiol* 2022; **dyac149**.

4. Labrecque J, Swanson SA. Understanding the Assumptions Underlying Instrumental Variable Analyses: a Brief Review of Falsification Strategies and Related Tools. *Curr Epidemiol Rep* 2018; **5**: 214-20.
5. Hernán MA, Robins JM. Instruments for Causal Inference: An Epidemiologist's Dream? *Epidemiology* 2006; **17**: 360-72.
6. Richmond RC, Davey Smith G. Mendelian Randomization: Concepts and Scope. *Cold Spring Harb Perspect Med* 2022; **12**: a040501.
7. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018; **362**: k601.
8. Mbutiwi FIN, Dessy T, Sylvestre MP. Mendelian Randomization: A Review of Methods for the Prevention, Assessment, and Discussion of Pleiotropy in Studies Using the Fat Mass and Obesity-Associated Gene as an Instrument for Adiposity. *Front Genet* 2022; **13**: 803238.
9. Kivimäki M, Jokela M, Hamer M, et al. Examining overweight and obesity as risk factors for common mental disorders using fat mass and obesity-associated (FTO) genotype-instrumented analysis: The Whitehall II Study, 1985-2004. *Am J Epidemiol* 2011; **173**: 421-9.
10. Painter JN, O'Mara TA, Marquart L, et al. Genetic Risk Score Mendelian Randomization Shows that Obesity Measured as Body Mass Index, but not Waist:Hip Ratio, Is Causal for Endometrial Cancer. *Cancer Epidemiol Biomarkers Prev* 2016; **25**: 1503-10.
11. Walter S, Glymour MM, Koenen K, et al. Do genetic risk scores for body mass index predict risk of phobic anxiety? Evidence for a shared genetic risk factor. *Psychol Med* 2015; **45**: 181-91.
12. Walter S, Kubzansky LD, Koenen KC, et al. Revisiting Mendelian randomization studies of the effect of body mass index on depression. *Am J Med Genet B Neuropsychiatr Genet* 2015; **168B**: 108-15.
13. Burgess S. Re: "credible mendelian randomization studies: Approaches for evaluating the instrumental variable assumptions". *Am J Epidemiol* 2012; **176**: 456-7.
14. Morrison J, Knoblauch N, Marcus JH, Stephens M, He X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat Genet* 2020; **52**: 740-7.

15. Burgess S, Thompson SG. Bias in causal estimates from Mendelian randomization studies with weak instruments. *Stat Med* 2011; **30**: 1312-23.
16. Cai B, Small DS, Have TRT. Two-stage instrumental variable methods for estimating the causal odds ratio: Analysis of bias. *Stat Med* 2011; **30**: 1809-24.
17. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* 2017; **26**: 2333-55.
18. Vansteelandt S, Bowden J, Babanezhad M, Goetghebeur E. Vansteelandt, S., Bowden, J., Babanezhad, M., and Goetghebeur, E. 2011. On instrumental variables estimation of causal odds ratios. *Statistical Science* 2011; **26**: 403–22.
19. Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *Am J Epidemiol* 2004; **159**: 702-6.
20. Chen W, Qian L, Shi J, Franklin M. Comparing performance between log-binomial and robust Poisson regression models for estimating risk ratios under model misspecification. *BMC Med Res Methodol* 2018; **18**: 63.
21. Wang J. Dose as Instrumental Variable in Exposure–Safety Analysis Using Count Models. *J Biopharm Stat* 2012; **22**: 565-81.
22. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol* 2013; **42**: 1134-44.
23. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011; **40**: 755-64.
24. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008; **27**: 1133-63.
25. Fall T, Hagg S, Magi R, et al. The Role of Adiposity in Cardiometabolic Traits: A Mendelian Randomization Analysis. *PLoS Med* 2013; **10**: e1001474.
26. Fall T, Hagg S, Ploner A, et al. Age- and sex-specific causal effects of adiposity on cardiovascular risk factors. *Diabetes* 2015; **64**: 1841-52.
27. Holmes MV, Lange LA, Palmer T, et al. Causal effects of body mass index on cardiometabolic traits and events: a Mendelian randomization analysis. *Am J Hum Genet* 2014; **94**: 198-208.

28. Kivimaki M, Smith GD, Timpson NJ, et al. Lifetime body mass index and later atherosclerosis risk in young adults: examining causal links using Mendelian randomization in the Cardiovascular Risk in Young Finns study. *Eur Heart J* 2008; **29**: 2552-60.
29. Lyall DM, Celis-Morales C, Ward J, et al. Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank A Mendelian Randomization Study. *JAMA Cardiol* 2017; **2**: 882-9.
30. Wade KH, Chiesa ST, Hughes AD, et al. Assessing the causal role of body mass index on cardiovascular health in young adults: Mendelian randomization and recall-by-genotype analyses. *Circulation* 2018; **138**: 2187-201.
31. Wurtz P, Wang Q, Kangas AJ, et al. Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. *PLoS Med* 2014; **11**: e1001765.
32. Tyrrell J, Mulugeta A, Wood AR, et al. Using genetics to understand the causal influence of higher BMI on depression. *Int J Epidemiol* 2019; **48**: 834-48.
33. Jokela M, Elovainio M, Keltikangas-Järvinen L, et al. Body mass index and depressive symptoms: instrumental-variables regression with genetic risk score. *Genes, brain, and behavior* 2012; **11**: 942-8.
34. Lawlor DA, Harbord RM, Tybjaerg-Hansen A, et al. Using genetic loci to understand the relationship between adiposity and psychological distress: A Mendelian Randomization study in the Copenhagen General Population Study of 53221 adults. *J Intern Med* 2011; **269**: 525-37.
35. Hung CF, Rivera M, Craddock N, et al. Relationship between obesity and the risk of clinically significant depression: Mendelian randomisation study. *Br J Psychiatry* 2014; **205**: 24-8.
36. Lee MR, Lim YH, Hong YC. Causal association of body mass index with hypertension using a Mendelian randomization design. *Medicine (Baltimore)* 2018; **97**: e11252.
37. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Stat Med* 2016; **35**: 1880-906.
38. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol* 2014; **43**: 922-9.

39. Abbasi A, Deetman PE, Corpeleijn E, et al. Bilirubin as a potential causal factor in type 2 diabetes risk: a Mendelian randomization study. *Diabetes* 2015; **64**: 1459-69.
40. Kunutsor SK, Frysz M, Verweij N, Kieneker LM, Bakker SJL, Dullaart RPF. Circulating total bilirubin and risk of non-alcoholic fatty liver disease in the PREVEND study: observational findings and a Mendelian randomization study. *Eur J Epidemiol* 2020; **35**: 123-37.
41. Lee YS, Cho Y, Burgess S, et al. Serum gamma-glutamyl transferase and risk of type 2 diabetes in the general Korean population: a Mendelian randomization study. *Hum Mol Genet* 2016; **25**: 3877-86.
42. Barning F, Abarin T. Assessing the Causality Factors in the Association between (Abdominal) Obesity and Physical Activity among the Newfoundland Population-A Mendelian Randomization Analysis. *Genet Epigenet* 2016; **8**: 15-24.
43. Klovaite J, Benn M, Nordestgaard BG. Obesity as a causal risk factor for deep venous thrombosis: a Mendelian randomization study. *J Intern Med* 2015; **277**: 573-84.
44. Tan AG, Kifley A, Flood VM, et al. Evaluating the associations between obesity and age-related cataract: a Mendelian randomization study. *Am J Clin Nutr* 2019; **110**: 969-76.
45. Boucher-Roy D. *Détection de l'invalidité et estimation d'un effet causal en présence d'instruments invalides dans un contexte de randomisation mendélienne* [Masters dissertation]. Montréal (QC): Université de Montréal; 2022.

Table 1

Simulation Cases 1 (left panel) and 2 (right panel): Performance of the LPC method in assessing the validity of a single SNP as IV in three scenarios allowing for different sample sizes and a null causal effect ($\beta_{YX} = 0$).

n	Mean <i>F</i> stat	Continuous outcome					Binary outcome				
		Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	$\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%)	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	$\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%)
Scenario 1. Valid IV											
2,000 ¹	37.1	0.00	0.17	0.20	0.02	11.6	-0.02	0.54	0.19	0.07	34.5
5,000	91.2	0.00	0.11	0.20	0.01	2.9	-0.01	0.33	0.19	0.04	26.6
10,000	181.3	0.00	0.08	0.20	0.01	0.4	0.00	0.23	0.19	0.03	19.9
20,000	361.2	0.00	0.05	0.20	0.01	0.0	0.00	0.16	0.19	0.02	11.6
30,000	540.5	0.00	0.04	0.20	0.01	0.0	0.00	0.13	0.19	0.02	6.8
Scenario 2. Invalid IV through association with confounder (indirect pleiotropy)											
2,000 ¹	65.1	0.25	0.13	0.20	0.02	64.8	0.26	0.39	0.20	0.07	56.4
5,000	161.5	0.25	0.08	0.20	0.01	73.8	0.24	0.24	0.20	0.04	57.6
10,000	320.9	0.25	0.06	0.20	0.01	81.5	0.24	0.17	0.20	0.03	61.3
20,000	640.2	0.25	0.04	0.20	0.01	89.7	0.25	0.12	0.20	0.02	65.4
30,000	960.9	0.25	0.03	0.20	0.01	94.2	0.25	0.10	0.20	0.02	68.8
Scenario 3. Invalid IV through direct association with outcome (direct pleiotropy)											
2,000 ¹	37.1	0.68	0.19	0.21	0.02	99.8	0.68	0.52	0.20	0.07	82.6
5,000	90.6	0.67	0.12	0.21	0.01	100.0	0.66	0.32	0.20	0.04	92.9
10,000	181.1	0.67	0.08	0.21	0.01	100.0	0.66	0.22	0.20	0.03	98.2
20,000	361.2	0.67	0.06	0.21	0.01	100.0	0.65	0.16	0.20	0.02	99.8
30,000	541.9	0.67	0.05	0.21	0.01	100.0	0.65	0.13	0.20	0.02	100.0

SNP = single-nucleotide polymorphism; IV = instrumental variable; *F* stat = first stage *F* statistic; $\hat{\beta}_{MR}$ = MR estimate across simulations; SE_{MR} = standard error of $\hat{\beta}_{MR}$; $\hat{\beta}_{CO}$ = conventional estimate across simulations; SE_{CO} = standard error of $\hat{\beta}_{CO}$; $\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%) = percentage of simulated data sets in which the LPC method detects invalid IV.

¹Corresponds to replications of results presented in Burgess (2012).¹³

See **Table S2** for results allowing for positive ($\beta_{XY} = 0.4$) and negative causal effect ($\beta_{XY} = -0.4$).

Table 2

Simulation Cases 3 (left panel) and 4 (right panel): Performance of the LPC method in assessing the validity of the genetic risk score (GRS) as IV in three scenarios allowing different sample sizes and proportion of pleiotropic SNPs in the GRS, and a null causal effect ($\beta_{YX} = 0$).

n	Mean F stat	Continuous outcome					Binary outcome				
		Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	$\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%)	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	$\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%)
Scenario 1. Valid IV											
2,000	38.8	-0.01	0.17	0.49	0.02	0.0	-0.01	0.46	0.47	0.06	10.7
5,000	94.6	-0.01	0.10	0.49	0.01	0.0	-0.00	0.29	0.46	0.04	2.6
10,000	189.2	-0.00	0.08	0.49	0.01	0.0	0.01	0.19	0.46	0.02	0.6
20,000	379.4	-0.00	0.05	0.49	0.01	0.0	-0.01	0.12	0.46	0.02	0.0
30,000	567.8	-0.00	0.04	0.49	0.01	0.0	0.00	0.10	0.46	0.01	0.0
Scenario 2. Invalid IV through association with confounder (indirect pleiotropy)											
10% pleiotropic SNPs											
2,000	38.6	0.09	0.15	0.49	0.02	0.6	0.07	0.35	0.47	0.06	16.2
5,000	95.7	0.09	0.09	0.49	0.01	0.1	0.10	0.21	0.47	0.03	7.1
10,000	190.3	0.10	0.08	0.49	0.01	0.0	0.09	0.17	0.47	0.03	1.6
20,000	379.2	0.10	0.05	0.49	0.01	0.0	0.10	0.13	0.47	0.02	0.6
30,000	568.8	0.10	0.04	0.49	0.01	0.0	0.10	0.10	0.47	0.02	0.1
30% pleiotropic SNPs											
2,000	39.2	0.30	0.16	0.50	0.02	14.1	0.27	0.51	0.47	0.06	31.3
5,000	95.2	0.29	0.08	0.50	0.01	5.3	0.29	0.28	0.47	0.03	25.2
10,000	189.8	0.30	0.07	0.50	0.01	4.3	0.28	0.19	0.47	0.03	15.7
20,000	377.7	0.30	0.04	0.50	0.01	3.7	0.29	0.12	0.47	0.02	10.1
30,000	567.8	0.30	0.04	0.50	0.01	3.5	0.28	0.10	0.47	0.01	7.6
50% pleiotropic SNPs											
2,000	38.7	0.51	0.15	0.50	0.02	51.8	0.51	0.35	0.47	0.06	54.7
5,000	95.8	0.50	0.09	0.50	0.01	50.7	0.47	0.22	0.48	0.04	48.6
10,000	189.4	0.50	0.06	0.50	0.01	49.4	0.48	0.16	0.47	0.03	52.8
20,000	379.6	0.50	0.05	0.50	0.01	50.3	0.47	0.13	0.47	0.02	51.0
30,000	570.3	0.50	0.04	0.50	0.01	50.9	0.47	0.10	0.47	0.01	50.7
Scenario 3. Invalid IV through direct association with outcome (direct pleiotropy)											
10% pleiotropic SNPs											
2,000	39.0	0.09	0.22	0.49	0.02	0.5	0.07	0.46	0.47	0.06	17.1
5,000	95.2	0.09	0.10	0.49	0.01	0.0	0.09	0.27	0.47	0.04	7.8
10,000	190.4	0.10	0.07	0.49	0.01	0.0	0.08	0.17	0.47	0.03	1.4
20,000	376.4	0.10	0.05	0.49	0.01	0.0	0.10	0.13	0.47	0.02	0.1

30,000	567.7	0.10	0.04	0.49	0.01	0.0	0.09	0.11	0.47	0.02	0.0
30% pleiotropic SNPs											
2,000	39.2	0.29	0.18	0.50	0.02	11.9	0.29	0.28	0.47	0.05	32.5
5,000	95.4	0.30	0.10	0.50	0.01	6.4	0.28	0.25	0.47	0.04	25.0
10,000	189.9	0.30	0.08	0.50	0.01	2.9	0.29	0.19	0.47	0.03	18.7
20,000	378.1	0.31	0.05	0.50	0.01	3.9	0.29	0.12	0.47	0.02	10.7
30,000	570.0	0.30	0.04	0.50	0.01	2.6	0.29	0.10	0.47	0.01	8.6
50% pleiotropic SNPs											
2,000	38.4	0.49	0.18	0.50	0.02	48.4	0.50	0.48	0.48	0.06	52.5
5,000	97.2	0.50	0.11	0.50	0.01	51.2	0.47	0.23	0.47	0.03	50.6
10,000	189.3	0.50	0.06	0.50	0.01	49.8	0.47	0.17	0.47	0.02	49.2
20,000	381.7	0.50	0.04	0.50	0.01	50.0	0.46	0.12	0.47	0.02	48.5
30,000	568.6	0.50	0.04	0.50	0.01	50.8	0.47	0.09	0.47	0.01	49.7

IV = instrumental variable; SNP = single-nucleotide polymorphism; F stat = first stage F statistic; $\hat{\beta}_{MR}$ = MR estimate across simulations; SE_{MR} = standard error of $\hat{\beta}_{MR}$; $\hat{\beta}_{CO}$ = conventional estimate across simulations; SE_{CO} = standard error of $\hat{\beta}_{CO}$; $\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%) = percentage of simulated data sets in which the LPC method detects invalid IV.

See **Table S3** for results allowing for positive ($\beta_{XY} = 0.4$) and negative causal effect ($\beta_{XY} = -0.4$).

Table 3

Comparison of the LPC method with alternative approaches in assessing the validity of the instruments in selected MR studies.

Outcome	Study	IV	n	Illustration of the LPC method			Suspicion of IV invalidity in the study ^c
				Conventional estimate ^a	MR estimate ^a	Suspicion of invalid IV ^b	
Continuous outcomes							
Systolic blood pressure	Fall et al., 2013 ²⁵	<i>FTO</i> rs9939609	147,644	0.903	0.892	No	No
	Fall et al., 2015 ²⁶	GRS based on 32 SNPs	117,781/67,553 ^d	0.22	0.16	No	No
	Holmes et al., 2014 ²⁷	GRS based on 14 SNPs	30,136	0.66	0.70	Yes	No
	Kivimaki et al., 2008 ²⁸	<i>FTO</i> rs9939609	2,204	0.36	0.95	Yes	No
	Lyall et al., 2017 ²⁹	GRS based on 93 SNPs	111,637	3.11/3.02 ^e	1.53/1.65 ^e	No	No
	Wade et al., 2018 ³⁰	GRS based on 97 SNPs	3,108	0.86	0.79	No	No
	Wurtz et al., 2014 ³¹	GRS based on 32 SNPs	12,635	0.973	1.09	Yes	No
Diastolic blood pressure	Fall et al., 2013 ²⁵	<i>FTO</i> rs9939609	130,380	0.619	0.490	No	No
	Fall et al., 2015 ²⁶	GRS based on 32 SNPs	117,230/66,997 ^d	0.25	0.15	No	No
	Lyall et al., 2017 ²⁹	GRS based on 93 SNPs	111,638	2.75/2.82 ^e	1.23/1.37 ^e	No	No
	Wade et al., 2018 ³⁰	GRS based on 97 SNPs	3,108	0.52	0.29	No	No
	Wurtz et al., 2014 ³¹	GRS based on 32 SNPs	12,627	0.788	1.12	Yes	No
Phobic anxiety score	Walter et al., 2015 ¹¹	<i>FTO</i> rs1558902	9,608	0.047	0.32	Yes	Yes
	Walter et al., 2015 ¹¹	GRS based on 32 SNPs	9,608	0.047	0.06	Yes	No
Depression score	Walter et al., 2015 ¹²	<i>FTO</i> rs1558902	6,989	0.024	0.064	Yes	Yes
	Walter et al., 2015 ¹²	GRS based on 32 SNPs	6,989	0.024	0.013	No	No
Depressive symptoms score	Jokela et al., 2012 ³³	GRS based on 31 SNPs	1,731	0.47/0.33 ^f	1.08/1.96 ^f	Yes	No
Common mental disorders score	Kivimaki et al., 2011 ⁹	<i>FTO</i> rs1421085	2,981	0.009	0.171 ^g	Yes	Yes
Binary outcomes							
Hypertension	Fall et al., 2013 ²⁵	<i>FTO</i> rs9939609	155,191	1.126	1.128	Yes	No
	Lee et al., 2018 ³⁶	Six different GRS based on 2 to 6 SNPs	8,832	1.15/1.19 ^h	1.13 to 1.26	No/Yes ⁱ	No
	Lyall et al., 2017 ²⁹	GRS based on 93 SNPs	119,859	1.76/1.75 ^e	1.65/1.64 ^e	No	No
Depression	Tyrrell et al., 2019 ³²	GRS based on 73 SNPs	287,503	1.16	1.18	Yes	No

Major depression	Hung et al., 2014 ³⁵	<i>FTO</i> rs3751812	3,222	0.05*	-0.03*	No	No
	Hung et al., 2014 ³⁵	GRS based on 32 SNPs	2,521	0.05*	-0.02*	No	No
Stress/nervous	Lawlor et al., 2011 ³⁴	<i>FTO</i> rs9939609	53,221	0.97	0.65	No	No
Not accomplishing	Lawlor et al., 2011 ³⁴	<i>FTO</i> rs9939609	53,221	1.11	0.69	No	No
Feels like giving up	Lawlor et al., 2011 ³⁴	<i>FTO</i> rs9939609	53,221	1.10	0.63	No	No
Antidepressant medication	Lawlor et al., 2011 ³⁴	<i>FTO</i> rs9939609	53,221	1.04	0.57	No	No

Abbreviations: IV, instrumental variable; LPC, Leveraging positive confounding (method); MR, Mendelian randomization; *FTO*, fat mass and obesity-associated (gene); GRS, genetic risk score; SNPs, single-nucleotide polymorphisms; 2SLS, two-stage least squares; 2SRI, two-stage residual inclusion (or control function); 2SPS, two-stage predictor substitution.

^a Estimates represent the beta coefficients for continuous outcomes and odds ratio or beta coefficients (marked with an asterisk) for binary outcomes per 1 unit increase in body mass index.

^b The conclusion of the LPC method assumes net positive confounding between BMI (exposure) and the outcome of interest.

^c The suspicion of IV invalidity in the study was based on alternative approaches provided by the authors.

^d n of the conventional analysis/n of the MR analysis.

^e Estimates from partially/fully adjusted models.

^f Estimates using adult/adolescent BMI as exposure.

^g Estimates from the unadjusted model using mean BMI as exposure in men.

^h Estimates from unadjusted/adjusted models

ⁱ Depending on the GRS, the LPC method identified invalid IVs.

Supplementary Tables

Table S1

Summary of parameters used in the simulations.

	Simulation Cases			
	1	2	3	4
<i>Y</i>	Continuous	Binary	Continuous	Binary
IV	Single SNP		Unweighted GRS	
Number of SNP(s)	1		25	
SNP creation	SNP generated as binary (0,1) dividing the population into two equal subgroups		Each SNP generated as 0, 1, and 2 (number of effect alleles), with effect allele frequency of 0.3	
Coefficient of $G - X$ association (β_{XG})	0.3		0.06	
Coefficient of $U - X$ (β_{XU}) or $U - Y$ (β_{YU}) association	0.5		1	
Coefficient of $G - U$ (β_{GU} , scenario 2 of indirect pleiotropy) or $G - Y$ (β_{YG} , scenario 3 of direct pleiotropy) association	0.2		0.06	
True causal effect of X on Y (β_{YX})	0 (null), 0.4 (positive), and -0.4 (negative), respectively			
Simulated sample sizes	Initially 2,000, then increased to 5,000, 10,000, 20,000 and 30,000 subjects			
Number of iterations	10,000		1,000	

IV, instrumental variable; SNP, single-nucleotide polymorphism; GRS, genetic risk score.

G is the IV, X the exposure, Y the outcome, and U a confounder on the association between X and Y .

Table S2

Simulation Cases 1 (left panel) and 2 (right panel): Performance of the LPC method in assessing the validity of a single SNP as IV in three scenarios allowing different sample sizes, and positive and negative causal effect.

n	Continuous outcome										Binary outcome									
	Positive causal effect ($\beta_{YX} = 0.4$)					Negative causal effect ($\beta_{YX} = -0.4$)					Positive causal effect ($\beta_{YX} = 0.4$)					Negative causal effect ($\beta_{YX} = -0.4$)				
	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	$\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%)	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	$\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%)	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	$\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%)	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	$\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%)
Scenario 1. Valid IV																				
2,000 ¹	0.39	0.18	0.60	0.02	11.5	-0.40	0.17	-0.20	0.02	11.8	0.38	0.54	0.59	0.07	34.6	-0.39	0.54	-0.20	0.07	35.2
5,000	0.40	0.11	0.60	0.01	2.6	-0.40	0.11	-0.20	0.01	2.7	0.39	0.33	0.58	0.04	27.1	-0.39	0.33	-0.20	0.04	27.8
10,000	0.40	0.08	0.60	0.01	0.4	-0.40	0.08	-0.20	0.01	0.4	0.39	0.23	0.58	0.03	19.5	-0.39	0.23	-0.20	0.03	20.1
20,000	0.40	0.05	0.60	0.01	0.0	-0.40	0.05	-0.20	0.01	0.0	0.39	0.16	0.58	0.02	10.9	-0.40	0.16	-0.20	0.02	11.8
30,000	0.40	0.04	0.60	0.01	0.0	-0.40	0.04	-0.20	0.01	0.0	0.39	0.13	0.58	0.02	6.7	-0.39	0.13	-0.20	0.02	6.8
Scenario 2. Invalid IV through association with confounder (indirect pleiotropy)																				
2,000 ¹	0.65	0.13	0.60	0.02	65.4	-0.15	0.13	-0.20	0.02	65.3	0.64	0.38	0.59	0.07	55.0	-0.15	0.39	-0.19	0.07	54.8
5,000	0.65	0.08	0.60	0.01	74.1	-0.15	0.08	-0.20	0.01	74.4	0.64	0.24	0.59	0.04	58.7	-0.15	0.24	-0.19	0.04	56.9
10,000	0.65	0.06	0.60	0.01	81.9	-0.15	0.06	-0.20	0.01	81.3	0.64	0.17	0.59	0.03	61.0	-0.15	0.17	-0.19	0.03	61.0
20,000	0.65	0.04	0.60	0.01	89.9	-0.15	0.04	-0.20	0.01	89.8	0.64	0.12	0.59	0.02	65.6	-0.15	0.12	-0.19	0.02	66.1
30,000	0.65	0.03	0.60	0.01	94.4	-0.15	0.03	-0.20	0.01	94.1	0.64	0.10	0.59	0.02	69.9	-0.15	0.10	-0.19	0.02	68.6
Scenario 3. Invalid IV through direct association with outcome (direct pleiotropy)																				
2,000 ¹	1.09	0.19	0.61	0.02	99.9	0.28	0.19	-0.19	0.02	99.8	1.05	0.52	0.60	0.07	81.9	0.28	0.52	-0.19	0.07	82.4
5,000	1.07	0.12	0.61	0.01	100.0	0.27	0.12	-0.19	0.01	100.0	1.05	0.32	0.59	0.04	92.6	0.26	0.32	-0.19	0.04	92.6
10,000	1.07	0.08	0.61	0.01	100.0	0.27	0.08	-0.19	0.01	100.0	1.05	0.22	0.59	0.03	98.0	0.27	0.22	-0.19	0.03	98.2
20,000	1.07	0.06	0.61	0.01	100.0	0.27	0.06	-0.19	0.01	100.0	1.05	0.16	0.59	0.02	99.9	0.26	0.16	-0.19	0.02	99.9
30,000	1.07	0.05	0.61	0.01	100.0	0.27	0.05	-0.19	0.01	100.0	1.04	0.13	0.59	0.02	100.0	0.26	0.13	-0.19	0.02	100.0

SNP = single-nucleotide polymorphism; IV = instrumental variable; $\hat{\beta}_{MR}$ = MR estimate across simulations; SE_{MR} = standard error of $\hat{\beta}_{MR}$; $\hat{\beta}_{CO}$ = conventional estimate across simulations; SE_{CO} = standard error of $\hat{\beta}_{CO}$; $\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%) = percentage of simulated data sets in which the LPC method detects invalid IV.

¹Corresponds to replications of results presented in Burgess (2012).⁴¹

Table S3

Simulation Cases 3 (left panel) and 4 (right panel): Performance of the LPC method in assessing the validity of the genetic risk score (GRS) as IV in three scenarios allowing different sample sizes and proportion of pleiotropic SNPs in the GRS, and positive and negative causal effect.

n	Continuous outcome										Binary outcome									
	Positive causal effect ($\beta_{YX} = 0.4$)					Negative causal effect ($\beta_{YX} = -0.4$)					Positive causal effect ($\beta_{YX} = 0.4$)					Negative causal effect ($\beta_{YX} = -0.4$)				
	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	Mean $\hat{\beta}_{MR} > \hat{\beta}_{CO} (\%)$	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	Mean $\hat{\beta}_{MR} > \hat{\beta}_{CO} (\%)$	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	Mean $\hat{\beta}_{MR} > \hat{\beta}_{CO} (\%)$	Mean $\hat{\beta}_{MR}$	Mean SE_{MR}	Mean $\hat{\beta}_{CO}$	Mean SE_{CO}	Mean $\hat{\beta}_{MR} > \hat{\beta}_{CO} (\%)$
Scenario 1. Valid IV																				
2,000	0.39	0.18	0.89	0.02	0.0	-0.42	0.16	0.09	0.02	0.0	0.38	0.53	0.84	0.07	14.7	-0.38	0.43	0.09	0.06	14.2
5,000	0.39	0.09	0.89	0.01	0.0	-0.41	0.11	0.09	0.01	0.0	0.38	0.23	0.84	0.04	3.9	-0.40	0.23	0.09	0.03	3.3
10,000	0.40	0.08	0.89	0.01	0.0	-0.40	0.08	0.09	0.01	0.0	0.37	0.16	0.84	0.03	0.5	-0.38	0.19	0.09	0.03	0.4
20,000	0.40	0.05	0.89	0.01	0.0	-0.40	0.05	0.09	0.01	0.0	0.38	0.13	0.84	0.02	0.0	-0.39	0.15	0.09	0.02	0.0
30,000	0.40	0.04	0.89	0.01	0.0	-0.40	0.04	0.09	0.01	0.0	0.38	0.11	0.84	0.02	0.0	-0.38	0.11	0.09	0.01	0.0
Scenario 2. Invalid IV through association with confounder (indirect pleiotropy)																				
<i>10% pleiotropic SNPs</i>																				
2,000	0.49	0.14	0.89	0.02	0.4	-0.31	0.19	0.09	0.02	0.6	0.48	0.37	0.84	0.07	17.3	-0.30	0.37	0.09	0.05	16.8
5,000	0.50	0.10	0.89	0.01	0.0	-0.31	0.09	0.09	0.01	0.0	0.48	0.26	0.84	0.04	8.6	-0.30	0.26	0.09	0.03	5.5
10,000	0.49	0.07	0.89	0.01	0.0	-0.30	0.06	0.09	0.01	0.0	0.47	0.16	0.84	0.03	3.1	-0.29	0.18	0.09	0.02	1.5
20,000	0.50	0.05	0.89	0.01	0.0	-0.30	0.05	0.09	0.01	0.0	0.47	0.13	0.84	0.02	0.3	-0.29	0.13	0.09	0.02	0.4
30,000	0.50	0.04	0.89	0.01	0.0	-0.30	0.04	0.09	0.01	0.0	0.47	0.10	0.84	0.02	0.1	-0.29	0.10	0.09	0.01	0.0
<i>30% pleiotropic SNPs</i>																				
2,000	0.70	0.16	0.90	0.02	13.1	-0.11	0.19	0.10	0.02	11.5	0.64	0.35	0.84	0.07	31.3	-0.11	0.49	0.09	0.06	30.6
5,000	0.70	0.10	0.90	0.01	5.5	-0.11	0.09	0.10	0.01	6.6	0.68	0.26	0.84	0.04	28.3	-0.11	0.25	0.09	0.03	22.4
10,000	0.70	0.07	0.90	0.01	4.2	-0.11	0.07	0.10	0.01	3.0	0.66	0.17	0.85	0.03	19.9	-0.09	0.19	0.09	0.03	17.3
20,000	0.70	0.04	0.90	0.01	3.0	-0.10	0.04	0.10	0.01	2.8	0.66	0.14	0.85	0.02	11.1	-0.10	0.13	0.09	0.02	9.3
30,000	0.70	0.04	0.90	0.01	2.1	-0.10	0.04	0.10	0.01	2.0	0.66	0.11	0.84	0.02	9.6	-0.10	0.11	0.09	0.01	7.0
<i>50% pleiotropic SNPs</i>																				
2,000	0.91	0.15	0.90	0.02	50.0	0.10	0.16	0.10	0.02	49.0	0.87	0.45	0.85	0.07	51.4	0.09	0.35	0.10	0.05	49.0
5,000	0.90	0.08	0.90	0.01	49.8	0.10	0.08	0.10	0.01	50.0	0.86	0.21	0.85	0.04	50.8	0.08	0.23	0.10	0.03	50.2

10,000	0.90	0.06	0.90	0.01	51.9	0.10	0.07	0.10	0.01	49.0	0.85	0.17	0.85	0.03	50.7	0.10	0.20	0.10	0.03	54.6
20,000	0.90	0.05	0.90	0.01	48.0	0.10	0.04	0.10	0.01	52.0	0.85	0.15	0.85	0.02	52.6	0.09	0.11	0.10	0.02	48.9
30,000	0.90	0.04	0.90	0.01	50.3	0.10	0.04	0.10	0.01	50.8	0.84	0.11	0.85	0.02	48.6	0.10	0.09	0.10	0.01	51.4
Scenario 3. Invalid IV through direct association with outcome (direct pleiotropy)																				
<i>10% pleiotropic SNPs</i>																				
2,000	0.49	0.17	0.89	0.02	0.5	-0.31	0.17	0.09	0.02	0.8	0.46	0.41	0.84	0.06	18.2	-0.28	0.44	0.09	0.05	16.5
5,000	0.50	0.08	0.89	0.01	0.0	-0.31	0.08	0.09	0.01	0.0	0.45	0.29	0.84	0.04	5.6	-0.29	0.24	0.09	0.03	5.5
10,000	0.50	0.07	0.89	0.01	0.0	-0.30	0.06	0.09	0.01	0.0	0.47	0.18	0.84	0.03	2.3	-0.30	0.18	0.09	0.02	2.1
20,000	0.50	0.04	0.89	0.01	0.0	-0.31	0.05	0.09	0.01	0.0	0.47	0.14	0.84	0.02	0.3	-0.28	0.12	0.09	0.02	0.1
30,000	0.50	0.04	0.89	0.01	0.0	-0.30	0.04	0.09	0.01	0.0	0.47	0.11	0.84	0.02	0.1	-0.29	0.11	0.09	0.01	0.0
<i>30% pleiotropic SNPs</i>																				
2,000	0.69	0.16	0.90	0.02	12.6	-0.11	0.15	0.10	0.02	12.5	0.64	0.57	0.85	0.06	31.8	-0.12	0.29	0.09	0.05	30.6
5,000	0.70	0.09	0.90	0.01	6.7	-0.11	0.10	0.10	0.01	6.3	0.64	0.25	0.84	0.04	22.2	-0.08	0.25	0.09	0.03	25.1
10,000	0.70	0.08	0.90	0.01	4.5	-0.10	0.07	0.10	0.01	3.0	0.65	0.20	0.84	0.03	15.9	-0.10	0.16	0.09	0.02	15.8
20,000	0.70	0.05	0.90	0.01	3.4	-0.10	0.05	0.10	0.01	3.1	0.65	0.13	0.84	0.02	10.6	-0.10	0.13	0.09	0.02	8.9
30,000	0.70	0.04	0.90	0.01	2.4	-0.10	0.04	0.10	0.01	3.3	0.66	0.10	0.84	0.02	7.5	-0.10	0.10	0.09	0.01	7.7
<i>50% pleiotropic SNPs</i>																				
2,000	0.89	0.18	0.90	0.02	48.0	0.11	0.17	0.10	0.02	52.0	0.85	0.46	0.85	0.06	49.1	0.12	0.47	0.09	0.05	51.9
5,000	0.90	0.09	0.90	0.01	51.3	0.10	0.10	0.10	0.01	50.8	0.86	0.27	0.85	0.04	49.8	0.09	0.21	0.10	0.03	49.6
10,000	0.90	0.06	0.90	0.01	50.9	0.10	0.08	0.10	0.01	51.2	0.84	0.17	0.85	0.03	49.7	0.10	0.16	0.09	0.03	50.5
20,000	0.90	0.05	0.90	0.01	51.7	0.10	0.05	0.10	0.01	49.7	0.84	0.14	0.85	0.02	47.5	0.09	0.12	0.09	0.02	49.6
30,000	0.90	0.04	0.90	0.01	49.0	0.10	0.04	0.10	0.01	49.1	0.85	0.12	0.85	0.02	49.9	0.10	0.10	0.10	0.01	51.5

SNP = single-nucleotide polymorphism; IV = instrumental variable; $\hat{\beta}_{MR}$ = MR estimate across simulations; SE_{MR} = standard error of $\hat{\beta}_{MR}$; $\hat{\beta}_{CO}$ = conventional estimate across simulations; SE_{CO} = standard error of $\hat{\beta}_{CO}$; $\hat{\beta}_{MR} > \hat{\beta}_{CO}$ (%) = percentage of simulated data sets in which the LPC method detects invalid IV.

4.3. Article 3

Accounting for antihypertensive medication in Mendelian randomization studies of blood pressure: methodological considerations in the Canadian Longitudinal Study on Aging

Fiston Ikwa Ndol Mbutiwi^{1,2}

Marie-Pierre Sylvestre^{1,3}

¹University of Montreal Hospital Research Centre (CRCHUM), Montreal, QC, Canada

²Faculty of Medicine, University of Kikwit, Kikwit, Democratic Republic of the Congo

³Department of Social and Preventive Medicine, University of Montreal Public Health School (ESPUM), Montreal, QC, Canada

Publié dans la revue

Journal of Hypertension 2023;41(8):1281-1289. doi: 10.1097/HJH.0000000000003464.

Abstract

Background: Mendelian randomization (MR) studies investigating determinants of blood pressure (BP) do not account for antihypertensive medication consistently, which may explain discrepancies across studies. We performed an MR study of the association between body mass index (BMI) and systolic BP (SBP) using five methods to account for antihypertensive medication and evaluated their impact on the estimation of the causal effect and on the assessment of the invalidity of the instruments used in MR.

Methods: Baseline and follow-up data on 20,430 participants from the Canadian Longitudinal Study on Aging (CLSA) Comprehensive cohort (2011-2018) were used. The five methods to account for antihypertensive medication in the MR study were: (i) no correction, (ii) adjustment for antihypertensive medication as a covariate in models, (iii) exclusion of treated subjects, (iv) addition of a constant value of 15 mmHg to measured values of SBP in treated subjects, and (v) using hypertension as a binary outcome.

Results: The magnitude of the estimated MR causal effect for SBP (mmHg) varied across the methods of accounting for antihypertensive medication effects ranging from 0.68 (effect per 1 kg/m² increase in BMI) in scenario adjusting MR models for medication covariate to 1.35 in that adding 15 mmHg to measured SBP in treated subjects. Conversely, the assessment of the validity of the instruments did not differ across methods of accounting for antihypertensive medication.

Conclusions: Methods to account for antihypertensive medication in MR studies may affect the estimation of the causal effects and must be selected with caution.

Keywords: Mendelian randomization; instrumental variable; blood pressure; hypertension; body mass index; antihypertensive medication; CLSA.

Abbreviations: BMI, body mass index; BP, blood pressure; CLSA, Canadian Longitudinal Study on Aging; DBP, diastolic blood pressure; *FTO*, fat mass and obesity-associated (gene); GRS, genetic risk score; MR, mendelian randomization; PASE, Physical Activity Scale for the Elderly; SBP, systolic blood pressure; SNP, single-nucleotide polymorphism.

Introduction

Hypertension is one of the most widespread health burdens worldwide and the leading cause for cardiovascular morbidity and all-cause mortality globally [1]. In 2019, the age-standardized prevalence of hypertension in adults age 30-79 years was 32% in women and 34% in men worldwide and is expected to increase significantly by 2050 [2,3].

Observational studies investigating modifiable risk factors for elevated blood pressure (BP) are often limited in their ability to properly address confounding by lifestyle and sociodemographic factors. Several authors resorted to using Mendelian randomization (MR), a statistical method that uses genetic variants as instruments to estimate the causal association between an exposure and an outcome in the presence of unmeasured confounding [4]. A key feature of the MR literature on BP modelling is the inconsistent accounting for antihypertensive medication across studies. For example, out of the 10 MR studies of the association between body mass index (BMI) and BP identified in a recent review [5], five studies ignored the antihypertensive medication issue [6-10], while two studies adjusted for antihypertensive medication in the MR models [11,12]. The remaining three studies imputed the value of measured BP in treated participants by adding a constant value of respectively 15 mmHg to systolic BP (SBP) and 10 mmHg to diastolic BP (DBP) [13,14] or 10 mmHg to SBP and 5 mmHg to DBP [15].

Correctly accounting for antihypertensive medication in studies of BP is paramount because BP measured in individuals treated with antihypertensive medication is likely lower than the value they would have if they were not treated [16,17]. Consequently, simply modelling measured BP may distort the true association with a risk factor [16]. Such distortion is likely to be important because the use of antihypertensive medication is common (e.g., about 80% of Canadians diagnosed with hypertension used antihypertensive medication in 2012-2015) [18].

Tobin et al. [16] used simulations and real data to compare the accuracy and statistical power to detect an association between a genetic variant and SBP in conventional regression models under ten methods to account for antihypertensive medication. Their study suggests that imputing the observed value of BP in treated subjects (whether by adding a constant value of 15 mmHg to the measured SBP or using a censored normal regression model) outperformed alternative methods such as excluding treated subjects, ignoring the medication, or including it in the model as

covariate. These observations were illustrated in subsequent work assessing the association between diabetes or obesity and SBP [19] and extended in studies that further considered antihypertensive drug classes or number of antihypertensive medications [17]. Adding a constant value to measured BP is currently the most commonly used imputation method to account for antihypertensive medication [20-26], probably owing to its simplicity.

Interestingly, no study thus far has evaluated methods of accounting for antihypertensive medication in BP modelling in the MR setting. Yet, the findings from conventional regression analyses may not translate to MR studies. For example, MR studies usually require larger sample size because they rely on genetic variants as instruments and such variants typically explain a small proportion of the variance in the exposure [27]. Hence, methods that exclude individuals under medication may affect MR studies to a larger extent than conventional analyses. In addition, antihypertensive medication may be associated with the instruments used in MR studies and thus cause a violation of the exclusion restriction assumption requesting that instruments must only affect the outcome through the exposure [28]. For example, the fat mass and obesity-associated (*FTO*) gene is a commonly used instrument in MR studies of the impact of adiposity on BP. *FTO* has also been linked to mental health conditions (e.g. dementia) [29-31] which may in turn be related to adherence to antihypertensive medication [32,33]. This would create a path from the instrument (*FTO*) to the outcome (BP) that does not pass through the exposure (adiposity) and violate the exclusion restriction. While methods are available to detect violations of the exclusion restriction assumption in MR studies [5], little is known about their performance in conjunction with the methods for accounting for antihypertensive medication described previously.

In this article, we use data from the Canadian Longitudinal Study on Aging (CLSA) to compare how to account for antihypertensive medication effects in an MR study of the association between adiposity (based on BMI), a known risk factor for BP [1], and SBP. We used the single-nucleotide polymorphism (SNP) rs1558902 in the *FTO* gene, the strongest signal associated with BMI [34], as well as a genetic risk score as the two BMI instruments. We consider five alternative methods of accounting for antihypertensive medication and investigate the impact on the estimated MR causal effect and detection of violation of the exclusion restriction assumption.

Methods

Study participants

We used baseline (2011-2015) and first follow-up (2015-2018) data from the CLSA Comprehensive cohort [35]. The CLSA Comprehensive cohort included 30,097 participants ages 45-85 years at baseline, randomly selected from seven Canadian provinces. Participants were interviewed in person at home, then physically examined at 11 data collection sites where blood and urine samples were provided. A total of 26,622 participants from the Comprehensive cohort provided genetic material and 27,765 provided data in the first follow-up cycle. All CLSA participants were volunteers and provided informed written consent. The CLSA protocol was approved by several research ethics boards across Canada [36]. The current project received approval from the research ethics board of the Centre Hospitalier de l'Université de Montréal (2021-8642, 20.148-YP).

Our analytical sample comprised participants who provided phenotypic (baseline and first follow-up) and genetic data (n=24,679). We excluded 1,784 participants of non-European ancestry and 1,862 participants with invalid responses or missing/unavailable data for the variable of interest described below. Because the association between BMI and SBP was not linear for BMI values > 40 kg/m², we further removed 630 participants with BMI > 40 kg/m² from the sample. While MR studies can accommodate non-linear relationships [37], they require more than one parameter to capture the association between BMI and SBP. By restricting the analysis to the range of values for which the association between BMI and SBP is linear, we facilitate the comparisons of the five methods of accounting for antihypertensive medication. The flow chart describing the final analytical sample (n=20,430) is presented in **Figure S1**.

Measures

Baseline BMI (kg/m²), calculated as weight (kg) divided by the square of height (m²), was used as the exposure of interest. Weight and standing height were measured in shoeless participants using a 140-10 Healthweigh digital physician scale (ranged up to 250 kg) and Seca 213 stadiometer, respectively [36].

Data on BP measurements, medical history of hypertension diagnosis or use of antihypertensive medication were obtained from the first follow-up examination. SBP and DBP were measured using an automated non-invasive BP monitor model BpTRU™ BPM200 (VSM MedTech Ltd., Coquitlam, BC, Canada) [36]. Six measurements were taken, and we considered the average of the last five SBP and DBP measurements [38]. Hypertension was defined as taking the value of 1 if at least one of the two following criteria were met: (i) SBP \geq 140 mmHg or DBP \geq 90 mmHg, and/or (ii) self-reported diagnosis of hypertension and current use of antihypertensive medication, and 0 otherwise [39].

The CLSA Comprehensive cohort genotyping was carried out using the Affymetrix UK Biobank Axiom array and provided data on 794,409 markers, leading to whole-genome imputed data on approximately 308 million genetic variants [40]. Marker- and sample-based quality control included tests for detection of discordant genotype frequency between batches, departure from Hardy-Weinberg equilibrium, discordance of genotyping across control replicates, and for discordance in genotype frequency between males and females, as well as minor allele frequency $>$ 0.01, and marker-wise missingness $<$ 0.01. Additional details on the quality control process and genetic variants imputation appear in Forgetta et al. [40].

We selected 97 single-nucleotide polymorphisms (SNPs) associated with BMI at genome-wide significance level ($P < 5 \times 10^{-8}$) in a large meta-analysis of 125 studies including 339,224 individuals [41]. The 97 SNPs were available in the CLSA Comprehensive cohort genetic dataset (**Table S1**). We recoded the SNPs to represent the number of alleles associated with an increase in BMI, and then harmonized alleles by comparing allele frequency with that reported by Locke et al. [41]. Ten SNPs (rs11165643, rs9641123, rs3736485, rs7715256, rs2075650, rs6477694, rs17724992, rs11847697, rs1528435, and rs6465468) were excluded because of their estimated association with BMI was negative in the CLSA. The remaining 87 SNPs were used to construct an unweighted GRS by summing the number of BMI-increasing alleles in each participant [42]. We used both the single *FTO* genotype (rs1558902) and the GRS based on 87 SNPs (which included *FTO*), as the two instruments for BMI. We further constructed an unweighted GRS based on all 97 BMI-associated SNPs used in sensitivity analysis.

We considered covariables associated with BMI and/or BP including baseline values for: (i) age; (ii) sex; (iii) marital status ('with partner' (living with a partner in a common-law relationship) vs.

‘no partner’ (single, never married or never lived with a partner/widowed/divorced/separated)) [43]; (iv) education level (graduate from high/secondary school or not); (v) smoking status (current/former daily smoker vs. other) [44]; (vi) Frequency of alcohol consumption in the past 12 months (‘drinker’ (at least once a month) or ‘non-drinker’ (less than once a month)) [45]²³¹; and (vii) physical activity using the Physical Activity Scale for the Elderly (PASE) [46].

We also considered the top ten principal components of the principal components analysis from the CLSA European ancestry subset to adjust for population stratification [40,47].

Statistical analysis

Baseline characteristics of participants were described using mean (standard deviation, SD) or frequency (n, %) distinguishing included participants from other participants excluded from the study.

Primary analyses

We modelled the associations between BMI and both SBP (continuous variable) and hypertension (binary indicator) with conventional linear and robust Poisson for common events [48,49] regression, respectively. We then performed MR analyses using the two-stage least squares method for SBP and the two-stage residual inclusion approach [50] with robust Poisson regression in stage 2 when hypertension was the outcome [51]. We considered two sets of MR analysis that respectively used *FTO* and the GRS as instruments. Effect estimates are expressed per 1 kg/m² increase in BMI. All conventional and MR models were adjusted for age, sex, and top ten European ancestry principal components. Standard errors were obtained using a sandwich estimator [52].

The methods of accounting for antihypertensive medication described below pertain to the analysis of SBP only because the variable hypertension already takes medication into account. We considered five methods from Tobin et al. [16], the first three of each using SBP values as measured in the CLSA as the outcome. The first method (*Uncorrected SBP*) corresponds to an analysis that ignores antihypertensive medication. The second method (*Medication-adjusted SBP*) refers to models in which a binary indicator for antihypertensive medication is included as an adjustment covariate. The third method (*Treated subjects excluded SBP*) is an analysis in which participants who are treated with antihypertensive medication were excluded from the analysis. The last two

methods corresponded to transformations of the SBP variable. In the *Fixed addition (15 mmHg)* SBP method we added a constant value of 15 mmHg to measured values of SBP in treated participants. Finally, we considered a binary indicator of hypertension that combined SBP values, and the use of antihypertensive medication as described in the section on variables above.

Assessment of MR assumptions

The validity of MR depends on three key assumptions, namely that : (i) the instrument is associated with the exposure, (ii) the instrument is not associated with the outcome due to confounding, and (iii) the instrument only affects the outcome via the exposure, known as the exclusion restriction assumption [28].

For assumption (i), we assessed the strength of the instruments using the F-statistic from the linear regression of BMI on each instrument. We restricted the analysis to participants of European genetic ancestry and adjusted models for the top ten principal components to limit the likelihood of violating assumption (ii) [47].

The main source of violation of assumption (iii) is pleiotropy [53], which could involve the antihypertensive medication variable. We assessed the validity of the instruments using four methods that operate under different assumptions [54]. We first investigated the presence of directional pleiotropy using the MR-Egger intercept test on the 87 BMI-associated SNPs included in the GRS [55]. We then considered three methods applicable to single instruments (GRS or single SNPs). Specifically, the method proposed by Glymour et al. [56] (referred to as “leveraging positive confounding” method [28] hereafter) which assumes net positive confounding between the exposure and the outcome to posit that observing an MR estimate that is larger than the corresponding conventional estimate suggests the presence of an alternative path from the instrument to the outcome that does not pass through the exposure and thus may indicate instrument invalidity. Second, we adjusted the associations between the instruments and SBP/hypertension for BMI to investigate whether the estimated coefficients of the instruments were different from the null, which would suggest a violation of the restriction assumption. Finally, we used regression analysis to investigate whether the instruments were associated with the binary indicator for antihypertensive medication and/or potential confounders in a model also including BMI. Such an association would also suggest a violation of the exclusion restriction assumption.

Analyses were performed using R version 4.2.1 and Stata/IC version 14.2 (StataCorp, College Station, Texas, USA) and included analytic weights provided in the CLSA datasets.

Results

Baseline characteristics of participants and BP distribution at follow-up

Baseline characteristics of participants included and not included in our analytical sample were relatively comparable (**Table 1**). The prevalence of hypertension was 32.1% at follow-up 1 and 80% of hypertensive participants reported antihypertensive medication use. SBP was similar in subjects included in the study (118.7 mm Hg) and in those excluded (119.6 mm Hg). It was slightly higher in participants using antihypertensive medication (125.6 mmHg) than in others (116.4 mmHg).

Primary analysis

Table 2 reports the associations between BMI and SBP obtained using conventional and MR methods under the five scenarios of accounting for antihypertensive medication. A similar pattern was observed across analytical methods (i.e., conventional regression, MR with GRS as instrument or MR with *FTO* as instrument) in which adjusting models for medication led to estimates that were the closest to the null, while adding 15 mmHg to measured values of SBP led to larger estimates. Within each scenario, estimates from MR with GRS as instrument led to estimates with the smallest magnitude. Estimates obtained using MR with *FTO* as instrument were more than twice the magnitude of those obtained with GRS as the instrument. The relative variation between conventional and *FTO*-based RM estimates within a scenario was largest for the scenario excluding treated subjects, followed by the scenario adjusting for medication.

The precision of the estimates, as determined by the width of their 95% confidence intervals, was higher for conventional analysis than for MR models, as expected from the use of 2SLS with instruments [57]. MR models with the GRS as the instrument were more precise than those using *FTO* as a single instrument, reflecting the relative strength of both instruments (F statistics of 298 for GRS vs. 80 for *FTO*). As expected, the scenario in which treated subjects were excluded was associated with a lower precision than other scenarios, reflecting the exclusion of approximately a

third of the participants. The decrease in precision was larger in MR studies, especially with *FTO* as the instrument, than in conventional analyses.

In a sensitivity analysis, the MR estimates using the GRS based on all selected 97 BMI-associated SNPs were similar to those using the GRS based on 87 SNPs in all the scenarios (**Table S2**).

MR assumptions

Table 3 reports the results from the investigations of the restriction exclusion assumption using the MR-Egger intercept test and the leveraging positive confounding method. The estimated coefficients and standard errors from SNPs-BP outcomes and SNPs-BMI associations used to perform MR-Egger intercept test are provided in **Tables S3, S4, and S5**. The MR-Egger intercept test did not detect any violations in any of the scenarios of accounting for antihypertensive medication ($p>0.05$). The leveraging positive confounding method identified *FTO* as an invalid instrument for BMI across scenarios, while the GRS was considered valid in all scenarios except for the *Hypertension* scenario.

Table 4 reports the estimated associations between each of the two instruments and (i) the BP outcomes while adjusting for BMI; (ii) antihypertensive medication; and (iii) covariables associated with BMI and/or BP outcomes. Adjustment for BMI (the exposure) in models regressing the BP outcomes on the GRS led to estimates corresponding to GRS that were very close to the null and precise, suggesting that there were not path from the instrument to the outcome outside of the exposure. In opposition, adjusting for BMI in models regressing the BP outcomes on *FTO* led to estimates that were different from zero suggesting a violation of the restriction exclusion criterion. Further analyses suggested that while the GRS was not associated with the antihypertensive medication variable and any of the potential confounders of the BMI and BP outcome associations, *FTO* was associated with the antihypertensive medication variable, suggesting that at least part of the path from *FTO* to BP outcomes was passing through antihypertensive medication.

Discussion

We used data from the CLSA to compare five methods to account for antihypertensive medication in MR study designed to estimate the causal effect of BMI on SBP. We considered both a GRS and

the single *FTO* SNP as instruments for BMI and examine the impact of the methods of accounting for antihypertensive medication on the MR causal estimates and on the detection of violation of the exclusion restriction assumption. We report two main findings. First, the magnitude of the estimated MR causal effects varied across the methods of accounting for antihypertensive medication. For example, compared with the MR estimates obtained by adding 15 mmHg to SBP in treated subjects, estimates from models adjusting for medication as a covariate were 26% and 50% smaller for the MR analyses using *FTO* and the GRS as instruments, respectively. Second, the conclusion on the validity of the instruments in MR did not change depending on the method of accounting for antihypertensive medication applied.

In a simulation study, Tobin et al. [16] investigated the impact of different methods to account for antihypertensive medication in conventional regression analyses on the magnitude of the regression coefficient measuring the association between a genetic variant and SBP and its corresponding statistical power. The study suggested that imputation methods, including adding a constant value of 15 mmHg to measured SBP in treated subjects outperformed alternative approaches such as ignoring antihypertensive medication, adjusting for antihypertensive medication as a covariate in the models or excluding treated subjects for which estimates were biased towards the null. Our analysis of the CLSA using conventional regression models are consistent with the simulation study from Tobin et al. in terms of the relative magnitude of the estimated regression coefficients across methods [16]. Our MR results are also consistent with analyses from Tobin et al. [16], suggesting that the observations from the simulation study on conventional analyses may extend to MR studies.

Our results suggest that MR estimates obtained after adjusting models for medication use were attenuated in comparison to results obtained by adding a 15 mmHg to SBP, as advocated by Tobin et al. [16]. Previous studies justified adding antihypertensive medication as a covariate in models regressing continuous BP on risk factors to adjust for confounding [58], or to increase the precision of the association between risk factors and BP by explaining some of the variation in BP [50]. As others have noted [16], antihypertensive medication is likely not a confounder of the association between BMI and BP because it does not cause BMI. The use of antihypertensive medication is in fact likely impacted by risk factors for BP that affect adherence to medication, including sociodemographic factors (e.g., age, income), lifestyle behaviors (e.g. physical activity, alcohol

consumption), and medical conditions (e.g., obesity) [33,59]. Antihypertensive medication may thus appear to be an intermediate variable between either the instrument or the exposure (BMI) and SBP. This is supported by our results showing that the association between BMI and SBP is attenuated after antihypertensive medication is included in the model. If medication use is an intermediate variable on the pathway from the instrument and/or BMI and SBP, then adjusting for antihypertensive medication would be ill-advised as it may attenuate the estimated association between BMI and SBP [16]. More research is needed to better characterize the role of antihypertensive medication covariate with respects to risk factors for BP and techniques such as using multivariable MR may help to decompose the causal effect of an exposure on an outcome and estimate both the direct part and those parts mediated through other intermediate covariates [60].

The assessment of the validity of the instruments in scenarios using SBP as a continuous outcome was robust across both methods of accounting for antihypertensive medication and approaches used to detect invalid instruments. We intentionally selected *FTO* to illustrate our methods because *FTO* had previously been suspected to affect SBP through pathways other than BMI [61]. As expected, our analyses suggested that *FTO* was an invalid instrument for BMI because it was associated with SBP through pathways involving antihypertensive medication use. A plausible mechanism explaining this result is that *FTO* affects mental health conditions (e.g. dementia) independently of BMI [29-31] which in turn may affect adherence to antihypertensive medication treatment [32,33]. In contrast to *FTO*, the GRS was not identified as an invalid instrument, even if *FTO* was one of SNPs in the GRS. It thus appears that the inclusion of the SNP on *FTO* in the 87-SNP GRS does not invalidate the GRS as an instrument for BMI likely because the pleiotropic effects of *FTO* is balanced out by that of other SNPs [62].

Our GRS-based MR results are consistent with previous MR studies investigating the causal association between BMI and SBP/hypertension, regardless of the method used to correct for antihypertensive medication. Specifically, our estimated risk ratio of hypertension of 1.11 95%CI (1.07, 1.16) is close to the corresponding odds ratio (OR) reported by Larsson et al. [63] (OR=1.10) and Lee et al. [64] (OR ranging from 1.13 to 1.26 using six instruments). Estimated MR coefficients (β) obtained in scenario ignoring the use of antihypertensive medication on continuous measure of SBP ($\beta=0.85$ 95% CI(0.47, 1.24)) are also consistent with those reported in Wade et al. [9] ($\beta=0.79$)

and Holmes et al. [7] ($\beta=0.70$) who used a similar approach. Lyall et al. [11] reported estimated coefficients ranging from 1.53 to 1.65 when adjusting for medication using a covariate in the models, while Censin et al. [13] reported estimated coefficient of 3.74 in men and 3.69 in women after adding 15 mmHg in the measured SBP values in treated participants. Lyall et al. [11] and Censin et al. [13] reported their results in 1-SD units, which prevents direct comparisons with our results.

Limitations of the present study include that we restricted our investigation to five methods of accounting for antihypertensive medication. Alternative approaches such as imputation methods using censored normal regression models [16] or based on the number- or pharmacological class of drugs used [17] were not investigated because they have not been previously used in MR studies. In addition, we did not perform statistical simulations to systematically assess the impact of the methods on bias and precision of the estimates of the association between BMI and SBP in MR studies. Although our results align with the simulation study of Tobin et al. [16] suggesting that findings from conventional regression models may translate to MR studies, simulation studies are required to confirm this hypothesis. While self-reported hypertension in the CLSA may be less accurate in diagnosing hypertensive participants [65], combining both the self-reporting hypertension and the use of antihypertensive medication in defining hypertension would reduce such misclassifications. Finally, while restricting our analyses to individuals with BMI values ≤ 40 kg/m² facilitated the comparisons of methods, it may have affected the generalizability of our estimates.

In summary, MR studies of BP need to take antihypertensive medication into account and to choose the method with caution. We encourage analysis to use different approaches as sensitivity analyses and to investigate the associations between medication and instruments and/or other variables of interest, including the exposure.

Author Contributions

Conception and design of the study: F.I.N.M. and M.P.S. Data analysis: F.I.N.M. Drafting the first version of the manuscript: F.I.N.M. Critical revision of the manuscript: M.P.S. Updating previous versions of the manuscript: F.I.N.M. Reading and approval of the final manuscript: F.I.N.M. and M.P.S.

Conflicts of interest: There are no conflicts of interest.

Data Availability Statement

Data are available from the Canadian Longitudinal Study on Aging (www.clsa-elcv.ca) for researchers who meet the criteria for access to de-identified CLSA data.

Support: This research was made possible using the data/biospecimens collected by the Canadian Longitudinal Study on Aging (CLSA). Funding for the Canadian Longitudinal Study on Aging (CLSA) is provided by the Government of Canada through the Canadian Institutes of Health Research (CIHR) under grant reference: LSA 94473 and the Canada Foundation for Innovation, as well as the following provinces, Newfoundland, Nova Scotia, Quebec, Ontario, Manitoba, Alberta, and British Columbia. This research has been conducted using the CLSA dataset Comprehensive Baseline Version 6.0, Comprehensive Follow-up 1 Version 3.0 and Genome-wide Genetic Data Version 3.0, under Application Number 2010004. The CLSA is led by Drs. Parminder Raina, Christina Wolfson and Susan Kirkland. F.I.N.M. was supported by a PhD scholarship from the Fonds de recherche du Québec – Santé (FRQS) (file number 36057). M.P.S. is supported by a Chercheur-Boursier Junior II career award from the FRQS (file number 34875).

References

1. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. *Nat Rev Nephrol* 2020; 16:223-237.
2. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *Lancet* 2021; 398:957-980.
3. Sudharsanan N, Geldsetzer P. Impact of Coming Demographic Changes on the Number of Adults in Need of Care for Hypertension in Brazil, China, India, Indonesia, Mexico, and South Africa. *Hypertension* 2019; 73:770-776.
4. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008; 27:1133-1163.
5. Mbutiwi FIN, Dessy T, Sylvestre MP. Mendelian Randomization: A Review of Methods for the Prevention, Assessment, and Discussion of Pleiotropy in Studies Using the Fat Mass and Obesity-Associated Gene as an Instrument for Adiposity. *Front Genet* 2022; 13:803238.
6. Dale CE, Fatemifar G, Palmer TM, White J, Prieto-Merino D, Zabaneh D, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. *Circulation* 2017; 135:2373-2388.
7. Holmes MV, Lange LA, Palmer T, Lanktree MB, North KE, Almqvister B, et al. Causal effects of body mass index on cardiometabolic traits and events: a Mendelian randomization analysis. *Am J Hum Genet* 2014; 94:198-208.
8. Kivimaki M, Smith GD, Timpson NJ, Lawlor DA, Batty GD, Kahonen M, et al. Lifetime body mass index and later atherosclerosis risk in young adults: examining causal links using Mendelian randomization in the Cardiovascular Risk in Young Finns study. *Eur Heart J* 2008; 29:2552-2560.
9. Wade KH, Chiesa ST, Hughes AD, Chaturvedi N, Charakida M, Rapala A, et al. Assessing the causal role of body mass index on cardiovascular health in young adults: Mendelian randomization and recall-by-genotype analyses. *Circulation* 2018; 138:2187-2201.

10. Wurtz P, Wang Q, Kangas AJ, Richmond RC, Skarp J, Tiainen M, et al. Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. *PLoS Med* 2014; 11:e1001765.
11. Lyall DM, Celis-Morales C, Ward J, Iliodromiti S, Anderson JJ, Gill JMR, et al. Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank A Mendelian Randomization Study. *JAMA Cardiol* 2017; 2:882-889.
12. Fall T, Hagg S, Magi R, Ploner A, Fischer K, Horikoshi M, et al. The Role of Adiposity in Cardiometabolic Traits: A Mendelian Randomization Analysis. *PLoS Med* 2013; 10:e1001474.
13. Censin JC, Peters SAE, Bovijn J, Ferreira T, Pulit SL, Magi R, et al. Causal relationships between obesity and the leading causes of death in women and men. *PLoS Genet* 2019; 15:e1008405.
14. Fall T, Hagg S, Ploner A, Magi R, Fischer K, Draisma HH, et al. Age- and sex-specific causal effects of adiposity on cardiovascular risk factors. *Diabetes* 2015; 64:1841-1852.
15. Timpson NJ, Harbord R, Davey Smith G, Zacho J, Tybjaerg-Hansen A, Nordestgaard BG. Does greater adiposity increase blood pressure and hypertension risk?: Mendelian randomization using the FTO/MC4R genotype. *Hypertension* 2009; 54:84-90.
16. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med* 2005; 24:2911-2935.
17. Rana BK, Dhamija A, Panizzon MS, Spoon KM, Vasilopoulos T, Franz CE, et al. Imputing observed blood pressure for antihypertensive treatment: impact on population and genetic analyses. *Am J Hypertens* 2014; 27:828-837.
18. DeGuire J, Clarke J, Rouleau K, Roy J, Bushnik T. Blood pressure and hypertension. *Health Rep* 2019; 30:14-21.
19. Tanamas SK, Hanson RL, Nelson RG, Knowler WC. Effect of different methods of accounting for antihypertensive treatment when assessing the relationship between diabetes or obesity and systolic blood pressure. *J Diabetes Complications* 2017; 31:693-699.

20. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 2011; 478:103-109.
21. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, et al. Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet* 2009; 41:666-676.
22. Sabatti C, Service SK, Hartikainen A-L, Pouta A, Ripatti S, Brodsky J, et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* 2009; 41:35-46.
23. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet* 2018; 50:1412-1425.
24. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet* 2017; 49:403-415.
25. Padmanabhan S, Melander O, Johnson T, Di Blasio AM, Lee WK, Gentilini D, et al. Genome-wide association study of blood pressure extremes identifies variant near UMOD associated with hypertension. *PLoS Genet* 2010; 6:e1001177.
26. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, et al. Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet* 2009; 41:666-676.
27. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol* 2014; 43:922-929.
28. Labrecque J, Swanson SA. Understanding the Assumptions Underlying Instrumental Variable Analyses: a Brief Review of Falsification Strategies and Related Tools. *Curr Epidemiol Rep* 2018; 5:214-220.
29. Walter S, Glymour MM, Koenen K, Liang L, Tchetgen Tchetgen EJ, Cornelis M, et al. Do genetic risk scores for body mass index predict risk of phobic anxiety? Evidence for a shared genetic risk factor. *Psychol Med* 2015; 45:181-191.

30. Kivimäki M, Jokela M, Hamer M, Geddes J, Ebmeier K, Kumari M, et al. Examining overweight and obesity as risk factors for common mental disorders using fat mass and obesity-associated (FTO) genotype-instrumented analysis: The Whitehall II Study, 1985-2004. *Am J Epidemiol* 2011; 173:421-429.
31. Keller L, Xu W, Wang HX, Winblad B, Fratiglioni L, Graff C. The obesity related gene, FTO, interacts with APOE, and is associated with Alzheimer's disease risk: a prospective cohort study. *J Alzheimers Dis* 2011; 23:461-469.
32. Calderón-Larrañaga A, Diaz E, Poblador-Plou B, Gimeno-Feliu LA, Abad-Díez JM, Prados-Torres A. Non-adherence to antihypertensive medication: The role of mental and physical comorbidity. *Int J Cardiol* 2016; 207:310-316.
33. Burnier M, Egan BM. Adherence in Hypertension. *Circ Res* 2019; 124:1124-1140.
34. Gill R, Stratigopoulos G, Lee JH, Leibel RL. Functional genomic characterization of the FTO locus in African Americans. *Physiol Genomics* 2019; 51:517-528.
35. Raina P, Wolfson C, Kirkland S, Griffith LE, Balion C, Cossette B, et al. Cohort Profile: The Canadian Longitudinal Study on Aging (CLSA). *Int J Epidemiol* 2019; 48:1752-1753j.
36. Raina PS, Wolfson C, Kirkland S. Canadian Longitudinal Study on Aging (CLSA). Protocol (version 3). Hamilton, ON, Canada: CLSA; 2008.
37. Staley JR, Burgess S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. *Genet Epidemiol* 2017; 41:341-352.
38. Canadian Longitudinal Study on Aging. Pulse Rate and Blood Pressure Measurement. Document no SOP_DCS_0008, version 2.4; 2017 [cited 15 Aug 2022]. Available from: <https://www.clsa-elcv.ca/doc/2303>.
39. van de Vijver SJM, Oti SO, Agyemang C, Gomez GB, Kyobutungi C. Prevalence, awareness, treatment and control of hypertension among slum dwellers in Nairobi, Kenya. *J Hypertens* 2013; 31:1018-1024.
40. Forgetta V, Li R, Darmond-Zwaig C, Belisle A, Balion C, Roshandel D, et al. Cohort profile: genomic data for 26 622 individuals from the Canadian Longitudinal Study on Aging (CLSA). *BMJ Open* 2022; 12:e059021.

41. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; 518:197-206.
42. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol* 2013; 42:1134-1144.
43. Costanian C, Edgell H, Ardern CI, Tamim H. Hormone therapy use in the Canadian Longitudinal Study on Aging: a cross-sectional analysis. *Menopause* 2018; 25:46-53.
44. Griffith LE, Raina P, Kanters D, Hogan D, Patterson C, Papaioannou A, et al. Frailty differences across population characteristics associated with health inequality: a cross-sectional analysis of baseline data from the Canadian Longitudinal Study on Aging (CLSA). *BMJ Open* 2021; 11:e047945.
45. Mian O, Anderson LN, Belsky DW, Gonzalez A, Ma J, Sloboda DM, et al. Associations of Adverse Childhood Experiences with Frailty in Older Adults: A Cross-Sectional Analysis of Data from the Canadian Longitudinal Study on Aging. *Gerontology* 2021:1-10.
46. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol* 1993; 46:153-162.
47. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; 38:904-909.
48. Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *Am J Epidemiol* 2004; 159:702-706.
49. Chen W, Qian L, Shi J, Franklin M. Comparing performance between log-binomial and robust Poisson regression models for estimating risk ratios under model misspecification. *BMC Med Res Methodol* 2018; 18:63.
50. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* 2017; 26:2333-2355.
51. Wang J. Dose as Instrumental Variable in Exposure–Safety Analysis Using Count Models. *J Biopharm Stat* 2012; 22:565-581.
52. Hardin JW. The Robust Variance Estimator for Two-stage Models. *Stata J* 2002; 2:253-266.
53. Dixon P, Hollingworth W, Harrison S, Davies NM, Smith GD. Mendelian Randomization analysis of the causal effect of adiposity on hospital costs. *J Health Econ* 2020; 70:102300.

54. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet* 2018; 27:R195-R208.
55. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* 2017; 32:377-389.
56. Glymour MM, Tchetgen Tchetgen EJ, Robins JM. Credible mendelian randomization studies: Approaches for evaluating the instrumental variable assumptions. *Am J Epidemiol* 2012; 175:332-339.
57. Burgess S, Thompson SG. Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation. Boca Raton (FL): Chapman and Hall/CRC; 2015.
58. Sun S, Lulla A, Sioda M, Winglee K, Wu MC, Jacobs DR, et al. Gut Microbiota Composition and Blood Pressure. *Hypertension* 2019; 73:998-1006.
59. Abbas H, Kurdi M, de Vries F, van Onzenoort HAW, Driessen JHM, Watfa M, et al. Factors Associated with Antihypertensive Medication Non-Adherence: A Cross-Sectional Study Among Lebanese Hypertensive Adults. *Patient Prefer Adherence* 2020; 14:663-673.
60. Sanderson E. Multivariable Mendelian Randomization and Mediation. *Cold Spring Harb Perspect Med* 2021; 11:a038984.
61. Pausova Z, Syme C, Abrahamowicz M, Xiao Y, Leonard GT, Perron M, et al. A common variant of the FTO gene is associated with not only increased adiposity but also elevated blood pressure in french canadians. *Circ Cardiovasc Genet* 2009; 2:260-269.
62. Davey Smith G. Random allocation in observational data: how small but robust effects could facilitate hypothesis-free causal inference. *Epidemiology* 2011; 22:460-463.
63. Larsson SC, Back M, Rees JMB, Mason AM, Burgess S. Body mass index and body composition in relation to 14 cardiovascular conditions in UK Biobank: a Mendelian randomization study. *Eur Heart J* 2020; 41:221-226.
64. Lee MR, Lim YH, Hong YC. Causal association of body mass index with hypertension using a Mendelian randomization design. *Medicine (Baltimore)* 2018; 97:e11252.
65. Gonçalves VSS, Andrade KRC, Carvalho KMB, Silva MT, Pereira MG, Galvao TF. Accuracy of self-reported hypertension: a systematic review and meta-analysis. *J Hypertens* 2018; 36:970-978.

Table 1

Baseline characteristics of participants by study inclusion status, CLSA Comprehensive cohort, 2011-2018.

Characteristic	Included participants			Excluded participants		
	n	% ^a	Mean (SD) ^a	n	% ^a	Mean (SD) ^a
Total	20,430	100	-	9,667	100	-
Age, years	20,430		59.3 (9.7)	9,667		59.7 (10.5)
Sex, Male	10,328	50.5		4,449	46.3	
Marital status						
With partner	14,557	78.6		6,094	72.1	
No partner	5,873	21.4		3,565	27.8	
Refused	0	0.0		8	0.1	
Graduate from high/secondary school						
Yes	18,486	92.0		8,361	88.5	
No	1,944	8.0		1,268	11.3	
Don't know/No answer	0	0.0		36	0.3	
Missing	0	0.0		2	0.0	
Current/former daily smoker						
Yes	9,616	44.8		4,483	43.9	
No	10,814	55.2		5,008	54.4	
Don't know/No answer	0	0.0		175	1.7	
Refused	0	0.0		1	0.0	
Alcohol consumption						
Non-drinker	4,498	20.9		3,347	33.4	
Drinker	15,932	79.1		6,307	66.5	
Don't know/No answer	0	0.0		11	0.1	
Refused	0	0.0		2	0.0	
PASE score	20,430		154.4 (76.3)	8,034		140.5 (79.5)
Missing	0		-	1,633		-
BMI, kg/m ²	20,430		27.4 (4.4)	9,531		29.4 (7.1)
Missing	0		-	136		-

Abbreviations: SD, standard deviation; PASE, Physical Activity Scale for the Elderly; BMI, body mass index.

^a Percentage, mean, and SD are estimated using analytic weights.

Table 2

Conventional and MR estimates of the association between body mass index (BMI) with blood pressure outcomes in different scenarios of accounting for antihypertensive medication effects, CLSA Comprehensive cohort, 2011-2018.

	n	Estimates ^a		
		Conventional	MR using GRS as instrument	MR using <i>FTO</i> as instrument
<i>SBP (mmHg) outcome</i>				
Uncorrected SBP ^b	20,430	1.01 (0.96, 1.05)	0.85 (0.45, 1.26)	2.16 (1.30, 3.01)
Medication-adjusted SBP ^c	20,430	0.90 (0.86, 0.95)	0.68 (0.24, 1.13)	2.09 (1.14, 3.05)
Treated subjects excluded SBP ^d	14,337	1.11 (1.05, 1.16)	0.90 (0.40, 1.41)	2.76 (1.55, 3.98)
Fixed addition (15 mmHg) SBP ^e	20,430	1.36 (1.31, 1.41)	1.35 (0.89, 1.81)	2.82 (1.84, 3.81)
<i>Binary outcome</i>				
Hypertension ^f	20,430	1.09 (1.08, 1.09)	1.11 (1.07, 1.16)	1.15 (1.07, 1.24)

Abbreviations: MR, mendelian randomization; GRS, genetic risk score; *FTO*, fat mass and obesity-associated (genotype); SBP, systolic blood pressure.

^a Conventional estimates are from linear (SBP as outcome) and robust Poisson (hypertension as outcome) regressions. MR estimates used the two-stage least squares method (SBP as outcome) and the two-stage residual inclusion approach (hypertension as outcome) with robust Poisson regression at the second stage. All models used analytic weights and were adjusted for age, sex, and top ten European ancestry principal components. Effect estimates represent the beta coefficient for SBP (mmHg) and risk ratio for hypertension (95% confidence interval) per 1 kg/m² increase in BMI.

^b Measured SBP was the outcome, and the use of antihypertensive medication issue ignored.

^c Measured SBP was the outcome, and the use of antihypertensive medication included in the models as an adjustment covariate.

^d Measured SBP was the outcome, and participants using antihypertensive medication were excluded from the analysis.

^e Consisted of adding a constant value of 15 mmHg to the measured SBP in treated subjects.

^f Hypertension was defined as (i) SBP \geq 140 mmHg or diastolic blood pressure \geq 90 mmHg, and/or (ii) self-reported diagnosis of hypertension and current use of antihypertensive medication.

Table 3

Assessment of the validity of the BMI-associated instruments in different scenarios of accounting for antihypertensive medication using MR-Egger intercept test and the leveraging positive confounding method, CLSA Comprehensive cohort, 2011-2018.

	n	MR-Egger intercept test ^a			Leveraging positive confounding method ^b				
		β	P value	Suspicion of directional pleiotropy	Conventional estimates	GRS as instrument		FTO as instrument	
						MR estimates	Suspicion of invalidity	MR estimates	Suspicion of invalidity
<i>SBP (mmHg) outcome</i>									
Uncorrected SBP ^c	20,430	-0.02	0.600	No	1.01	0.85	No	2.16	Yes
Medication-adjusted SBP ^d	20,430	-0.00	0.921	No	0.90	0.68	No	2.09	Yes
Treated subjects excluded SBP ^e	14,337	-0.03	0.497	No	1.11	0.90	No	2.76	Yes
Fixed addition (15 mmHg) SBP ^f	20,430	-0.02	0.737	No	1.36	1.35	No	2.82	Yes
<i>Binary outcome</i>									
Hypertension ^g	20,430	0.00	0.425	No	1.09	1.11	Yes	1.15	Yes

Abbreviations: GRS, genetic risk score; *FTO*, fat mass and obesity-associated (gene); MR, Mendelian randomization; SBP, systolic blood pressure.

^a The MR-Egger intercept test is based on one-sample summary statistics from associations between each of the 87 single-nucleotide polymorphisms included in the GRS.

^b States that the instrument is invalid if the MR estimate is larger than the conventional estimate assuming net positive confounding between the exposure (BMI) and the outcome (SBP/hypertension) [56]. Conventional and MR estimates represent the beta coefficient for SBP (mmHg) and risk ratio for hypertension per 1 kg/m² increase in BMI (reported in Table 2).

^c Measured SBP was the outcome, and the use of antihypertensive medication issue ignored.

^d Measured SBP was the outcome, and the use of antihypertensive medication included in the models as an adjustment covariate. Summary statistics (beta and standard errors) used to perform MR-Egger intercept test were derived from the SNPs-BMI and SNPs-measured SBP associations both adjusted for antihypertensive medication (Table S4).

^e Measured SBP was the outcome, and participants using antihypertensive medication were excluded from the analysis.

^f Consisted of adding a constant value of 15 mmHg to the measured SBP in treated subjects.

^g Hypertension was defined as (i) SBP \geq 140 mmHg or diastolic blood pressure \geq 90 mmHg, and/or (ii) self-reported diagnosis of hypertension and current use of antihypertensive medication.

Note: All analyses used analytic weights.

Table 4

Associations between the instruments and blood pressure outcomes in different scenarios of accounting for antihypertensive medication, antihypertensive medication covariate, and potential confounders, CLSA Comprehensive cohort, 2011-2018.

Trait/analysis ^a	n	GRS		FTO	
		β (95% CI)	P value	β (95% CI)	P value
<i>BP outcomes in models adjusted for BMI</i>					
Uncorrected SBP, mmHg ^b	20,430	-0.02 (-0.06, 0.02)	0.261	0.50 (0.20, 0.80)	1.21 x 10 ⁻⁰³
Medication-adjusted SBP, mmHg ^c	20,430	-0.02 (-0.06, 0.01)	0.184	0.44 (0.14, 0.73)	3.78 x 10 ⁻⁰³
Treated subjects excluded SBP, mmHg ^d	14,337	-0.03 (-0.07, 0.01)	0.194	0.56 (0.22, 0.91)	1.16 x 10 ⁻⁰³
Fixed addition (15 mmHg) SBP, mmHg ^e	20,430	-0.01 (-0.05, 0.03)	0.525	0.63 (0.28, 0.98)	3.99 x 10 ⁻⁰⁴
Hypertension ^f	20,430	0.00 (-0.00, 0.01)	0.556	0.04 (-0.01, 0.08)	0.082
<i>Antihypertensive medication</i>					
Unadjusted model	20,430	0.01 (0.01, 0.02)	1.99 x 10 ⁻⁰⁷	0.10 (0.05, 0.14)	2.29 x 10 ⁻⁰⁵
Model adjusted for BMI	20,430	0.00 (-0.00, 0.01)	0.326	0.05 (0.00, 0.10)	0.044
<i>Potential confounders</i>					
Age, years	20,430	-0.02 (-0.04, 0.01)	0.146	0.07 (-0.12, 0.26)	0.489
Sex, Male	20,430	0.00 (-0.00, 0.01)	0.181	0.03 (-0.01, 0.07)	0.149
Marital status, No partner	20,430	0.00 (-0.01, 0.01)	0.729	0.03 (-0.02, 0.07)	0.299
Alcohol consumption, Drinker	20,430	-0.00 (-0.01, 0.01)	0.740	-0.00 (-0.05, 0.05)	0.965
Current/former daily smoker	20,430	0.00 (-0.00, 0.01)	0.168	0.01 (-0.04, 0.05)	0.800
No graduate from high/secondary school	20,430	-0.00 (-0.01, 0.01)	0.848	-0.06 (-0.13, 0.01)	0.105
PASE score	20,430	0.02 (-0.16, 0.19)	0.868	0.27 (-1.25, 1.79)	0.729

Abbreviations: GRS, genetic risk score; FTO, fat mass and obesity-associated (genotype); SBP, systolic blood pressure; BMI, body mass index; PASE, Physical Activity Scale for the Elderly.

^a Associations were estimated using linear or logistic regression, as appropriated, with BMI-instrument being the exposure. Estimates applied analytic weights and are given per 1 effect allele increase in GRS/FTO rs1558902 genotype.

^b Measured SBP was the outcome, and the use of antihypertensive medication issue ignored.

^c Measured SBP was the outcome, and the use of antihypertensive medication included in the models as an adjustment covariate.

^d Measured SBP was the outcome, and participants using antihypertensive medication were excluded from the analysis.

^e Consisted of adding a constant value of 15 mmHg to the measured SBP in treated subjects.

^f Hypertension was defined as (i) SBP \geq 140 mmHg or diastolic blood pressure \geq 90 mmHg, and/or (ii) self-reported diagnosis of hypertension and current use of antihypertensive medication.

Online Supplemental Data

Supplement to: Accounting for antihypertensive medication in Mendelian randomization studies of blood pressure: methodological considerations in the Canadian Longitudinal Study on Aging

Fiston Ikwa Ndol Mbutiwi^{a,b} and Marie-Pierre Sylvestre^{a,c}

^a University of Montreal Hospital Research Centre (CRCHUM), Montréal, Quebec, Canada.

^b Faculty of Medicine, University of Kikwit, Kikwit, Democratic Republic of the Congo.

^c Department of Social and Preventive Medicine, University of Montreal Public Health School (ESPUM), Montréal, Canada.

Table of Contents	Page
Figure S1. Flowchart of selection of study participants, CLSA Comprehensive cohort, 2011-2018.....	205
Table S1. Associations of 97 BMI-associated SNPs with BMI, CLSA Comprehensive cohort, 2011-2018 (n = 20,430).....	206
Table S2. MR estimates of the association between BMI with blood pressure outcomes in different scenarios of accounting for antihypertensive medication using the GRS based on 97 BMI-associated SNPs as instrument, CLSA Comprehensive cohort, 2011-2018.....	209
Table S3. Associations of 87 BMI-associated SNPs included in the genetic risk score with BMI and with SBP/hypertension, CLSA Comprehensive cohort, 2011-2018 (n = 20,430).....	210
Table S4. Associations of 87 BMI-associated SNPs included in the GRS with BMI and with measured SBP both adjusted for antihypertensive medication covariate, CLSA Comprehensive cohort, 2011-2018 (n = 20,430).....	214
Table S5. Associations of 87 BMI-associated SNPs included in the GRS with BMI and with measured SBP after exclusion of participants using antihypertensive medication, CLSA Comprehensive cohort, 2011-2018 (n = 14,337).....	217
Reference	220

Figure S1

Flowchart of selection of study participants, CLSA Comprehensive cohort, 2011-2018.

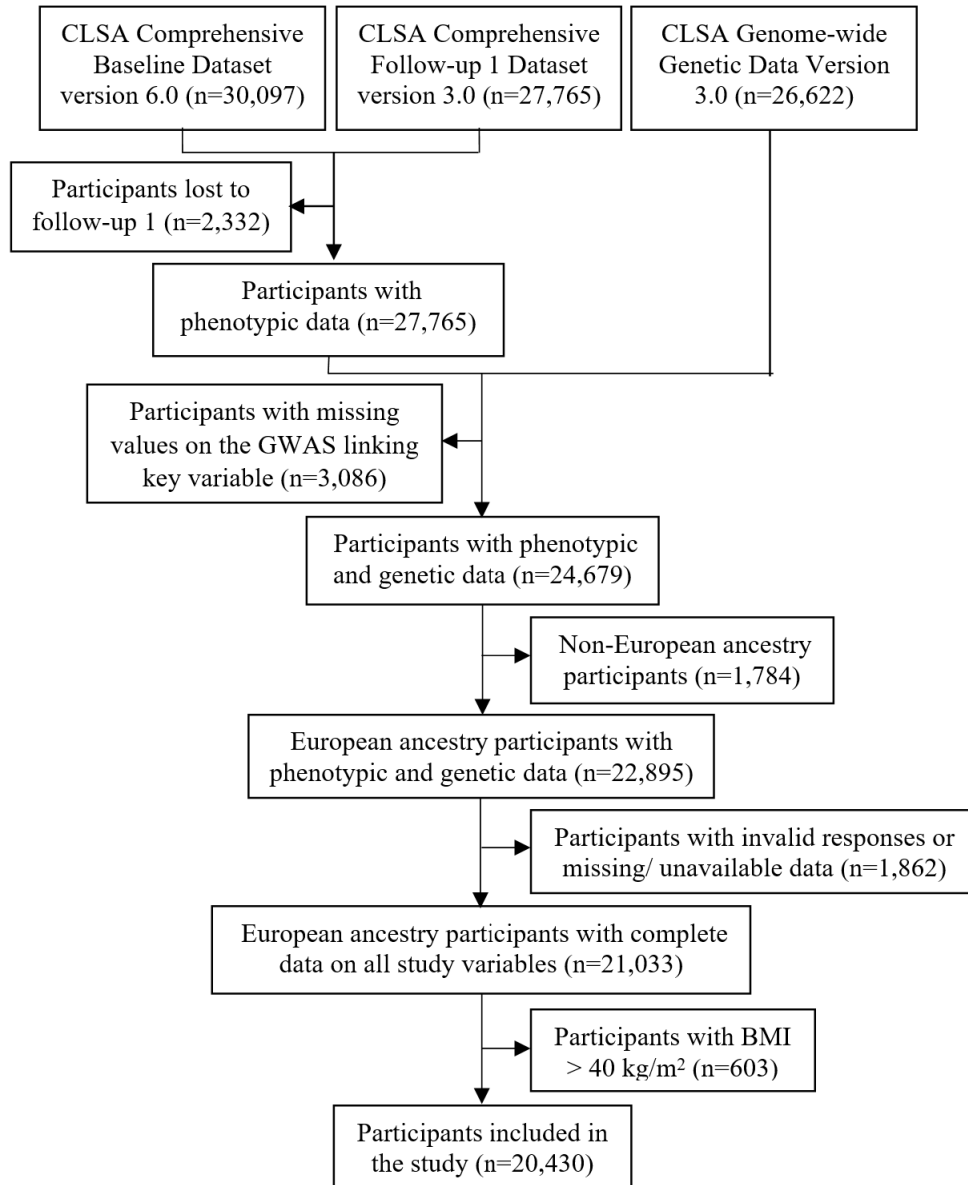


Table S1

Associations of 97 BMI-associated SNPs with BMI, CLSA Comprehensive cohort, 2011-2018 (n = 20,430).

id	SNP ^a	Chr	Position ^b	EA	OA	EAF	BMI, kg/m ²		
							β	SE	P value
1	rs1558902	16	53 803 574	A	T	0.399	0.402	0.045	3.63E-19
2	rs6567160	18	57 829 135	C	T	0.231	0.333	0.052	2.01E-10
3	rs13021737	2	632 348	G	A	0.821	0.159	0.057	5.13E-03
4	rs10938397	4	45 182 527	G	A	0.435	0.083	0.044	6.12E-02
5	rs543874	1	177 889 480	G	A	0.194	0.167	0.055	2.45E-03
6	rs2207139	6	50 845 490	G	A	0.171	0.150	0.058	1.01E-02
7	rs11030104	11	27 684 517	A	G	0.796	0.218	0.054	5.70E-05
8	rs3101336	1	72 751 185	C	T	0.623	0.095	0.045	3.48E-02
9	rs7138803	12	50 247 468	A	G	0.381	0.159	0.045	4.11E-04
10	rs10182181	2	25 150 296	G	A	0.476	0.134	0.044	2.43E-03
11	rs3888190	16	28 889 486	A	C	0.389	0.167	0.045	1.99E-04
12	rs1516725	3	185 824 004	C	T	0.865	0.257	0.064	5.57E-05
13	rs12446632	16	19 935 389	G	A	0.856	0.129	0.063	3.95E-02
14	rs2287019	19	46 202 172	C	T	0.811	0.140	0.056	1.20E-02
15	rs16951275	15	68 077 168	T	C	0.780	0.237	0.053	7.58E-06
16	rs3817334	11	47 650 993	T	C	0.405	0.142	0.045	1.48E-03
17	rs2112347	5	75 015 242	T	G	0.633	0.128	0.046	5.14E-03
18	rs12566985	1	75 002 193	G	A	0.429	0.049	0.044	2.69E-01
19	rs3810291	19	47 569 003	A	G	0.679	0.160	0.047	7.69E-04
20	rs7141420	14	79 899 454	T	C	0.524	0.121	0.044	5.60E-03
21	rs13078960	3	85 807 590	G	T	0.209	0.035	0.054	5.20E-01
22	rs10968576	9	28 414 339	G	A	0.315	0.074	0.047	1.17E-01
23	rs17024393	1	110 154 688	C	T	0.026	0.332	0.136	1.45E-02
24	rs657452	1	49 589 847	A	G	0.398	0.057	0.045	2.09E-01
25	rs12429545	13	54 102 206	A	G	0.119	0.111	0.068	1.01E-01
26	rs12286929	11	115 022 404	G	A	0.523	0.031	0.044	4.77E-01
27	rs13107325	4	103 188 709	T	C	0.071	0.265	0.085	1.88E-03
28	rs7903146	10	114 758 349	C	T	0.705	0.068	0.048	1.58E-01
29	rs10132280	14	25 928 179	C	A	0.702	0.103	0.048	3.03E-02
30	rs17405819	8	76 806 584	T	C	0.693	0.078	0.047	9.96E-02
31	rs6091540	20	51 087 862	C	T	0.711	0.027	0.048	5.77E-01
32	rs1016287	2	59 305 625	T	C	0.290	0.166	0.048	6.20E-04

33	rs4256980	11	8 673 939	G	C	0.644	0.031	0.046	4.95E-01
34	rs17094222	10	102 395 440	C	T	0.205	0.096	0.054	7.66E-02
35	rs12401738	1	78 446 761	A	G	0.357	0.071	0.046	1.21E-01
36	rs7599312	2	213 413 231	G	A	0.735	0.035	0.050	4.75E-01
37	rs2365389	3	61 236 462	C	T	0.585	0.107	0.045	1.64E-02
38	rs205262	6	34 563 164	G	A	0.276	0.040	0.049	4.14E-01
39	rs2820292	1	201 784 287	C	A	0.561	0.083	0.044	6.07E-02
40	rs12885454	14	29 736 838	C	A	0.649	0.099	0.046	3.13E-02
41	rs12016871	13	28 017 782	T	C	0.182	0.105	0.057	6.37E-02
42	rs16851483	3	141 275 436	T	G	0.064	0.200	0.090	2.65E-02
43	rs1167827	7	75 163 169	G	A	0.563	0.066	0.044	1.37E-01
44	rs758747	16	3 627 358	T	C	0.293	0.063	0.048	1.87E-01
45	rs1928295	9	120 378 483	T	C	0.561	0.051	0.044	2.46E-01
46	rs9925964	16	31 129 895	A	G	0.623	0.028	0.045	5.38E-01
47	rs11126666	2	26 928 811	A	G	0.265	0.066	0.050	1.83E-01
48	rs2650492	16	28 333 411	A	G	0.293	0.177	0.048	2.51E-04
49	rs6804842	3	25 106 437	G	A	0.585	0.044	0.044	3.23E-01
50	rs12940622	17	78 615 571	G	A	0.567	0.091	0.044	3.96E-02
51	rs7164727	15	73 093 991	T	C	0.683	0.018	0.047	7.09E-01
52	rs4740619	9	15 634 326	T	C	0.549	0.019	0.044	6.70E-01
53	rs492400	2	219 349 752	C	T	0.423	0.128	0.044	3.99E-03
54	rs13191362	6	163 033 350	A	G	0.874	0.066	0.066	3.16E-01
55	rs17001654	4	77 129 568	G	C	0.156	0.065	0.061	2.87E-01
56	rs11191560	10	104 869 038	C	T	0.082	0.107	0.080	1.79E-01
57	rs2080454	16	49 062 590	C	A	0.381	0.055	0.045	2.26E-01
58	rs2176040	2	227 092 802	A	G	0.361	0.003	0.046	9.39E-01
59	rs1000940	17	5 283 252	G	A	0.295	0.024	0.048	6.14E-01
60	rs2033529	6	40 348 653	G	A	0.281	0.061	0.049	2.13E-01
61	rs11583200	1	50 559 820	C	T	0.387	0.056	0.045	2.17E-01
62	rs7239883	18	40 147 671	G	A	0.393	0.095	0.045	3.37E-02
63	rs2836754	21	40 291 740	C	T	0.626	0.017	0.045	7.00E-01
64	rs9400239	6	108 977 663	C	T	0.695	0.119	0.048	1.23E-02
65	rs10733682	9	129 460 914	A	G	0.478	0.021	0.044	6.40E-01
66	rs11688816	2	63 053 048	G	A	0.526	0.017	0.044	6.95E-01
67	rs11057405	12	122 781 897	G	A	0.897	0.042	0.072	5.57E-01
68	rs9914578	17	2 005 136	G	C	0.198	0.104	0.055	5.78E-02
69	rs977747	1	47 684 677	T	G	0.404	0.060	0.045	1.81E-01
70	rs2121279	2	143 043 285	T	C	0.124	0.176	0.066	7.90E-03

71	rs29941	19	34 309 532	G	A	0.679	0.145	0.047	1.96E-03
72	rs11727676	4	145 659 064	T	C	0.906	0.142	0.075	5.81E-02
73	rs3849570	3	81 792 112	A	C	0.342	0.039	0.046	4.00E-01
74	rs9374842	6	120 185 665	T	C	0.762	0.050	0.051	3.35E-01
75	rs4787491	16	30 015 337	G	A	0.531	0.080	0.044	6.63E-02
76	rs1441264	13	79 580 919	A	G	0.591	0.115	0.045	1.00E-02
77	rs7899106	10	87410904	G	A	0.045	0.317	0.105	2.45E-03
78	rs2176598	11	43 864 278	T	C	0.246	0.116	0.051	2.41E-02
79	rs2245368	7	76 608 143	C	T	0.170	0.083	0.058	1.54E-01
80	rs17203016	2	208 255 518	G	A	0.194	0.046	0.056	4.09E-01
81	rs7243357	18	56 883 319	T	G	0.825	0.120	0.058	3.92E-02
82	rs16907751	8	81 375 457	C	T	0.895	0.161	0.071	2.38E-02
83	rs1808579	18	21 104 888	C	T	0.517	0.049	0.044	2.65E-01
84	rs13201877	6	137 675 541	G	A	0.135	0.089	0.064	1.66E-01
85	rs2033732	8	85 079 709	C	T	0.746	0.007	0.050	8.95E-01
86	rs9540493	13	66 205 704	A	G	0.443	0.017	0.044	6.98E-01
87	rs1460676	2	164 567 689	C	T	0.157	0.071	0.060	2.40E-01
88	rs11165643	1	96 924 097	T	C	0.589	-0.005	0.045	9.08E-01
89	rs9641123	7	93 197 732	C	G	0.413	-0.024	0.045	5.84E-01
90	rs3736485	15	51 748 610	A	G	0.456	-0.005	0.044	9.06E-01
91	rs7715256	5	153 537 893	G	T	0.419	-0.024	0.044	5.89E-01
92	rs2075650	19	45 395 619	A	G	0.859	-0.052	0.063	4.09E-01
93	rs6477694	9	111 932 342	C	T	0.349	-0.012	0.046	7.95E-01
94	rs17724992	19	18 454 825	A	G	0.733	-0.050	0.050	3.17E-01
95	rs11847697	14	30 515 112	T	C	0.042	-0.197	0.110	7.47E-02
96	rs1528435	2	181 550 962	T	C	0.618	-0.035	0.045	4.31E-01
97	rs6465468	7	95 169 514	T	G	0.316	-0.074	0.047	1.20E-01

Abbreviations: SNP, single-nucleotide polymorphism; Chr, chromosome; EA, effect allele; OA, other allele; EAF, effect allele frequency; BMI, body mass index; SE, standard error.

^a BMI-associated SNPs are those reported by Locke et al. (2015) [1].

^b Positions are reported in reference to human genome build GRCh37/hg19 and effect allele harmonized with Locke et al. (2015) [1].

Notes: Estimates are from linear regression models of BMI on each SNP using analytic weights.

Table S2

MR estimates of the association between BMI with blood pressure outcomes in different scenarios of accounting for antihypertensive medication using the GRS based on 97 BMI-associated SNPs as instrument, CLSA Comprehensive cohort, 2011-2018.

	n	Effect estimates^a
<i>SBP (mmHg) outcome</i>		
Uncorrected SBP ^b	20,430	0.84 (0.42, 1.26)
Medication-adjusted SBP ^c	20,430	0.67 (0.21, 1.12)
Treated subjects excluded SBP ^d	14,337	0.88 (0.36, 1.41)
Fixed addition (15 mmHg) SBP ^e	20,430	1.36 (0.88, 1.83)
<i>Binary outcome</i>		
Hypertension ^f	20,430	1.12 (1.07, 1.16)

Abbreviation: SBP, systolic blood pressure.

^a MR estimates used the two-stage least squares method (SBP as outcome) and the two-stage residual inclusion approach (hypertension as outcome) with robust Poisson regression at the second stage. All models used analytic weights and were adjusted for age, sex, and top ten European ancestry principal components. Effect estimates represent the beta coefficient for SBP (mmHg) and risk ratio for hypertension (95% confidence interval) per 1 kg/m² increase in BMI.

^b Measured SBP was the outcome, and the use of antihypertensive medication issue ignored.

^c Measured SBP was the outcome, and the use of antihypertensive medication included in the models as an adjustment covariate.

^d Measured SBP was the outcome, and participants using antihypertensive medication were excluded from the analysis.

^e Consisted of adding a constant value of 15 mmHg to the measured SBP in treated subjects.

^f Hypertension was defined as (i) SBP \geq 140 mmHg or diastolic blood pressure \geq 90 mmHg, and/or (ii) self-reported diagnosis of hypertension and current use of antihypertensive medication.

Table S3

Associations of 87 BMI-associated SNPs included in the genetic risk score with BMI and with SBP/hypertension, CLSA Comprehensive cohort, 2011-2018 (n = 20,430).

id	SNP	BMI, kg/m ²			Uncorrected SBP, mmHg ^a			Fixed addition (15 mmHg) SBP, mmHg ^b			Hypertension ^c		
		β	SE	P value	β	SE	P value	β	SE	P value	β	SE	P value
1	rs1558902	0.402	0.045	3.63E-19	0.924	0.160	8.08E-09	1.204	0.188	1.63E-10	0.062	0.016	1.24E-04
2	rs6567160	0.333	0.052	2.01E-10	0.113	0.187	5.46E-01	0.228	0.219	3.00E-01	0.022	0.019	2.48E-01
3	rs13021737	0.159	0.057	5.13E-03	0.338	0.203	9.55E-02	0.433	0.239	6.93E-02	0.036	0.021	7.96E-02
4	rs10938397	0.083	0.044	6.12E-02	-0.077	0.158	6.28E-01	0.007	0.185	9.71E-01	0.011	0.016	4.75E-01
5	rs543874	0.167	0.055	2.45E-03	-0.161	0.196	4.13E-01	-0.163	0.231	4.80E-01	0.002	0.020	9.25E-01
6	rs2207139	0.150	0.058	1.01E-02	0.305	0.208	1.43E-01	0.413	0.245	9.14E-02	0.022	0.021	2.81E-01
7	rs11030104	0.218	0.054	5.70E-05	0.033	0.193	8.64E-01	-0.005	0.227	9.82E-01	-0.017	0.020	3.98E-01
8	rs3101336	0.095	0.045	3.48E-02	0.073	0.160	6.50E-01	0.113	0.188	5.48E-01	0.003	0.016	8.63E-01
9	rs7138803	0.159	0.045	4.11E-04	0.251	0.160	1.18E-01	0.343	0.189	6.87E-02	0.016	0.016	3.35E-01
10	rs10182181	0.134	0.044	2.43E-03	-0.172	0.157	2.73E-01	-0.195	0.185	2.92E-01	-0.017	0.016	2.73E-01
11	rs3888190	0.167	0.045	1.99E-04	0.198	0.160	2.15E-01	0.226	0.188	2.30E-01	0.005	0.016	7.58E-01
12	rs1516725	0.257	0.064	5.57E-05	0.275	0.227	2.25E-01	0.382	0.267	1.52E-01	0.029	0.023	2.12E-01
13	rs12446632	0.129	0.063	3.95E-02	0.015	0.224	9.48E-01	-0.092	0.264	7.27E-01	-0.010	0.022	6.57E-01
14	rs2287019	0.140	0.056	1.20E-02	0.148	0.199	4.55E-01	0.096	0.234	6.82E-01	-0.012	0.020	5.32E-01
15	rs16951275	0.237	0.053	7.58E-06	0.218	0.189	2.49E-01	0.405	0.222	6.82E-02	0.046	0.020	1.73E-02
16	rs3817334	0.142	0.045	1.48E-03	-0.306	0.159	5.40E-02	-0.289	0.187	1.22E-01	-0.016	0.016	3.25E-01
17	rs2112347	0.128	0.046	5.14E-03	0.150	0.163	3.58E-01	0.203	0.192	2.90E-01	0.013	0.016	4.44E-01
18	rs12566985	0.049	0.044	2.69E-01	0.179	0.158	2.56E-01	0.284	0.186	1.25E-01	0.015	0.016	3.47E-01
19	rs3810291	0.160	0.047	7.69E-04	0.055	0.169	7.44E-01	0.112	0.199	5.74E-01	0.013	0.017	4.45E-01
20	rs7141420	0.121	0.044	5.60E-03	-0.144	0.156	3.58E-01	-0.246	0.184	1.81E-01	-0.025	0.016	1.16E-01
21	rs13078960	0.035	0.054	5.20E-01	0.095	0.194	6.25E-01	0.169	0.228	4.59E-01	0.026	0.019	1.77E-01

22	rs10968576	0.074	0.047	1.17E-01	-0.041	0.167	8.05E-01	-0.016	0.197	9.33E-01	0.015	0.017	3.82E-01
23	rs17024393	0.332	0.136	1.45E-02	0.826	0.485	8.83E-02	0.774	0.570	1.75E-01	0.008	0.049	8.77E-01
24	rs657452	0.057	0.045	2.09E-01	-0.031	0.161	8.46E-01	-0.075	0.189	6.92E-01	-0.001	0.016	9.29E-01
25	rs12429545	0.111	0.068	1.01E-01	-0.208	0.242	3.90E-01	-0.221	0.284	4.37E-01	0.007	0.024	7.66E-01
26	rs12286929	0.031	0.044	4.77E-01	-0.085	0.157	5.89E-01	-0.164	0.185	3.76E-01	-0.021	0.016	1.93E-01
27	rs13107325	0.265	0.085	1.88E-03	-0.766	0.304	1.17E-02	-0.942	0.357	8.39E-03	-0.045	0.031	1.50E-01
28	rs7903146	0.068	0.048	1.58E-01	-0.092	0.172	5.91E-01	-0.128	0.202	5.25E-01	-0.010	0.017	5.66E-01
29	rs10132280	0.103	0.048	3.03E-02	0.283	0.170	9.62E-02	0.284	0.200	1.56E-01	0.012	0.017	4.89E-01
30	rs17405819	0.078	0.047	9.96E-02	0.318	0.169	6.01E-02	0.434	0.199	2.93E-02	0.026	0.017	1.28E-01
31	rs6091540	0.027	0.048	5.77E-01	0.042	0.173	8.07E-01	0.167	0.203	4.10E-01	0.016	0.018	3.58E-01
32	rs1016287	0.166	0.048	6.20E-04	0.632	0.172	2.51E-04	0.758	0.203	1.87E-04	0.033	0.017	5.79E-02
33	rs4256980	0.031	0.046	4.95E-01	0.139	0.163	3.94E-01	0.274	0.192	1.53E-01	0.038	0.017	2.06E-02
34	rs17094222	0.096	0.054	7.66E-02	0.126	0.194	5.15E-01	0.098	0.228	6.68E-01	-0.014	0.020	4.91E-01
35	rs12401738	0.071	0.046	1.21E-01	0.214	0.164	1.90E-01	0.188	0.193	3.29E-01	-0.007	0.017	6.93E-01
36	rs7599312	0.035	0.050	4.75E-01	0.136	0.177	4.41E-01	0.075	0.208	7.18E-01	0.003	0.018	8.80E-01
37	rs2365389	0.107	0.045	1.64E-02	-0.143	0.159	3.69E-01	-0.174	0.187	3.52E-01	-0.006	0.016	7.28E-01
38	rs205262	0.040	0.049	4.14E-01	0.141	0.175	4.21E-01	0.296	0.205	1.49E-01	0.033	0.017	6.09E-02
39	rs2820292	0.083	0.044	6.07E-02	-0.076	0.158	6.29E-01	0.020	0.185	9.12E-01	0.013	0.016	4.16E-01
40	rs12885454	0.099	0.046	3.13E-02	0.195	0.163	2.31E-01	0.136	0.192	4.80E-01	-0.007	0.016	6.51E-01
41	rs12016871	0.105	0.057	6.37E-02	-0.050	0.203	8.06E-01	-0.015	0.238	9.49E-01	-0.006	0.021	7.81E-01
42	rs16851483	0.200	0.090	2.65E-02	-0.247	0.322	4.43E-01	0.048	0.379	8.98E-01	0.034	0.032	2.90E-01
43	rs1167827	0.066	0.044	1.37E-01	0.162	0.158	3.03E-01	0.072	0.185	6.97E-01	-0.018	0.016	2.52E-01
44	rs758747	0.063	0.048	1.87E-01	-0.027	0.171	8.74E-01	0.004	0.201	9.84E-01	-0.013	0.017	4.66E-01
45	rs1928295	0.051	0.044	2.46E-01	0.031	0.157	8.44E-01	0.087	0.185	6.39E-01	0.011	0.016	4.82E-01
46	rs9925964	0.028	0.045	5.38E-01	-0.016	0.161	9.22E-01	-0.093	0.190	6.25E-01	-0.017	0.016	2.97E-01
47	rs11126666	0.066	0.050	1.83E-01	0.459	0.177	9.63E-03	0.689	0.208	9.37E-04	0.058	0.018	1.02E-03
48	rs2650492	0.177	0.048	2.51E-04	0.144	0.172	4.03E-01	0.303	0.202	1.35E-01	0.026	0.017	1.39E-01
49	rs6804842	0.044	0.044	3.23E-01	-0.345	0.158	2.91E-02	-0.354	0.186	5.68E-02	-0.006	0.016	6.91E-01

50	rs12940622	0.091	0.044	3.96E-02	0.060	0.157	7.02E-01	0.113	0.185	5.42E-01	0.019	0.016	2.29E-01
51	rs7164727	0.018	0.047	7.09E-01	-0.135	0.168	4.22E-01	-0.123	0.197	5.34E-01	0.000	0.017	9.98E-01
52	rs4740619	0.019	0.044	6.70E-01	0.310	0.157	4.78E-02	0.350	0.184	5.75E-02	0.018	0.016	2.52E-01
53	rs492400	0.128	0.044	3.99E-03	0.229	0.158	1.47E-01	0.297	0.186	1.11E-01	0.010	0.016	5.26E-01
54	rs13191362	0.066	0.066	3.16E-01	0.039	0.235	8.70E-01	0.099	0.277	7.21E-01	0.028	0.024	2.38E-01
55	rs17001654	0.065	0.061	2.87E-01	-0.004	0.216	9.84E-01	-0.024	0.254	9.26E-01	0.011	0.022	5.99E-01
56	rs11191560	0.107	0.080	1.79E-01	-0.572	0.285	4.44E-02	-0.831	0.335	1.30E-02	-0.063	0.030	3.43E-02
57	rs2080454	0.055	0.045	2.26E-01	0.139	0.161	3.90E-01	0.100	0.189	5.96E-01	0.005	0.016	7.73E-01
58	rs2176040	0.003	0.046	9.39E-01	0.388	0.162	1.67E-02	0.355	0.191	6.30E-02	-0.008	0.016	6.07E-01
59	rs1000940	0.024	0.048	6.14E-01	0.043	0.171	8.03E-01	0.029	0.201	8.87E-01	0.012	0.017	4.90E-01
60	rs2033529	0.061	0.049	2.13E-01	-0.067	0.173	7.00E-01	-0.096	0.204	6.39E-01	0.004	0.018	8.31E-01
61	rs11583200	0.056	0.045	2.17E-01	0.205	0.162	2.05E-01	0.282	0.190	1.38E-01	0.021	0.016	2.00E-01
62	rs7239883	0.095	0.045	3.37E-02	0.309	0.160	5.34E-02	0.321	0.188	8.79E-02	0.012	0.016	4.41E-01
63	rs2836754	0.017	0.045	7.00E-01	0.142	0.162	3.81E-01	0.322	0.190	8.97E-02	0.026	0.016	1.07E-01
64	rs9400239	0.119	0.048	1.23E-02	0.067	0.170	6.94E-01	0.138	0.200	4.90E-01	0.007	0.017	6.71E-01
65	rs10733682	0.021	0.044	6.40E-01	-0.331	0.156	3.43E-02	-0.298	0.184	1.05E-01	0.005	0.016	7.31E-01
66	rs11688816	0.017	0.044	6.95E-01	-0.140	0.157	3.74E-01	-0.219	0.184	2.36E-01	-0.013	0.016	3.97E-01
67	rs11057405	0.042	0.072	5.57E-01	-0.352	0.255	1.68E-01	-0.366	0.300	2.23E-01	-0.014	0.025	5.94E-01
68	rs9914578	0.104	0.055	5.78E-02	0.504	0.195	9.83E-03	0.658	0.230	4.15E-03	0.040	0.019	3.84E-02
69	rs977747	0.060	0.045	1.81E-01	-0.063	0.159	6.93E-01	-0.019	0.187	9.21E-01	0.015	0.016	3.53E-01
70	rs2121279	0.176	0.066	7.90E-03	0.000	0.237	9.99E-01	0.029	0.278	9.17E-01	0.010	0.024	6.73E-01
71	rs29941	0.145	0.047	1.96E-03	0.022	0.167	8.97E-01	0.126	0.197	5.22E-01	0.033	0.017	5.33E-02
72	rs11727676	0.142	0.075	5.81E-02	-0.431	0.267	1.07E-01	-0.404	0.314	1.98E-01	-0.003	0.027	9.20E-01
73	rs3849570	0.039	0.046	4.00E-01	0.025	0.164	8.79E-01	-0.013	0.193	9.48E-01	-0.004	0.017	8.10E-01
74	rs9374842	0.050	0.051	3.35E-01	-0.357	0.183	5.14E-02	-0.334	0.215	1.21E-01	0.002	0.019	9.20E-01
75	rs4787491	0.080	0.044	6.63E-02	0.105	0.156	4.99E-01	0.151	0.183	4.08E-01	0.009	0.016	5.62E-01
76	rs1441264	0.115	0.045	1.00E-02	0.172	0.159	2.79E-01	0.253	0.187	1.77E-01	0.023	0.016	1.57E-01
77	rs7899106	0.317	0.105	2.45E-03	0.397	0.373	2.87E-01	0.725	0.439	9.84E-02	0.044	0.037	2.36E-01

78	rs2176598	0.116	0.051	2.41E-02	0.076	0.183	6.79E-01	0.227	0.215	2.90E-01	0.025	0.018	1.72E-01
79	rs2245368	0.083	0.058	1.54E-01	-0.103	0.207	6.19E-01	-0.124	0.244	6.12E-01	-0.013	0.021	5.33E-01
80	rs17203016	0.046	0.056	4.09E-01	-0.010	0.199	9.60E-01	0.067	0.234	7.76E-01	0.024	0.020	2.23E-01
81	rs7243357	0.120	0.058	3.92E-02	0.354	0.207	8.66E-02	0.552	0.243	2.30E-02	0.050	0.021	1.94E-02
82	rs16907751	0.161	0.071	2.38E-02	0.337	0.254	1.85E-01	0.367	0.299	2.19E-01	0.026	0.026	3.18E-01
83	rs1808579	0.049	0.044	2.65E-01	0.230	0.156	1.40E-01	0.335	0.183	6.71E-02	0.020	0.016	2.11E-01
84	rs13201877	0.089	0.064	1.66E-01	0.234	0.230	3.08E-01	0.472	0.270	8.00E-02	0.054	0.023	1.57E-02
85	rs2033732	0.007	0.050	8.95E-01	-0.008	0.180	9.65E-01	0.122	0.212	5.65E-01	0.033	0.018	7.25E-02
86	rs9540493	0.017	0.044	6.98E-01	0.020	0.157	8.98E-01	-0.045	0.184	8.06E-01	-0.011	0.016	4.84E-01
87	rs1460676	0.071	0.060	2.40E-01	0.518	0.215	1.61E-02	0.707	0.253	5.25E-03	0.037	0.021	8.06E-02

Abbreviations: SNP, single-nucleotide polymorphism; BMI, body mass index; SBP, systolic blood pressure; SE, standard error.

^a Consisted of simply using measured SBP as the outcome and ignoring the medication issue.

^b Consisted of adding a constant value of 15 mmHg to measured SBP in treated subjects.

^c Hypertension was defined as (i) SBP \geq 140 mmHg or diastolic blood pressure \geq 90 mmHg, and/or (ii) self-reported diagnosis of hypertension and current use of antihypertensive medication.

Notes: Estimates used linear regression for BMI and for SBP and robust Poisson regression for hypertension. All analyses included analytic weights.

Table S4

Associations of 87 BMI-associated SNPs included in the GRS with BMI and with measured SBP both adjusted for antihypertensive medication covariate, CLSA Comprehensive cohort, 2011-2018 (n = 20,430).

id	SNP	BMI, kg/m ²			Medication-adjusted SBP, mmHg		
		β	SE	P value	β	SE	P value
1	rs1558902	0.355	0.043	3.58E-16	0.753	0.155	1.20E-06
2	rs6567160	0.313	0.051	6.18E-10	0.042	0.180	8.15E-01
3	rs13021737	0.143	0.055	9.36E-03	0.280	0.196	1.54E-01
4	rs10938397	0.069	0.043	1.09E-01	-0.128	0.153	4.03E-01
5	rs543874	0.167	0.053	1.72E-03	-0.159	0.190	4.01E-01
6	rs2207139	0.132	0.057	1.98E-02	0.238	0.201	2.37E-01
7	rs11030104	0.224	0.052	1.87E-05	0.056	0.187	7.63E-01
8	rs3101336	0.088	0.044	4.31E-02	0.048	0.155	7.56E-01
9	rs7138803	0.143	0.044	1.01E-03	0.194	0.155	2.11E-01
10	rs10182181	0.137	0.043	1.28E-03	-0.158	0.152	2.97E-01
11	rs3888190	0.162	0.043	1.87E-04	0.182	0.155	2.41E-01
12	rs1516725	0.238	0.062	1.09E-04	0.210	0.220	3.39E-01
13	rs12446632	0.148	0.061	1.54E-02	0.080	0.217	7.13E-01
14	rs2287019	0.149	0.054	5.76E-03	0.181	0.192	3.47E-01
15	rs16951275	0.206	0.051	6.17E-05	0.104	0.183	5.71E-01
16	rs3817334	0.139	0.043	1.30E-03	-0.317	0.154	3.94E-02
17	rs2112347	0.119	0.044	7.21E-03	0.117	0.158	4.57E-01
18	rs12566985	0.031	0.043	4.67E-01	0.115	0.153	4.51E-01
19	rs3810291	0.150	0.046	1.09E-03	0.021	0.164	8.99E-01
20	rs7141420	0.139	0.042	1.08E-03	-0.081	0.151	5.91E-01
21	rs13078960	0.023	0.053	6.69E-01	0.050	0.188	7.91E-01
22	rs10968576	0.069	0.045	1.27E-01	-0.057	0.162	7.26E-01
23	rs17024393	0.341	0.132	9.52E-03	0.858	0.469	6.72E-02
24	rs657452	0.064	0.044	1.42E-01	-0.004	0.155	9.77E-01
25	rs12429545	0.114	0.066	8.39E-02	-0.200	0.234	3.93E-01
26	rs12286929	0.045	0.043	2.95E-01	-0.037	0.152	8.10E-01
27	rs13107325	0.295	0.083	3.53E-04	-0.658	0.294	2.51E-02
28	rs7903146	0.074	0.047	1.12E-01	-0.070	0.166	6.73E-01
29	rs10132280	0.103	0.046	2.54E-02	0.283	0.165	8.60E-02
30	rs17405819	0.059	0.046	2.02E-01	0.248	0.164	1.31E-01

31	rs6091540	0.006	0.047	9.01E-01	-0.034	0.167	8.37E-01
32	rs1016287	0.144	0.047	2.07E-03	0.555	0.167	8.91E-04
33	rs4256980	0.008	0.044	8.51E-01	0.056	0.158	7.21E-01
34	rs17094222	0.101	0.053	5.48E-02	0.143	0.188	4.44E-01
35	rs12401738	0.076	0.044	8.85E-02	0.231	0.158	1.46E-01
36	rs7599312	0.046	0.048	3.40E-01	0.174	0.171	3.10E-01
37	rs2365389	0.113	0.043	9.28E-03	-0.124	0.154	4.21E-01
38	rs205262	0.014	0.047	7.73E-01	0.045	0.169	7.89E-01
39	rs2820292	0.066	0.043	1.20E-01	-0.135	0.152	3.74E-01
40	rs12885454	0.109	0.044	1.42E-02	0.232	0.158	1.42E-01
41	rs12016871	0.100	0.055	7.04E-02	-0.071	0.196	7.18E-01
42	rs16851483	0.150	0.087	8.57E-02	-0.429	0.312	1.69E-01
43	rs1167827	0.081	0.043	5.83E-02	0.218	0.153	1.53E-01
44	rs758747	0.058	0.046	2.12E-01	-0.046	0.166	7.79E-01
45	rs1928295	0.042	0.043	3.29E-01	-0.003	0.152	9.83E-01
46	rs9925964	0.041	0.044	3.50E-01	0.031	0.156	8.41E-01
47	rs11126666	0.027	0.048	5.74E-01	0.318	0.171	6.38E-02
48	rs2650492	0.150	0.047	1.34E-03	0.047	0.167	7.79E-01
49	rs6804842	0.045	0.043	2.90E-01	-0.339	0.153	2.64E-02
50	rs12940622	0.082	0.043	5.53E-02	0.028	0.152	8.55E-01
51	rs7164727	0.016	0.046	7.33E-01	-0.142	0.162	3.81E-01
52	rs4740619	0.012	0.043	7.78E-01	0.286	0.152	5.94E-02
53	rs492400	0.116	0.043	6.77E-03	0.188	0.153	2.19E-01
54	rs13191362	0.056	0.064	3.82E-01	0.002	0.228	9.94E-01
55	rs17001654	0.068	0.059	2.48E-01	0.008	0.209	9.71E-01
56	rs11191560	0.151	0.077	5.06E-02	-0.414	0.275	1.33E-01
57	rs2080454	0.061	0.044	1.62E-01	0.162	0.156	2.99E-01
58	rs2176040	0.009	0.044	8.35E-01	0.409	0.157	9.18E-03
59	rs1000940	0.027	0.046	5.67E-01	0.051	0.165	7.57E-01
60	rs2033529	0.065	0.047	1.65E-01	-0.049	0.168	7.70E-01
61	rs11583200	0.043	0.044	3.29E-01	0.157	0.156	3.15E-01
62	rs7239883	0.093	0.043	3.18E-02	0.302	0.155	5.12E-02
63	rs2836754	-0.013	0.044	7.63E-01	0.031	0.156	8.43E-01
64	rs9400239	0.107	0.046	2.01E-02	0.023	0.164	8.88E-01
65	rs10733682	0.015	0.042	7.24E-01	-0.351	0.151	2.02E-02
66	rs11688816	0.031	0.043	4.72E-01	-0.091	0.152	5.49E-01
67	rs11057405	0.044	0.069	5.21E-01	-0.344	0.247	1.64E-01
68	rs9914578	0.078	0.053	1.42E-01	0.410	0.189	3.00E-02

69	rs977747	0.052	0.043	2.27E-01	-0.090	0.154	5.58E-01
70	rs2121279	0.171	0.064	7.67E-03	-0.018	0.229	9.37E-01
71	rs29941	0.128	0.045	4.96E-03	-0.042	0.162	7.95E-01
72	rs11727676	0.138	0.073	5.81E-02	-0.447	0.259	8.35E-02
73	rs3849570	0.045	0.045	3.11E-01	0.048	0.159	7.62E-01
74	rs9374842	0.046	0.050	3.59E-01	-0.371	0.177	3.64E-02
75	rs4787491	0.072	0.042	8.70E-02	0.077	0.150	6.10E-01
76	rs1441264	0.101	0.043	1.91E-02	0.123	0.154	4.24E-01
77	rs7899106	0.262	0.101	9.86E-03	0.196	0.361	5.87E-01
78	rs2176598	0.090	0.050	7.01E-02	-0.017	0.177	9.21E-01
79	rs2245368	0.086	0.056	1.24E-01	-0.090	0.200	6.53E-01
80	rs17203016	0.033	0.054	5.41E-01	-0.057	0.193	7.67E-01
81	rs7243357	0.086	0.056	1.26E-01	0.233	0.200	2.44E-01
82	rs16907751	0.156	0.069	2.38E-02	0.318	0.246	1.95E-01
83	rs1808579	0.031	0.042	4.66E-01	0.166	0.151	2.71E-01
84	rs13201877	0.049	0.062	4.33E-01	0.088	0.222	6.91E-01
85	rs2033732	-0.015	0.049	7.54E-01	-0.087	0.174	6.16E-01
86	rs9540493	0.028	0.043	5.09E-01	0.060	0.151	6.91E-01
87	rs1460676	0.039	0.058	5.04E-01	0.403	0.208	5.31E-02

Abbreviations: SNP, single-nucleotide polymorphism; BMI, body mass index; SBP, systolic blood pressure; SE, standard error.

Estimates are from linear regression models of BMI and of measured SBP on each SNP including antihypertensive medication as an adjustment covariate. All analyses used analytic weights.

Table S5

Associations of 87 BMI-associated SNPs included in the GRS with BMI and with measured SBP after exclusion of participants using antihypertensive medication, CLSA Comprehensive cohort, 2011-2018 (n = 14,337).

id	SNP	BMI, kg/m ²			Treated subjects excluded SBP, mmHg		
		β	SE	P value	β	SE	P value
1	rs1558902	0.346	0.052	2.03E-11	0.963	0.183	1.50E-07
2	rs6567160	0.273	0.060	5.30E-06	0.022	0.213	9.17E-01
3	rs13021737	0.181	0.065	5.35E-03	0.597	0.231	9.72E-03
4	rs10938397	0.015	0.051	7.72E-01	-0.159	0.180	3.76E-01
5	rs543874	0.163	0.063	9.50E-03	-0.243	0.224	2.77E-01
6	rs2207139	0.119	0.067	7.49E-02	0.302	0.238	2.04E-01
7	rs11030104	0.276	0.062	9.29E-06	-0.128	0.221	5.62E-01
8	rs3101336	0.093	0.051	7.13E-02	-0.106	0.182	5.60E-01
9	rs7138803	0.133	0.051	9.59E-03	0.077	0.183	6.75E-01
10	rs10182181	0.173	0.050	5.94E-04	-0.217	0.179	2.24E-01
11	rs3888190	0.151	0.051	3.14E-03	0.209	0.182	2.50E-01
12	rs1516725	0.230	0.072	1.51E-03	0.279	0.257	2.79E-01
13	rs12446632	0.168	0.072	1.95E-02	0.146	0.256	5.70E-01
14	rs2287019	0.159	0.064	1.27E-02	0.339	0.227	1.36E-01
15	rs16951275	0.231	0.061	1.41E-04	0.252	0.215	2.41E-01
16	rs3817334	0.138	0.051	7.27E-03	-0.363	0.182	4.61E-02
17	rs2112347	0.101	0.052	5.45E-02	0.129	0.186	4.88E-01
18	rs12566985	0.064	0.051	2.07E-01	0.038	0.180	8.33E-01
19	rs3810291	0.172	0.054	1.53E-03	0.104	0.193	5.90E-01
20	rs7141420	0.174	0.050	5.22E-04	0.013	0.178	9.42E-01
21	rs13078960	0.062	0.062	3.18E-01	0.240	0.222	2.79E-01
22	rs10968576	0.128	0.054	1.77E-02	-0.149	0.191	4.34E-01
23	rs17024393	0.239	0.155	1.24E-01	0.861	0.552	1.18E-01
24	rs657452	0.059	0.052	2.51E-01	0.154	0.183	4.00E-01
25	rs12429545	0.145	0.078	6.28E-02	-0.298	0.276	2.80E-01
26	rs12286929	0.033	0.050	5.10E-01	-0.093	0.179	6.03E-01
27	rs13107325	0.275	0.097	4.57E-03	-0.474	0.344	1.69E-01
28	rs7903146	0.094	0.055	8.96E-02	-0.080	0.196	6.83E-01
29	rs10132280	0.101	0.055	6.43E-02	0.370	0.194	5.70E-02
30	rs17405819	0.042	0.054	4.33E-01	0.344	0.192	7.34E-02

31	rs6091540	0.022	0.055	6.92E-01	0.129	0.196	5.11E-01
32	rs1016287	0.168	0.056	2.52E-03	0.456	0.197	2.08E-02
33	rs4256980	0.032	0.052	5.44E-01	0.100	0.186	5.89E-01
34	rs17094222	0.121	0.062	5.12E-02	0.158	0.220	4.75E-01
35	rs12401738	0.065	0.053	2.20E-01	0.251	0.187	1.79E-01
36	rs7599312	0.073	0.057	2.02E-01	0.229	0.202	2.57E-01
37	rs2365389	0.102	0.051	4.76E-02	-0.177	0.182	3.32E-01
38	rs205262	0.039	0.056	4.87E-01	-0.026	0.200	8.97E-01
39	rs2820292	0.074	0.051	1.47E-01	-0.079	0.180	6.61E-01
40	rs12885454	0.095	0.052	6.98E-02	0.127	0.185	4.91E-01
41	rs12016871	0.066	0.065	3.15E-01	-0.065	0.231	7.78E-01
42	rs16851483	0.089	0.105	3.94E-01	-0.447	0.372	2.30E-01
43	rs1167827	0.064	0.051	2.04E-01	0.132	0.179	4.62E-01
44	rs758747	0.066	0.055	2.30E-01	0.074	0.195	7.03E-01
45	rs1928295	0.009	0.051	8.53E-01	0.045	0.180	8.00E-01
46	rs9925964	0.047	0.052	3.65E-01	-0.130	0.184	4.81E-01
47	rs11126666	0.059	0.057	2.99E-01	0.469	0.203	2.09E-02
48	rs2650492	0.139	0.056	1.22E-02	0.109	0.197	5.79E-01
49	rs6804842	0.021	0.051	6.77E-01	-0.240	0.180	1.83E-01
50	rs12940622	0.072	0.050	1.51E-01	0.205	0.179	2.50E-01
51	rs7164727	0.073	0.054	1.78E-01	-0.084	0.192	6.64E-01
52	rs4740619	0.010	0.050	8.44E-01	0.346	0.178	5.23E-02
53	rs492400	0.080	0.051	1.14E-01	0.104	0.180	5.63E-01
54	rs13191362	0.065	0.076	3.87E-01	-0.197	0.269	4.64E-01
55	rs17001654	0.056	0.069	4.20E-01	0.193	0.245	4.32E-01
56	rs11191560	0.164	0.090	6.97E-02	-0.293	0.321	3.62E-01
57	rs2080454	0.044	0.052	3.93E-01	0.153	0.184	4.06E-01
58	rs2176040	-0.032	0.052	5.39E-01	0.410	0.185	2.68E-02
59	rs1000940	0.067	0.055	2.22E-01	0.177	0.195	3.64E-01
60	rs2033529	0.028	0.056	6.14E-01	-0.005	0.197	9.78E-01
61	rs11583200	0.052	0.052	3.21E-01	0.263	0.185	1.56E-01
62	rs7239883	0.064	0.051	2.09E-01	0.368	0.182	4.27E-02
63	rs2836754	-0.060	0.052	2.46E-01	0.045	0.184	8.06E-01
64	rs9400239	0.050	0.054	3.57E-01	0.068	0.193	7.25E-01
65	rs10733682	0.012	0.050	8.17E-01	-0.254	0.179	1.54E-01
66	rs11688816	0.022	0.050	6.63E-01	-0.152	0.178	3.95E-01
67	rs11057405	0.061	0.082	4.61E-01	-0.271	0.291	3.52E-01
68	rs9914578	0.132	0.063	3.69E-02	0.458	0.224	4.09E-02

69	rs977747	0.043	0.051	3.97E-01	0.002	0.182	9.89E-01
70	rs2121279	0.138	0.076	6.79E-02	-0.007	0.269	9.78E-01
71	rs29941	0.140	0.054	9.06E-03	-0.068	0.190	7.20E-01
72	rs11727676	0.076	0.086	3.75E-01	-0.326	0.305	2.85E-01
73	rs3849570	0.065	0.053	2.15E-01	0.114	0.187	5.40E-01
74	rs9374842	0.049	0.059	4.04E-01	-0.311	0.209	1.37E-01
75	rs4787491	0.048	0.050	3.39E-01	-0.059	0.177	7.40E-01
76	rs1441264	0.100	0.051	5.00E-02	0.170	0.181	3.48E-01
77	rs7899106	0.187	0.121	1.22E-01	-0.151	0.429	7.25E-01
78	rs2176598	0.046	0.059	4.37E-01	0.064	0.208	7.58E-01
79	rs2245368	0.092	0.066	1.67E-01	-0.159	0.236	5.01E-01
80	rs17203016	0.011	0.064	8.62E-01	0.025	0.228	9.14E-01
81	rs7243357	0.100	0.066	1.30E-01	0.233	0.234	3.19E-01
82	rs16907751	0.198	0.081	1.52E-02	0.110	0.289	7.05E-01
83	rs1808579	0.068	0.050	1.75E-01	0.139	0.177	4.33E-01
84	rs13201877	0.019	0.074	7.94E-01	-0.093	0.263	7.24E-01
85	rs2033732	0.001	0.058	9.86E-01	-0.087	0.205	6.71E-01
86	rs9540493	0.031	0.050	5.42E-01	-0.047	0.178	7.92E-01
87	rs1460676	0.027	0.070	6.99E-01	0.407	0.247	9.94E-02

Abbreviations: SNP, single-nucleotide polymorphism; BMI, body mass index; SBP, systolic blood pressure; SE, standard error.

Estimates are from linear regression models of BMI and of measured SBP on each SNP after exclusion of treated subjects. All analyses used analytic weights.

Reference

1. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; 518:197-206.

Chapitre 5 – Discussion et conclusion

En RM, les violations de l'hypothèse de restriction d'exclusion représentent la principale menace à la validité des instruments.¹³ Spécifiquement, la pléiotropie, phénomène biologique ubiquitaire dans le génome humain^{24,25} et par lequel un SNP affecte à la fois l'exposition et l'issue de manière distincte,^{13,14} est une cause principale de violation de la restriction d'exclusion.¹⁴ L'utilisation des SNPs pléiotropiques comme instruments dans une étude de RM peut induire un biais dans l'estimation de l'effet causal, une perte de puissance et une inflation de l'erreur de type I.^{26,29,167} Ainsi, la satisfaction de l'hypothèse de restriction d'exclusion est déterminante pour la crédibilité de la RM de gérer la confusion induite par des facteurs non mesurés dans les études observationnelles, qui est l'une des principales plus-values de cette approche vis-à-vis des modélisations par les régressions conventionnelles.

A travers ses trois études, la présente thèse a examiné certains défis méthodologiques pratiques relatifs à la vérification de l'hypothèse de restriction d'exclusion et à la validité des résultats dans une étude de RM. Plus spécifiquement, la thèse a (i) cartographié, via une revue de littérature, comment les auteurs des études de RM préviennent, évaluent ou contrôlent, et discutent des violations potentielles de la restriction d'exclusion dues notamment à la pléiotropie, (ii) évalué, au moyen de simulations, la performance de la méthode basée sur la confusion positive, proposée par Glymour et al.,⁵ à détecter des instruments invalides en RM, (iii) examiné, à partir d'une étude de RM estimant l'effet causal de l'IMC sur la PAS, l'impact de quelques méthodes courantes de gestion de la médication antihypertensive sur l'estimation de l'effet causal et la détection des violations potentielles de la restriction d'exclusion lorsqu'on modélise la PA en RM.

Trois principaux résultats peuvent être mis de l'avant dans cette thèse. Premièrement, comme l'étude 1 l'indique, plusieurs chercheurs en RM utilisent comme instruments de multiples SNPs indépendants (combinés ou non en un GRS) comme une stratégie pour prévenir ou minimiser le biais dû à des SNPs pléiotropiques. Aussi, les méthodes statistiques utilisées comme analyses complémentaires pour évaluer ou gérer des possibles violations de la restriction d'exclusion sont très hétérogènes, et ont toutes des limites ; d'où l'intérêt de trianguler plusieurs méthodes pour une inférence plus crédible. Deuxièmement, comme l'étude 2 le suggère, advenant une taille d'échantillon suffisante, la méthode basée sur la confusion positive est performante à détecter la

pléiotropie dans une étude de RM lorsque l'instrument est un SNP, mais peut s'avérer peu performante à détecter la pléiotropie d'un GRS. Troisièmement, comme l'étude 3 le montre, des décisions de modélisation peuvent avoir des impacts sur les résultats de la RM. Ainsi, si l'effet causal estimé de l'IMC sur la PAS en RM varie beaucoup en fonction de méthodes de gestion de la médication, les évaluations de la validité des instruments restent robustes.

Le présent chapitre discute les principaux résultats de notre thèse autour de cinq thématiques, tout en relevant leur pertinence et leur contribution aux connaissances actuelles. Par la suite, nous présentons l'impact de notre thèse en santé publique puis les forces et les limites de chacune des trois études de la thèse, avant de clôturer par une conclusion.

5.1. Discussion des principaux résultats de la thèse

5.1.1. Prévention du biais dû aux SNPs pléiotropiques dans les études de RM

La sélection des SNPs occupe une place importante dans l'effort pour la réduction du biais dû à la pléiotropie des instruments.³⁵ La façon par excellence d'éviter la pléiotropie dans une étude de RM est de ne pas inclure comme instruments des SNPs qui sont pléiotropiques pour l'issue d'intérêt, et cela requiert de la connaissance sur les mécanismes biologiques spécifiques des gènes.^{88,174,175} Notre thèse (étude 1) montre qu'en pratique, les auteurs ont sélectionné davantage plusieurs SNPs comme instruments et ce, exclusivement sur base de l'existence d'une association statistique entre les SNPs et l'exposition (IMC), ce qui accroît sensiblement les chances d'inclure des SNPs pléiotropiques dans une étude de RM.^{28,100} Ce résultat peut en partie s'expliquer par la nature polygénique de l'IMC qui ne dispose pas d'un gène codant proximal spécifique, et qu'il aurait été peut être différent si l'exposition considérée était une protéine pour laquelle les SNPs pouvaient être sélectionnés dans la région du gène codant (on parle de '*cis-variants*').³⁵ En même temps, ce résultat peut être transposable à divers autres facteurs de risque étudiés en santé publique qui, à l'exemple de l'IMC, sont des phénotypes complexes pour lesquels la connaissance sur la fonction spécifiques des gènes ne guidera probablement que très peu la sélection des instruments, car non disponible.^{98,99}

Aussi, notre revue a mis en exergue le choix de plusieurs auteurs d'utiliser plusieurs SNPs indépendants, combinés ou non en GRS, comme une stratégie pour atténuer l'impact de la

pléiotropie en RM, présumant une annulation mutuelle des effets pléiotropiques des différents SNPs lorsqu'ils sont utilisés ensemble.²³⁶ Cette notion d'utiliser plusieurs SNPs comme une stratégie de prévention contre la pléiotropie est controversée car l'annulation mutuelle des effets pléiotropiques n'est absolument pas garantie et n'est pas fondée sur des recommandations de récentes lignes directrices de la RM.^{35,85} D'autres chercheurs ont plutôt montré que l'inclusion des SNPs pléiotropiques dans un GRS biaise l'estimation de l'effet causal en RM, et ce biais augmente avec l'augmentation de la proportion des SNPs pléiotropiques dans le GRS.⁹⁴

Ainsi donc, dépendamment de la nature de l'exposition d'intérêt, la sélection des SNPs peut être une étape de vulnérabilité importante des études de RM à la pléiotropie du fait des limites des connaissances sur les mécanismes biologiques des SNPs qu'on sélectionne. D'où l'intérêt, pour les chercheurs de suivre les recommandations des lignes directrices de la RM,^{35,85} mises à jour régulièrement, et d'user de plusieurs méthodes statistiques disponibles pour évaluer la robustesse des résultats aux violations potentielles des hypothèses de la RM.

5.1.2. Évaluation de la plausibilité de l'hypothèse de restriction d'exclusion dans les études de RM

Outre la problématique de sélection des SNPs potentiellement pléiotropiques ci-haut évoquée, un défi important dans l'effort de limiter l'impact de la pléiotropie dans les études de RM réside dans l'impossibilité de prouver, avec les données sous-analyse, la satisfaction des hypothèses 2 (indépendance) et 3 (restriction d'exclusion) de la RM devant démontrer que les instruments sélectionnés sont valides.^{5,11} Par conséquent, les différentes méthodes statistiques proposées pour évaluer la plausibilité des hypothèses de RM permettent simplement de falsifier ces hypothèses, c.a.d. tenter de détecter une violation potentielle des hypothèses de RM, et/ou de contrôler le biais.^{7,34,42,88} Notre thèse (étude 1) montre qu'en pratique, la très grande majorité des auteurs appliquent différentes méthodes statistiques dans les analyses complémentaires pour évaluer la plausibilité de la restriction d'exclusion, conformément à la recommandation des lignes directrices de la RM.^{35,85} Cependant, certains défis pratiques peuvent affecter le choix, l'efficacité et l'interprétation de ces méthodes.

Premièrement, notre étude a montré qu'il existe une pléthore des méthodes disponibles dont la mécanique sous-tendant la détection de la violation de la restriction d'exclusion peut différer

sensiblement. Certaines méthodes, à l'exemple des méthodes robustes (ex. MR-Egger, etc.),^{29,97} ont fait l'objet d'évaluations plus approfondies et ont été validées, alors que d'autres stratégies (ex. estimer simplement l'association entre l'instrument et l'issue) sont susceptibles d'être sous-optimales à détecter une violation de la restriction d'exclusion.

Deuxièmement, toutes les méthodes recensées ont des faiblesses. Par exemple, le très populaire test d'intercept de la régression MR-Egger, qui était la méthode robuste la plus utilisée par les auteurs dans notre revue,⁴² peut sévèrement manquer de puissance pour détecter la pléiotropie directionnelle, et ce, même si la proportion des SNPs pléiotropiques est relativement importante. En effet, à travers leur étude de simulation, Burgess et al.²³⁷ ont montré que lorsque 30% des SNPs simulés étaient pléiotropiques, la puissance du test de l'intercept de la régression MR-Egger à détecter la pléiotropie directionnelle était autour de 13 à 15% lorsque l'hypothèse InSIDE (indépendance entre les effets pléiotropiques directes et les associations SNPs–exposition) était satisfaite.

Troisièmement, l'interprétation de certaines méthodes peut être difficile dans la mesure où plusieurs d'entre elles ne sont pas spécifiques à la pléiotropie ou à la restriction d'exclusion. En effet, plusieurs parmi ces méthodes peuvent aussi capturer les violations de l'hypothèse 2 (indépendance), en particulier la stratification de la population, dont la gestion est bien différente de celle de la pléiotropie (restriction de l'analyse aux sujets homogènes sur le plan de l'ascendance, ajustement des modèles de RM pour les composantes principales, etc.).¹⁶⁸ D'autres méthodes, à l'exemple des méthodes basées sur l'analyse de l'hétérogénéité entre les effets causaux obtenus avec différents instruments, peuvent être positives si l'ampleur des effets causaux diffèrent sensiblement, que cette variation soit due à la pléiotropie ou pas.¹²

Ainsi donc, compte tenu des limites des méthodes, il appartient aux chercheurs de faire usage de précaution dans le choix des méthodes les plus optimales en fonction des données en possession. Le choix d'une méthode ne doit pas se faire au hasard ; il doit être réfléchi, en tenant compte de la performance de la méthode mais aussi de sa mécanique à détecter la pléiotropie. L'utilisation concomitante de plusieurs méthodes qui fonctionnent sous des hypothèses différentes est indispensable et permettra d'accroître la crédibilité des résultats.^{176,177}

5.1.3. Performance de la méthode basée sur la confusion positive à détecter la pléiotropie des instruments dans les études de RM

Lorsque la méthode basée sur la confusion positive, proposée par Glymour et al.⁵ en 2012, a été évaluée au cours de la même année par Burgess.⁴¹ Il s'est dégagé le constat selon lequel, dans une analyse de RM avec un SNP comme instrument et une issue continue, la méthode basée sur la confusion positive était efficace à détecter la pléiotropie directe mais moins efficace (65% de succès) en cas de pléiotropie indirecte, en plus d'allouer des résultats faux positifs (dans 11% des cas).⁴¹

Notre étude 2 a répliqué cette étude et confirmé ces résultats. Par la suite, nous avons étendu ces simulations en considérant des tailles d'échantillon plus grandes que les 2000 sujets générés par Burgess.⁴¹ Nos résultats montrent que la performance de la méthode à détecter la pléiotropie indirecte s'est nettement améliorée atteignant 94% pour $n=30000$ et les détections faussement positives d'invalidité des instruments disparaissent dès $n=10000$. Comme attendu, l'augmentation de la taille d'échantillon a eu comme corolaire d'augmenter la précision des estimations, beaucoup plus dans l'analyse de RM (qui est plus imprécise) que dans l'analyse conventionnelle.¹²⁵ Précisons que Glymour et al.⁵ ont conceptualisé leur méthode à partir des paramètres de la population, en ignorant la variation échantillonnale.⁵ En réalité, étant donné qu'en pratique les analyses sont faites sur des échantillons finis, l'imprécision liée à l'estimation de l'effet par la régression conventionnelle et surtout par la RM peut faire que la méthode permette des fausses conclusions sur la présence (ou non, inversement) d'instruments invalides. Toutefois, lorsque la taille d'échantillon est très grande, la méthode basée sur la confusion positive pourrait tout de même convenir telle qu'elle, comme observé dans nos simulations. Nos résultats considérant une issue binaire montrent que la performance de la méthode basée sur la confusion positive à détecter l'invalidité si l'instrument est un SNP est similaire à celle qu'on a observée pour une issue continue, c.à.d., la méthode performe moins bien lorsque la pléiotropie est indirecte que directe, mais la performance s'améliore (augmente) avec l'augmentation de la taille d'échantillon.

Par la suite, nous avons innové en évaluant, pour la toute première fois, la méthode basée sur la confusion positive dans un contexte d'analyse de RM où l'instrument est un GRS. En effet, la méthode basée sur la confusion positive avait été utilisée en RM avec un GRS comme instrument³⁸

même si sa performance n'avait pas été spécifiquement évaluée dans ce contexte. En pratique, même si un GRS s'utilise de la même manière qu'un simple SNP, il est une combinaison de plusieurs SNPs, ce qui permet d'augmenter la force de l'instrument et donc la puissance statistique.^{58,93} Nos résultats suggèrent que, indépendamment de la nature continue ou binaire de l'issue, la méthode basée sur la confusion positive semble peu performante à détecter l'invalidité d'un GRS comme instrument, mais la performance à détecter l'invalidité augmente avec l'augmentation de la proportion des SNPs pléiotropiques dans le GRS. Ce résultat souligne l'intérêt d'être prudent dans l'interprétation de cette méthode dans une analyse de RM utilisant un GRS comme instrument car, comme on l'observe, cela requiert une proportion importante des SNPs pléiotropiques et donc un biais plus important pour que la performance de la méthode à détecter l'invalidité d'un GRS augmente. Par conséquent, il est possible, en dehors de tout phénomène de balancement de pléiotropie (annulation mutuelle des effets pléiotropiques des différents SNPs) au sein d'un GRS,²³⁶ que la méthode basée sur la confusion positive ne détecte pas l'invalidité d'un GRS affectant peu de SNPs du fait que le biais ne sera pas important.

Nos résultats suggèrent globalement que la méthode basée sur la confusion positive peut être un outil intéressant à considérer dans l'évaluation de la restriction d'exclusion dans une analyse de RM où l'instrument est un SNP unique, surtout si la taille d'échantillon est suffisante. Vu que des études de RM utilisant un SNP comme instrument sont toujours publiées,²³⁸⁻²⁴³ et compte tenu du fait que l'offre des méthodes de falsification des hypothèses de RM dans le cas d'un instrument simple est assez limitée,^{12,42} la méthode basée sur la confusion positive peut toujours trouver sa place. Dans le cas de GRS, des études futures sont nécessaires pour clarifier davantage la performance la méthode en considérant diverses situations et paramètres plausibles.

5.1.4. Impact des méthodes de gestion de la médication antihypertensive sur l'estimation de l'effet causal dans les études de RM modélisant la PA

Une évaluation antérieure des méthodes de gestion de la médication antihypertensive dans des modèles de régression conventionnelle analysant les déterminants de la PAS a suggéré entre autres que l'ajout d'une valeur constante de 15 mm Hg à la PAS chez les sujets traités aux antihypertenseurs était une meilleure méthode de gestion de la médication par rapport à d'autres stratégies comme ignorer la médication, ajuster pour la médication dans les modèles ou exclure les

sujets traités aux antihypertenseurs de l'analyse, qui biaisaient l'association vers la valeur nulle.⁵³ Notre étude 3 a examiné ces méthodes en RM en utilisant les données de l'ÉLCV et a montré que l'effet de l'IMC sur la PAS estimé en RM variait de manière importante selon les méthodes de gestion de la médication appliquée. Cela suggère que la façon dont on tient compte de la médication a un impact différentiel, et peut biaiser l'estimation de l'effet causal des déterminants de la PAS en RM. De plus, le fait que la variation relative des coefficients estimés dans nos modèles de RM était similaire avec celle des coefficients de nos modèles conventionnels et aussi cohérente avec l'étude de simulation de Tobin et al.,⁵³ cela peut suggérer que les observations de l'étude de simulation par les régressions conventionnelles pourraient s'appliquer aussi aux études de RM. Par conséquent, ajouter 15 mm Hg à la PAS chez les sujets traités aux antihypertenseurs, approche d'ailleurs appliquée dans plusieurs études de RM,^{50,51,244-248} pourrait aussi être une meilleure stratégie de gestion de la médication dans les études des déterminants de la PAS en RM.

Cependant, on note aussi que plusieurs autres études de RM de la PA ont appliqué des stratégies comme ignorer la médication,^{43-47,249} ajuster pour la médication dans les modèles,^{48,49,250} ou exclure les sujets traités aux antihypertenseurs de l'analyse,²⁵¹ même si ces stratégies se sont avérées peu optimales dans l'étude de simulation de Tobin et al.,⁵³ et ce, sans se préoccuper de leur impact potentiel sur les résultats de la RM. De même, lorsqu'on entreprend d'étudier les déterminants de la PA par une analyse de RM à deux échantillons, il est important de vérifier les modèles ayant généré les statistiques sommaires qu'on va utiliser. L'étude de RM par Borges et al.,²⁵² par exemple, a utilisé les statistiques sommaires sur la PAS et PAD du 'consortium international pour la pression artérielle' (*International Consortium for Blood Pressure, ICBP*),²⁵³ obtenues à partir des modèles de régression ayant inclus la médication antihypertensive comme une covariable d'ajustement. Ceci pourrait, au vu des conclusions de Tobin et al.,⁵³ contribuer à sous-estimer les paramètres sommaires de la PA et, par conséquent, l'association avec le facteur de risque.

Ainsi donc, dans le contexte d'augmentation de la prévalence de l'hypertension et d'utilisation accrue de médicaments antihypertenseurs à travers le monde,⁵⁵⁻⁵⁷ le choix de la stratégie de gestion de la variable médication dans les études de RM de la PA est primordiale vu qu'une mauvaise stratégie peut potentiellement biaiser l'estimation de l'effet causal en RM.

5.1.5. Utilisation spécifique de *FTO* comme instrument de l'IMC dans nos études

Un autre aspect important à mettre en exergue dans notre recherche est notre choix délibéré d'utiliser *FTO* comme instrument de l'adiposité, aussi bien comme critère d'inclusion des études dans la revue de littérature (étude 1) et donc dans l'illustration de la méthode basée sur la confusion positive (étude 2), que pour la modélisation de la PAS en RM (étude 3). Ce choix a été motivé, non seulement du fait que l'adiposité est un facteur de risque majeur contribuant significativement à l'incidence mondiale de plusieurs problèmes de santé,²⁵⁴ mais surtout du fait que des SNPs sur *FTO*, souvent sélectionnés comme instruments de l'IMC, sont suspectés de pléiotropie sur divers phénotypes de santé, incluant la PAS.^{37,199-202,205-207,210} En faisant ce choix, nous augmentons les chances d'inclure, dans l'article 1, des études de RM ayant une discussion plus élaborée sur la pléiotropie et d'examiner, dans l'étude 3, l'influence potentielle des méthodes de gestion de la médication sur la détection des violations de la restriction d'exclusion dans un contexte d'analyse de RM où l'instrument serait possiblement invalide.

Notre étude 1 a documenté des suspicions de pléiotropie de certains SNPs sur *FTO* majoritairement sur des issues de santé mentale (ex. bien-être subjectif,²⁰⁸ symptômes d'anxiété phobique,³⁹ troubles mentaux courants,³⁷ dépression⁴⁰), mais les mécanismes biologiques sous-tendant ces effets pléiotropiques n'étaient pas clairement élucidés. Toutefois, dans des études animales, il a été suggéré que *FTO* pourrait influencer les comportements de type anxiété et dépression via des altérations de la flore intestinale chez les souris.²⁰⁹ Notre étude 3 a suspecté *FTO* d'affecter la PAS indépendamment de l'IMC par des voies impliquant l'utilisation de la médication. Une explication plausible serait que *FTO* affecterait des conditions de santé mentale, comme précédemment évoqué, qui à leur tour influenceraient l'adhérence et donc l'utilisation des médicaments antihypertenseurs.^{61,62} Ainsi donc, même si le choix de *FTO* comme instrument de l'IMC en RM est naturellement dictée par leur forte association statistique,⁵⁹ il est important pour les chercheurs de faire preuve de prudence lorsque l'analyse porte sur la PA ou des phénotypes de santé mentale.

5.2. Impact de la thèse en santé publique

L'impact de la présente thèse en santé publique est à deux niveaux. Premièrement, sur le plan méthodologique, notre thèse aborde des défis complexes touchant à la validité des résultats de la RM, une approche analytique de plus en plus utilisée comme une alternative aux modèles de régression conventionnelle afin de limiter le biais dû notamment aux facteurs de confusion non mesurés dans les études observationnelles.⁸⁰ Par conséquent, notre thèse s'inscrit dans le transfert des connaissances méthodologiques vers les chercheurs qui exploitent les techniques de la RM dans le but de les outiller à la vérification de la validité de leurs résultats. Ceci est primordial pour la qualité des évidences mises à la disposition de la population et des décideurs politiques, dans un contexte où l'utilisation de la RM, qui s'attelait à l'estimation de l'effet causal des facteurs de risque environnementaux sur des issues de santé, s'étend à d'autres secteurs, dont le développement des médicaments^{68,255} et les maladies infectieuses.²⁵⁶⁻²⁶⁰ La recherche épidémiologique de qualité, en effet, est fondamentale pour l'exercice des fonctions essentielles de la santé publique, notamment par la production des meilleures connaissances devant orienter la prise de décisions et la mise en œuvre des politiques, programmes et actions de prévention des maladies, de surveillance et de promotion de la santé des personnes, et de la protection de l'environnement.²⁶¹

Deuxièmement, la présente thèse contribue à l'avancement des connaissances sur le lien de causalité entre l'obésité et l'hypertension, deux conditions de santé affectant, respectivement, 26,5%²⁶² et 23,2%⁵⁵ des canadiens adultes. Notre thèse est l'une des premières études canadiennes à estimer l'effet causal de l'IMC sur la PAS et l'hypertension. Vu le rôle majeur de l'adiposité dans la survenue de l'hypertension, un des fardeaux sanitaires les plus répandus dans le monde et principale cause de morbidité cardiovasculaire et de mortalité toutes causes confondues,^{263,264} et compte tenu de l'ampleur de l'adiposité chez les jeunes au Canada (27,5% d'enfants de 6 à 17 ans étaient en surpoids ou obèses en 2019²⁶⁵), les résultats de notre thèse devraient s'imbriquer dans les efforts de prévention de l'adiposité, particulièrement dans le jeune âge, en vue de limiter l'incidence et/ou les impacts de l'hypertension dans notre société.²⁶⁶ Finalement, notre recherche est porteuse puisque, même si elle a considéré spécifiquement le gène *FTO* comme instrument de l'IMC pour illustrer des défis méthodologiques pratiques ayant une incidence sur la validité des résultats de la RM, sa portée va bien au-delà de *FTO* et de l'IMC et s'applique à toute utilisation de la RM, peu importe l'instrument et l'exposition considérés.

5.3. Forces et limites de notre recherche

Étude 1

Il existe, certes, plusieurs revues des études de RM¹⁵⁻²³ ayant résumé divers aspects pratiques de la RM au cours de la dernière décennie, mais aucune d'entre elles n'a spécifiquement examiné la manière dont les auteurs préviennent ou gèrent le biais potentiel lié à la pléiotropie. L'étude 1 de la présente thèse est originale en étant la première revue de littérature à cartographier la gestion pratique des chercheurs des possibles violations de la restriction d'exclusion dues notamment à la pléiotropie. L'article suscite un intérêt indéniable dans la communauté scientifique avec plus de 3500 vues et téléchargements en 17 mois, et en étant plus téléchargé que 81% de tous les articles de *Frontiers*.²⁶⁷

Toutefois, plusieurs considérations peuvent limiter la portée de nos résultats. Premièrement, notre revue n'est pas systématique. Elle s'est limitée à comptabiliser les méthodes rapportées dans les articles sans s'intéresser à la qualité des études dans la mesure où l'objectif de notre recherche était essentiellement de recenser l'éventail des pratiques des chercheurs sur la gestion de violations potentielles de la restriction d'exclusion. Par conséquent, c'est la qualité de la méthode de gestion de l'hypothèse de restriction d'exclusion appliquée par les auteurs (considération abordée plus-bas) plutôt que la qualité même de l'article qui était une préoccupation pour nous. Aussi, la recherche documentaire n'a pas été étendue à d'autres sources (ex. données non publiées) en dehors des quatre moteurs de recherche exploités, de sorte que, même si la plupart des méthodes de gestion de la pléiotropie ont été capturées, il est possible que certaines méthodes, notamment plus récentes, n'aient pas été rapportées.

Deuxièmement, les méthodes de gestion de la restriction d'exclusion rapportées dans notre revue incluent des méthodes bien connues et validées, et d'autres stratégies qui peuvent être moins efficaces/optimales. Même si nous avons résumé chaque méthode recensée avec ses limites (voir **Table S6** de l'article 1), nous appelons les chercheurs à faire preuve de prudence dans le choix des meilleures méthodes pour les données dont ils disposent. Troisièmement, il est possible que le fait de restreindre l'inclusion dans la revue aux seules études de RM ayant l'IMC comme exposition puisse limiter la généralisation de certains résultats dans d'autres contextes d'études de RM, par

exemple, où l'exposition d'intérêt est une protéine disposant d'un gène codant proximal spécifique connu. Dans une telle situation, les SNPs seraient davantage sélectionnés sur base de la connaissance des mécanismes biologiques spécifiques du gène codant cette protéine plutôt que par l'approche statistique comme observée avec l'IMC.³⁵ Mais en même temps, dans un tel cas, la question de pléiotropie se poserait moins vu la spécificité de l'action du gène codant sur l'exposition protéique, qu'il ne le serait si un trait polygénique comme l'IMC est l'exposition.

Quatrièmement, enfin, nous avons rencontré des difficultés opérationnelles liées au reportage des informations dans les articles qui était parfois peu claire et incomplet, avec des discussions sur la pléiotropie souvent sommaires. Certains auteurs contactés pour obtenir des informations supplémentaires n'ont pas répondu à nos courriels. Un reportage plus complet, respectueux des lignes directrices de la RM,⁸⁵ aiderait à mieux entreprendre les futures études.

Étude 2

L'étude 2 de la présente thèse a utilisé des simulations statistiques pour approfondir l'évaluation de la performance de la méthode basée sur la confusion positive dans la détection de la pléiotropie des instruments en intégrant des contextes pratiques des études de RM (par ex. utilisation de GRS comme instrument, issue binaire) non étudiés antérieurement (étude initiale de Glymour et al.,⁵ ou l'évaluation antérieure par Burgess⁴¹). En tant que telle, cette étude est originale et génératrice de connaissance. Elle met à la disposition des chercheurs en RM des données probantes sur la performance de cette méthode en vue d'un usage éclairé dans leurs études respectives. Notre étude s'est basée principalement sur des simulations existantes et publiées,^{41,94} offrant la possibilité de comparer certains résultats de notre étude avec des travaux existants, ce qui est indispensable pour la triangulation des connaissances.

En outre, grâce à nos résultats indiquant une amélioration de la performance de la méthode avec l'augmentation de la précision dans le cas d'un SNP, notre étude suggère l'intérêt d'étudier davantage et d'incorporer, éventuellement, l'incertitude liée à l'estimation des coefficients comparés pour une meilleure performance de la méthode basée sur la confusion positive. En effet, même si Glymour et al.⁵ ont conceptualisé leur méthode en assumant un échantillon infini de sorte que les coefficients estimés soient égaux à leurs vraies valeurs dans la population, on sait qu'en pratique, les estimateurs de RM sont vulnérables au biais d'échantillon fini et seront moins précis

que les estimateurs de l'analyse conventionnelle.^{106,125} Par conséquent, il est possible que, dans certaines situations, la variation échantillonnale puisse mener à des conclusions erronées sur la présence ou non d'une invalidité de l'instrument. Ce sera le cas, notamment, lorsque la puissance statistique est faible (cas d'analyse de RM avec petit échantillon et un SNP comme instrument).

Néanmoins, l'étude comporte des limites. Premièrement, comme toute étude de simulation, l'interprétation des résultats doit être faite dans les limites des paramètres considérés. Même si nous avons étendu la portée de nos situations pour intégrer plus de situations pratiques par rapport à l'évaluation antérieure de Burgess,⁴¹ d'autres contextes, par exemple, incluant une variation de la force des instruments ou de l'amplitude des effets pléiotropiques, n'ont pas été étudiés et pourraient faire l'objet d'évaluations futures. Deuxièmement, pour des raisons pratiques, l'application de la méthode a été restreinte aux seules études ayant utilisé la PAS, la PAD, l'hypertension ou une condition de santé mentale comme issue d'intérêt, limitant ainsi la portée de notre évaluation. Toutefois, les défis de l'interprétation de la méthode révélés par notre illustration ne sont pas spécifiques aux issues retenues.

Troisièmement, diverses situations pratiques pouvant affecter l'interprétation de cette méthode n'ont pas été abordées dans nos simulations. Des études futures pourraient examiner l'ampleur des erreurs de détection d'invalidité de l'instrument pouvant être commises par cette méthode dans divers scénarios considérant la non-satisfaction de l'hypothèse de confusion positive, ce qui est une situation plausible en pratique. De même, l'impact d'autres causes de variations potentielles des estimations ponctuelles qu'on compare dans cette méthode, par exemple, l'effet d'autres types de biais sur les estimations, pourrait davantage être approfondi.

Étude 3

L'étude 3 de la présente thèse est la première étude à appliquer et comparer plusieurs méthodes de prise en compte de la médication dans la modélisation de la PA en RM. En tant que telle, cette étude est originale et a permis de montrer que les différentes méthodes de gestion de la médication peuvent affecter différenciellement l'estimation de l'effet causal d'un facteur de risque sur la PA en RM, et ce, indépendamment de la validité de l'instrument.

De plus, le fait de limiter l'ajustement de nos modèles pour les covariables âge et sexe (en plus des composantes principales pour la stratification de la population) est en alignement avec les lignes directrices de la RM³⁵ et pourrait avoir contribué à améliorer la précision des estimations sans risque de biais car l'âge et le sexe ne peuvent être ni collisionneurs, ni médiateurs sur un chemin causal.¹⁵²

En outre, même si notre étude est avant tout méthodologique, réalisée pour illustrer l'impact des méthodes de gestion de la médication, et bien que l'association entre l'IMC et la PAS/hypertension ait fait l'objet de plusieurs études antérieures,^{43-52,268-270} nos analyses sur les données de l'ÉLCV font de cette étude l'une des premières études à tester et confirmer, au moyen des techniques de RM, l'association causale entre l'IMC et la PAS/hypertension dans un groupe d'adultes canadiens dont 80% d'hypertendus consommaient des médicaments antihypertenseurs. Dans cette optique, cette étude peut servir de référence pour d'autres études futures, considérant l'ampleur de l'obésité et de l'hypertension au sein de la population canadienne (affectent environ un-quarts de la population canadienne adulte, respectivement).^{55,262}

Toutefois, certaines limites peuvent être présentées. Premièrement, l'étude n'a considéré qu'un nombre limité (cinq) de méthodes de gestion de la médication couramment appliquées dans les études de RM. Il existe d'autres méthodes d'imputation des valeurs mesurées de la PA évaluées dans l'étude de simulation Tobin et al.,⁵³ et même d'autres encore, plus raffinées, qui tiennent compte du nombre et de la nature (classe pharmacologique) des médicaments antihypertenseurs consommés⁵⁴ que nous n'avons pas examinées du fait qu'elles n'ont pas été utilisées antérieurement dans des études de RM. Des futurs travaux pourraient élargir l'évaluation à ces autres méthodes non prises en compte dans notre étude. Deuxièmement, notre étude n'a pas évalué systématiquement par des simulations l'impact des méthodes de gestion de la médication sur le biais et la précision des estimations de l'association entre l'IMC et la PAS en RM. Bien que nos résultats s'alignent sur l'étude de simulation de Tobin et al.⁵³ suggérant que les résultats des modèles de régression conventionnels peuvent s'appliquer aux études de RM, des études de simulation sont nécessaires pour confirmer cette hypothèse.

Troisièmement, le fait de restreindre nos analyses aux participants d'ascendance européenne et à ceux ayant des valeurs d'IMC ≤ 40 kg/m² peut affecter la généralisation de nos résultats (validité externe). Dans le premier cas, il est admis que la génétique de la PA peut différer entre groupes

ethniques. C'est le cas notamment des disparités de la PA observées entre les Afro-américains et les Caucasiens qui sont en partie dues à des différences de fréquences alléliques des gènes régulant l'homéostasie du sodium.²⁷¹ Ainsi donc, l'exclusion des sujets d'ascendance non européenne nous a permis de conserver, dans notre étude, une population homogène et prévenir le biais dû à la stratification de la population.¹¹⁶ Dans le second cas, nous nous sommes concentrés sur des valeurs de l'IMC ≤ 40 kg/m² car l'association entre l'IMC et la PAS y était linéaire, et donc un seul coefficient était estimé, ce qui a facilité la comparaison des méthodes de gestion de la médication. Des approches pour estimer des associations non-linéaires en RM existent,¹³⁶ mais requièrent d'estimer plus d'un paramètre pour mieux capturer les effets non linéaires, ce qui n'aurait pas facilité la comparaison des méthodes.

Quatrièmement, nous ne pouvons exclure un potentiel biais de sélection lié possiblement aux pertes des participants (pertes au suivi et données manquantes), à la restriction de la sélection sur la base de l'exposition (IMC) et à l'inclusion particulièrement importante des participants qui consommaient de l'alcool par rapport à ceux qui n'en consommaient pas (toutes proportions gardées). A cet effet, il est admis, comme d'ailleurs le cas dans les analyses par modélisation conventionnelle, que l'inférence relative à l'effet causal en RM peut être invalide si la population de l'étude n'est pas représentative de la population cible, en raison notamment de la sélection, de l'abandon, de l'analyse des sous-groupes, de l'attrition au cours du suivi ou des données manquantes.¹⁴⁷ Lors de nos analyses, nous avons exclu 9667 sujets, soit le tiers (32%) des 30097 participants de l'ÉLCV évalués au recrutement, du fait de la perte au suivi (n = 2332), des données manquantes ou invalides (n = 4948) ou d'autres critères d'inclusion fixés (être d'ascendance européenne et avoir un IMC ≤ 40 kg/m², n = 2387). De plus, même si le fait de restreindre l'inclusion aux sujets ayant un IMC ≤ 40 kg/m² nous a facilité la comparaison des méthodes (comme précédemment mentionné), la sélection des participants sur base de l'exposition (IMC) seule ou sur base de l'IMC et d'un facteur de confusion de l'association entre l'IMC et la PAS (ex. consommation d'alcool), peut être une source de biais de sélection dans notre étude. Hughes et al.¹⁴⁷ ont examiné différents mécanismes par lesquels la sélection des participants pouvait induire un biais dans une étude de RM. Ces auteurs ont montré que l'estimateur 2SLS de la RM est biaisé, entre autres, lorsque la sélection des participants se fait sur base de l'exposition seule (voir Figure 1, scénario D) ou, à la fois, sur base de l'exposition et d'un facteur de confusion de l'association entre l'exposition et l'issue (voir Figure 1, scénario E).¹⁴⁷ Cela s'explique du fait que dans un

graphique acyclique orienté de la RM (**Figure 1**), l'exposition (X) est un collisionneur entre l'instrument (G) et les facteurs de confusion (U) de l'association entre X et l'issue (Y). Ainsi, conditionner sur X (ou, à la fois, sur X et sur un U mesuré) peut biaiser l'estimateur de la RM car l'association G-Y sera confondue dans l'échantillon sélectionné.¹⁴⁷ Toutefois, vu que la distribution de l'IMC (exposition d'intérêt) et celle de la PAS (issue d'intérêt) sont similaires entre le groupe des sujets inclus dans notre échantillon analytique et celui des autres qui étaient exclus de l'étude, et vu que, dans l'échantillon analytique, la distribution de l'IMC (moyenne : 27,2 vs. 28,0 kg/m²) et celle de la PAS (moyenne : 118,5 vs. 119,7 mm Hg) sont similaires entre le groupe des sujets qui consommaient de l'alcool vs. celui des autres qui n'en consommaient pas, respectivement, l'impact d'un possible biais de sélection serait marginal.

Cinquièmement, il est possible qu'il se soit glissées des erreurs de mesure de l'IMC (exposition) ou de la PAS (issue), ou des erreurs de classification de l'hypertension (issue). Ces erreurs seraient non différentielles car elles seraient indépendantes vis-à-vis de l'issue (pour les erreurs de mesure de l'IMC) ou de l'exposition (pour les erreurs de mesure/classification de la PAS/hypertension).^{153,272} Dans la définition de l'hypertension, outre d'éventuelles erreurs pouvant résulter de la mesure des PAS et PAD, l'auto-reportage du diagnostic de l'hypertension peut être une source importante d'erreurs de classification. En effet, une revue systématique et méta-analyse de 22 études ayant évalué la validité de l'auto-reportage du diagnostic de l'hypertension à identifier correctement les sujets hypertendus, a rapporté une sensibilité de 42% et une spécificité de 90%.²⁷³ En combinant l'auto-reportage du diagnostic de l'hypertension et de l'utilisation de la médication antihypertensive dans notre définition de l'hypertension,²²⁰ cela devrait réduire d'éventuelles erreurs de classification liées à l'auto-rapportage de l'hypertension. Toutefois, il est important de souligner que d'éventuelles erreurs de mesure non différentielles de l'IMC ou de la PAS, ou même des erreurs de classification non différentielles de l'hypertension ne devraient pas affecter substantiellement nos résultats du fait de la robustesse des estimateurs de RM.¹⁵³

Sixièmement, en fin, l'impact potentiel des facteurs de confusion de l'association entre l'IMC et la PAS sur nos estimations de RM peut se discuter en distinguant les facteurs de confusion post-conception (qui ne peuvent causer le variant génétique) des facteurs de l'ascendance qui peuvent induire une confusion génétique entre l'instrument et l'issue (stratification de la population). L'approche de la RM est protégée de l'effet confondant des facteurs de confusion post-conception

du fait de l'instrumentation de l'exposition par les variants génétiques qui sont assignés dès la conception, et précèdent chronologiquement les variables confondantes post-conception.^{67-69,80} Cette protection contre la confusion constitue d'ailleurs l'une des principales plus-values de la RM par rapport aux modèles de régression conventionnelle.⁸⁰ Cependant, si l'instrument est invalide et affecte l'issue via ces facteurs de confusion (pléiotropie indirecte), l'effet causal estimé sera biaisé par violation de l'hypothèse 3 de la RM.³⁴ Nos analyses ont suggéré que les deux instruments de l'IMC utilisés (GRS et *FTO*) n'affectaient pas l'ensemble des potentiels facteurs de confusion mesurés. De plus, la triangulation de différentes méthodes d'évaluation de la restriction d'exclusion n'a pas identifié le GRS comme un instrument invalide. *FTO*, de sa part, a été suspecté d'être un instrument invalide affectant la PAS en dehors de l'IMC, non via les potentiels facteurs de confusion mesurés, mais plutôt via la variable médication qui serait un médiateur entre *FTO* et la PAS. Ceci suggère que les facteurs de confusion post-conception de l'association entre l'IMC et la PAS auraient, somme toute, un impact marginal sur nos résultats. Finalement, il est connu que la stratification de la population viole l'hypothèse 2 de la RM du fait de la confusion génétique entre l'instrument et l'issue.^{34,113} Dans notre étude, le fait de restreindre l'échantillon analytique au seul sous-groupe homogène des participants d'ascendance Européenne et d'ajuster les modèles pour les 10 premières composantes principales génomiques du sous-ensemble des participants d'ascendance européenne,²¹⁶ devrait contribuer à limiter substantiellement le biais potentiel de nos résultats pouvant résulter de la stratification de la population.^{115,116}

5.4. Conclusion

L'atteinte de la mission de la santé publique de réduire l'apparition des maladies et maintenir la population en santé implique l'accomplissement de plusieurs fonctions essentielles de la santé publique incluant la gestion et la planification en santé publique, à travers notamment l'utilisation des niveaux de preuves scientifiques dans l'élaboration et la mise en œuvre des politiques de santé publique et la recherche en santé publique.^{261,274} La promotion et la gestion de la recherche et des connaissances en santé sont donc essentielles en santé publique.²⁷⁵

La présente thèse s'inscrit dans le transfert des connaissances vers les chercheurs en RM en vue d'améliorer la qualité de la recherche et générer ainsi les meilleures données scientifiques devant soutenir la prise de décision en santé publique.²⁷⁶

La thèse met de l'avant les difficultés méthodologiques dans les applications de l'approche de la RM et l'importance de la triangulation de plusieurs méthodes lors de la vérification des hypothèses de RM et de réfléchir à chaque étape. La thèse met aussi de l'avant l'importance, pour les chercheurs, de bien rapporter les résultats de leurs études de RM pour permettre aux lecteurs de faire une lecture critique.

Le champ de la RM est en plein essor et des nouvelles méthodes sont souvent proposées. Il devient important non seulement de les évaluer, mais aussi d'en détailler l'utilisation et les hypothèses sous-jacentes pour une utilisation optimale en complément aux méthodes existantes.

Références bibliographiques

1. Rivas-Ruiz F, Perez-Vicente S, Gonzalez-Ramirez A. Bias in clinical epidemiological study designs. *Allergol Immunopathol (Madr)* 2013; **41**: 54-9.
2. Dumas O, Siroux V, Le Moual N, Varraso R. Approches d'analyse causale en épidémiologie. *Rev Epidemiol Sante Publique* 2014; **62**: 53-63.
3. Rutter M. Epidemiological methods to tackle causal questions. *Int J Epidemiol* 2009; **38**: 3-6.
4. Streeter AJ, Lin NX, Crathorne L, et al. Adjusting for unmeasured confounding in nonrandomized longitudinal studies: a methodological review. *J Clin Epidemiol* 2017; **87**: 23-34.
5. Glymour MM, Tchetgen Tchetgen EJ, Robins JM. Credible mendelian randomization studies: Approaches for evaluating the instrumental variable assumptions. *Am J Epidemiol* 2012; **175**: 332-9.
6. Spiga F, Gibson M, Dawson S, et al. Tools for assessing quality and risk of bias in Mendelian randomization studies: a systematic review. *Int J Epidemiol* 2022; **dyac149**.
7. Labrecque J, Swanson SA. Understanding the Assumptions Underlying Instrumental Variable Analyses: a Brief Review of Falsification Strategies and Related Tools. *Curr Epidemiol Rep* 2018; **5**: 214-20.
8. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nat Rev Methods Primers* 2022; **2**: 6.
9. Sekula P, Del Greco MF, Pattaro C, Kottgen A. Mendelian Randomization as an Approach to Assess Causality Using Observational Data. *J Am Soc Nephrol* 2016; **27**: 3253-65.
10. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; **32**: 1-22.
11. Hernán MA, Robins JM. Instruments for Causal Inference: An Epidemiologist's Dream? *Epidemiology* 2006; **17**: 360-72.
12. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet* 2018; **27**: R195-R208.

13. van Kippersluis H, Rietveld CA. Pleiotropy-robust Mendelian randomization. *Int J Epidemiol* 2018; **47**: 1279-88.
14. Dixon P, Hollingworth W, Harrison S, Davies NM, Smith GD. Mendelian Randomization analysis of the causal effect of adiposity on hospital costs. *J Health Econ* 2020; **70**: 102300.
15. Lor GCY, Risch HA, Fung WT, et al. Reporting and guidelines for mendelian randomization analysis: A systematic review of oncological studies. *Cancer Epidemiol* 2019; **62**: 101577.
16. Pingault J-B, Cecil C, Murray J, Munafo M, Viding E. Causal inference in psychopathology: Using Mendelian randomisation to identify environmental risk factors for psychopathology. *Psychopathol Rev* 2016; **4**: 4-25.
17. Goodarzi MO. Genetics of obesity: what genetic association studies have taught us about the biology of obesity and its complications. *Lancet Diabetes Endocrinol* 2018; **6**: 223-36.
18. Frayling TM, Stoneman CE. Mendelian randomisation in type 2 diabetes and coronary artery disease. *Curr Opin Genet Dev* 2018; **50**: 111-20.
19. Meng X, Li X, Timofeeva MN, et al. Phenome-wide Mendelian-randomization study of genetically determined vitamin D on multiple health outcomes using the UK Biobank study. *Int J Epidemiol* 2019; **48**: 1425-34.
20. Guo J-Z, Xiao Q, Gao S, Li X-Q, Wu Q-J, Gong T-T. Review of Mendelian randomization studies on ovarian cancer. *Front Oncol* 2021; **11**: 681396.
21. Bochud M, Rousson V. Usefulness of Mendelian randomization in observational epidemiology. *Int J Environ Res Public Health* 2010; **7**: 711-28.
22. Davies NM, Smith GD, Windmeijer F, Martin RM. Issues in the reporting and conduct of instrumental variable studies: a systematic review. *Epidemiology* 2013; **24**: 363-9.
23. Boef AG, Dekkers OM, le Cessie S. Mendelian randomization studies: a review of the approaches used and the quality of reporting. *Int J Epidemiol* 2015; **44**: 496-511.
24. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell* 2017; **169**: 1177-86.
25. Chesmore K, Bartlett J, Williams SM. The ubiquity of pleiotropy in human disease. *Hum Genet* 2018; **137**: 39-44.

26. Verbanck M, Chen C-Y, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* 2018; **50**: 693-8.
27. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nat Commun* 2020; **11**: 376.
28. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* 2017; **46**: 1985-98.
29. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015; **44**: 512-25.
30. Minelli C, Del Greco MF, van der Plaats DA, Bowden J, Sheehan NA, Thompson J. The use of two-sample methods for Mendelian randomization analyses on single large datasets. *Int J Epidemiol* 2021: dyab084.
31. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *Ann Statist* 2020; **48**: 1742-69.
32. Rees JMB, Wood AM, Dudbridge F, Burgess S. Robust methods in Mendelian randomization via penalization of heterogeneous causal estimates. *PLoS One* 2019; **14**: e0222362.
33. Thompson JR, Minelli C, Bowden J, et al. Mendelian randomization incorporating uncertainty about pleiotropy. *Stat Med* 2017; **36**: 4627-45.
34. Richmond RC, Davey Smith G. Mendelian Randomization: Concepts and Scope. *Cold Spring Harb Perspect Med* 2022; **12**: a040501.
35. Burgess S, Davey Smith G, Davies NM, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res* 2020; **4**: 186.
36. Davey Smith G, Davies NM, Dimou N, et al. STROBE-MR: Guidelines for strengthening the reporting of Mendelian randomization studies. *PeerJ Preprints* 2019; 7: e27857v1 [Prépublication]. 2019 [cité le 12 oct 2022]: [7 p.]. <https://doi.org/10.7287/peerj.preprints.27857v1>.

37. Kivimäki M, Jokela M, Hamer M, et al. Examining overweight and obesity as risk factors for common mental disorders using fat mass and obesity-associated (FTO) genotype-instrumented analysis: The Whitehall II Study, 1985-2004. *Am J Epidemiol* 2011; **173**: 421-9.
38. Painter JN, O'Mara TA, Marquart L, et al. Genetic Risk Score Mendelian Randomization Shows that Obesity Measured as Body Mass Index, but not Waist:Hip Ratio, Is Causal for Endometrial Cancer. *Cancer Epidemiol Biomarkers Prev* 2016; **25**: 1503-10.
39. Walter S, Glymour MM, Koenen K, et al. Do genetic risk scores for body mass index predict risk of phobic anxiety? Evidence for a shared genetic risk factor. *Psychol Med* 2015; **45**: 181-91.
40. Walter S, Kubzansky LD, Koenen KC, et al. Revisiting Mendelian randomization studies of the effect of body mass index on depression. *Am J Med Genet B Neuropsychiatr Genet* 2015; **168B**: 108-15.
41. Burgess S. Re: "credible mendelian randomization studies: Approaches for evaluating the instrumental variable assumptions". *Am J Epidemiol* 2012; **176**: 456-7.
42. Mbutiwi FIN, Dessy T, Sylvestre MP. Mendelian Randomization: A Review of Methods for the Prevention, Assessment, and Discussion of Pleiotropy in Studies Using the Fat Mass and Obesity-Associated Gene as an Instrument for Adiposity. *Front Genet* 2022; **13**: 803238.
43. Dale CE, Fatemifar G, Palmer TM, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. *Circulation* 2017; **135**: 2373-88.
44. Holmes MV, Lange LA, Palmer T, et al. Causal effects of body mass index on cardiometabolic traits and events: a Mendelian randomization analysis. *Am J Hum Genet* 2014; **94**: 198-208.
45. Kivimaki M, Smith GD, Timpson NJ, et al. Lifetime body mass index and later atherosclerosis risk in young adults: examining causal links using Mendelian randomization in the Cardiovascular Risk in Young Finns study. *Eur Heart J* 2008; **29**: 2552-60.
46. Wade KH, Chiesa ST, Hughes AD, et al. Assessing the causal role of body mass index on cardiovascular health in young adults: Mendelian randomization and recall-by-genotype analyses. *Circulation* 2018; **138**: 2187-201.

47. Wurtz P, Wang Q, Kangas AJ, et al. Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. *PLoS Med* 2014; **11**: e1001765.
48. Lyall DM, Celis-Morales C, Ward J, et al. Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank A Mendelian Randomization Study. *JAMA Cardiol* 2017; **2**: 882-9.
49. Fall T, Hagg S, Magi R, et al. The Role of Adiposity in Cardiometabolic Traits: A Mendelian Randomization Analysis. *PLoS Med* 2013; **10**: e1001474.
50. Censin JC, Peters SAE, Bovijn J, et al. Causal relationships between obesity and the leading causes of death in women and men. *PLoS Genet* 2019; **15**: e1008405.
51. Fall T, Hagg S, Ploner A, et al. Age- and sex-specific causal effects of adiposity on cardiovascular risk factors. *Diabetes* 2015; **64**: 1841-52.
52. Timpson NJ, Harbord R, Davey Smith G, Zacho J, Tybjaerg-Hansen A, Nordestgaard BG. Does greater adiposity increase blood pressure and hypertension risk?: Mendelian randomization using the FTO/MC4R genotype. *Hypertension* 2009; **54**: 84-90.
53. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med* 2005; **24**: 2911-35.
54. Rana BK, Dhamija A, Panizzon MS, et al. Imputing observed blood pressure for antihypertensive treatment: impact on population and genetic analyses. *Am J Hypertens* 2014; **27**: 828-37.
55. DeGuire J, Clarke J, Rouleau K, Roy J, Bushnik T. Blood pressure and hypertension. *Health Rep* 2019; **30**: 14-21.
56. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *Lancet* 2021; **398**: 957-80.
57. Sudharsanan N, Geldsetzer P. Impact of Coming Demographic Changes on the Number of Adults in Need of Care for Hypertension in Brazil, China, India, Indonesia, Mexico, and South Africa. *Hypertension* 2019; **73**: 770-6.

58. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol* 2014; **43**: 922-9.
59. Gill R, Stratigopoulos G, Lee JH, Leibel RL. Functional genomic characterization of the FTO locus in African Americans. *Physiol Genomics* 2019; **51**: 517-28.
60. Keller L, Xu W, Wang HX, Winblad B, Fratiglioni L, Graff C. The obesity related gene, FTO, interacts with APOE, and is associated with Alzheimer's disease risk: a prospective cohort study. *J Alzheimers Dis* 2011; **23**: 461-9.
61. Calderón-Larrañaga A, Diaz E, Poblador-Plou B, Gimeno-Feliu LA, Abad-Díez JM, Prados-Torres A. Non-adherence to antihypertensive medication: The role of mental and physical comorbidity. *Int J Cardiol* 2016; **207**: 310-6.
62. Burnier M, Egan BM. Adherence in Hypertension. *Circ Res* 2019; **124**: 1124-40.
63. Sargan JD. The Estimation of Economic Relationships using Instrumental Variables. *Econometrica* 1958; **26**: 393-415.
64. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000; **29**: 722-9.
65. Wehby GL, Ohsfeldt RL, Murray JC. 'Mendelian randomization' equals instrumental variable analysis with genetic instruments. *Stat Med* 2008; **27**: 2745-9.
66. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008; **27**: 1133-63.
67. Burgess S, Swanson SA, Labrecque JA. Are Mendelian randomization investigations immune from bias due to reverse causation? *Eur J Epidemiol* 2021; **36**: 253-7.
68. Zheng J, Baird D, Borges MC, et al. Recent Developments in Mendelian Randomization Studies. *Curr Epidemiol Rep* 2017; **4**: 330-45.
69. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol* 2013; **178**: 1177-84.
70. Gala H, Tomlinson I. The use of Mendelian randomisation to identify causal cancer risk factors: promise and limitations. *J Pathol* 2020; **250**: 541-54.

71. Taylor AE, Davies NM, Ware JJ, VanderWeele T, Smith GD, Munafò MR. Mendelian randomization in health research: Using appropriate genetic variants and avoiding biased estimates. *Econ Hum Biol* 2014; **13**: 99-106.
72. Katan MB. Apolipoprotein E isoforms, serum cholesterol, and cancer. *Lancet* 1986; **327**: 507-8.
73. Katan MB. Apolipoprotein E isoforms, serum cholesterol, and cancer. 1986. *Int J Epidemiol* 2004; **33**: 9.
74. Tobin MD, Minelli C, Burton PR, Thompson JR. Commentary: Development of Mendelian randomization: from hypothesis test to 'Mendelian deconfounding'. *Int J Epidemiol* 2004; **33**: 26-9.
75. VanderWeele TJ, Tchetgen EJT, Cornelis M, Kraft P. Methodological challenges in Mendelian randomization. *Epidemiology* 2014; **25**: 427-35.
76. Wheatley K, Gray R. Commentary: Mendelian randomization—an update on its use to evaluate allogeneic stem cell transplantation in leukaemia. *Int J Epidemiol* 2004; **33**: 15-7.
77. Speed MS, Jepsen OH, Borglum AD, Speed D, Ostergaard SD. Investigating the association between body fat and depression via Mendelian randomization. *Transl Psychiatry* 2019; **9**: 184.
78. Wang N, Lu M, Chen C, et al. Adiposity genetic risk score modifies the association between blood lead level and body mass index. *J Clin Endocrinol Metab* 2018; **103**: 4005-13.
79. Gray R, Wheatley K. How to avoid bias when comparing bone marrow transplantation with chemotherapy. *Bone Marrow Transplant* 1991; **7 Suppl 3**: 9-12.
80. Smith GD. Capitalizing on Mendelian randomization to assess the effects of treatments. *J R Soc Med* 2007; **100**: 432-5.
81. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* 2017; **26**: 2333-55.
82. Qi Q, Qi L. Lipoprotein(a) and cardiovascular disease in diabetic patients. *Clin Lipidol* 2012; **7**: 397-407.
83. Verduijn M, Siegerink B, Jager KJ, Zoccali C, Dekker FW. Mendelian randomization: use of genetics to enable causal inference in observational studies. *Nephrol Dial Transplant* 2010; **25**: 1394-8.

84. Smith GD, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ* 2005; **330**: 1076-9. Figure 1; p. 1076.
85. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. *JAMA* 2021; **326**: 1614-21.
86. Sanderson E. Multivariable Mendelian Randomization and Mediation. *Cold Spring Harb Perspect Med* 2021; **11**: a038984.
87. Davies NM, Hill WD, Anderson EL, Sanderson E, Deary IJ, Smith GD. Multivariable two-sample mendelian randomization estimates of the effects of intelligence and education on health. *ELife* 2019; **8**: e43990.
88. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018; **362**: k601.
89. Nussbaum RL, McInnes RR, Willard HF. *Thompson and Thompson Genetics in Medicine*. 8th ed. Philadelphia (US): Elsevier; 2016.
90. Le Morvan V, Formento JL, Milano G, Bonnet J, Robert J. Techniques de recherche des polymorphismes génétiques. *Oncologie* 2005; **7**: 7-16.
91. Palmer TM, Lawlor DA, Harbord RM, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res* 2012; **21**: 223-42.
92. Burgess S, Thompson SG. *Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation*. 1st ed. Boca Raton (FL): Chapman and Hall/CRC; 2015.
93. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Stat Med* 2016; **35**: 1880-906.
94. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol* 2013; **42**: 1134-44.
95. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011; **40**: 755-64.
96. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* 2017; **32**: 377-89.

97. Slob EA, Burgess S. A comparison of robust Mendelian randomization methods using summary data. *Genetic Epidemiology* 2020; **44**: 313-29.
98. Swerdlow DI, Kuchenbaecker KB, Shah S, et al. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. *Int J Epidemiol* 2016; **45**: 1600-16.
99. Danchin A, Fang G. Unknown unknowns: essential genes in quest for function. *Microb Biotechnol* 2016; **9**: 530-40.
100. Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: A review. *Res Synth Methods* 2019; **10**: 486-96.
101. Gagliano Taliun SA, Evans DM. Ten simple rules for conducting a mendelian randomization study. *PLoS Comput Biol* 2021; **17**: e1009238.
102. Lousdal ML. An introduction to instrumental variable assumptions, validation and estimation. *Emerg Themes Epidemiol* 2018; **15**: 1.
103. Millard LAC, Smith GD, Tilling K. The global randomization test: A Mendelian randomization falsification test for the exclusion restriction assumption. medRxiv 2022.05.03.22274459 [Prépublication]. 2022 [cité le 27 oct 2022]: [20 p.]. Disponible: <https://doi.org/10.1101/2022.05.03.22274459>.
104. Lyon MS, Millard LAC, Smith GD, Hartwig F, Gaunt TR, Tilling K. Examining the evidence for Mendelian randomization homogeneity assumption violation using instrument association with exposure variance. medRxiv 2022.09.12.22279854 [Prépublication]. 2022 [cité le 8 oct 2022]: [36 p.]. Disponible: <https://www.medrxiv.org/content/10.1101/2022.09.12.22279854v1#p-5>.
105. Staiger D, Stock JH. Instrumental Variables Regression with Weak Instruments. *Econometrica* 1997; **65**: 557-86.
106. Burgess S, Thompson SG. Bias in causal estimates from Mendelian randomization studies with weak instruments. *Stat Med* 2011; **30**: 1312-23.
107. Wang S, Kang H. Weak-instrument robust tests in two-sample summary-data Mendelian randomization. *Biometrics* 2022; **78**: 1699-713.

108. Sanderson E, Spiller W, Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat Med* 2021; **40**: 5434-52.
109. Bowden J, Del Greco MF, Minelli C, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *Int J Epidemiol* 2018; **48**: 728-42.
110. Zhao QY, Wang JS, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *Ann Stat* 2020; **48**: 1742-69.
111. Bowden J, Del Greco MF, Minelli C, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *Int J Epidemiol* 2019; **48**: 728-42.
112. Noyce AJ, Nalls MA. Mendelian Randomization — the Key to Understanding Aspects of Parkinson's Disease Causation? *Mov Disord* 2016; **31**: 478-83.
113. Sanderson E, Richardson TG, Hemani G, Davey Smith G. The use of negative control outcomes in Mendelian randomization to detect potential population stratification. *Int J Epidemiol* 2021; **50**: 1350-61.
114. Brumpton B, Sanderson E, Heilbron K, et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat Commun* 2020; **11**: 3519.
115. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**: 904-9.
116. Bouaziz M, Ambroise C, Guedj M. Accounting for population stratification in practice: a comparison of the main strategies dedicated to genome-wide association studies. *PLoS One* 2011; **6**: e28845.
117. Swanson SA, Hernán MA. Commentary: how to report instrumental variable analyses (suggestions welcome). *Epidemiology* 2013; **24**: 370-4.
118. Baiocchi M, Cheng J, Small DS. Tutorial in biostatistics: instrumental variable methods for causal inference. *Stat Med* 2014; **33**: 2297-340.

119. Sainani KL. Instrumental Variables: Uses and Limitations. *PM R* 2018; **10**: 303-8.
120. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol* 2016; **40**: 597-608.
121. Angrist JD, Krueger AB. Split-Sample Instrumental Variables Estimates of the Return to Schooling. *J Bus Econ Stat* 1995; **13**: 225-35.
122. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013; **37**: 658-65.
123. Morrison J, Knoblauch N, Marcus JH, Stephens M, He X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. *Nat Genet* 2020; **52**: 740-7.
124. Qi G, Chatterjee N. Mendelian randomization analysis using mixture models for robust and efficient estimation of causal effects. *Nat Commun* 2019; **10**: 1941.
125. Davies NM, Holmes MV, Smith GD. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018; **362**: k601.
126. Harbord RM, Didelez V, Palmer TM, Meng S, Sterne JAC, Sheehan NA. Severity of bias of a simple estimator of the causal odds ratio in Mendelian randomization studies. *Stat Med* 2013; **32**: 1246-58.
127. Allman PH, Aban I, Long DM, et al. A novel Mendelian randomization method with binary risk factor and outcome. *Genet Epidemiol* 2021; **45**: 549-60.
128. Didelez V, Meng S, Sheehan NA. Assumptions of IV Methods for Observational Epidemiology. *Statist Sci* 2010; **25**: 22-40.
129. Cummings P. The relative merits of risk ratios and odds ratios. *Arch Pediatr Adolesc Med* 2009; **163**: 438-45.
130. Burgess S. Software code for performing instrumental variable analyses for Mendelian randomization investigations [En ligne]. Cambridge (GB): Academia; 9 nov 2015 [cité le 04 nov 2022]. Disponible: https://www.academia.edu/19281072/Software_code_for_performing_instrumental_variable_analyses_for_Mendelian_randomization_investigations.
131. Vansteelandt S, Bowden J, Babanezhad M, Goetghebeur E. On Instrumental Variables Estimation of Causal Odds Ratios. *Statist Sci* 2011; **26**: 403-22.

132. Cai B, Small DS, Have TR. Two-stage instrumental variable methods for estimating the causal odds ratio: analysis of bias. *Stat Med* 2011; **30**: 1809-24.
133. Palmer TM, Thompson JR, Tobin MD, Sheehan NA, Burton PR. Adjusting for bias and unmeasured confounding in Mendelian randomization studies with binary responses. *Int J Epidemiol* 2008; **37**: 1161-8.
134. Greenland S, Pearl J. Adjustments and their consequences—collapsibility analysis using graphical models. *Int Stat Rev* 2011; **79**: 401-26.
135. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol* 2015; **30**: 543-52.
136. Staley JR, Burgess S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. *Genet Epidemiol* 2017; **41**: 341-52.
137. Zhou A, Selvanayagam JB, Hyppönen E. Non-linear Mendelian randomization analyses support a role for vitamin D deficiency in cardiovascular disease risk. *Eur Heart J* 2022; **43**: 1731-9.
138. Sun YQ, Burgess S, Staley JR, et al. Body mass index and all cause mortality in HUNT and UK Biobank studies: linear and non-linear mendelian randomisation analyses. *BMJ* 2019; **364**: 11042.
139. Sulc J, Sjaarda J, Kutalik Z. Polynomial Mendelian randomization reveals non-linear causal effects for obesity-related traits. *HGG Adv* 2022; **3**: 100124.
140. Burgess S, Davies NM, Thompson SG. Instrumental variable analysis with a nonlinear exposure-outcome relationship. *Epidemiology* 2014; **25**: 877-85.
141. Swanson SA, Tiemeier H, Ikram MA, Hernán MA. Nature as a Trialist?: Deconstructing the Analogy Between Mendelian Randomization and Randomized Trials. *Epidemiology* 2017; **28**: 653-9.
142. Shi J, Swanson SA, Kraft P, Rosner B, De Vivo I, Hernán MA. Mendelian Randomization With Repeated Measures of a Time-varying Exposure: An Application of Structural Mean Models. *Epidemiology* 2022; **33**: 84-94.

143. Labrecque JA, Swanson SA. Interpretation and Potential Biases of Mendelian Randomization Estimates With Time-Varying Exposures. *Am J Epidemiol* 2019; **188**: 231-8.
144. Sanderson E, Richardson TG, Morris TT, Tilling K, Davey Smith G. Estimation of causal effects of a time-varying exposure at multiple time points through multivariable mendelian randomization. *PLoS Genet* 2022; **18**: e1010290.
145. Canan C, Lesko C, Lau B. Instrumental Variable Analyses and Selection Bias. *Epidemiology* 2017; **28**: 396-8.
146. Tönnies T, Kahl S, Kuss O. Collider Bias in Observational Studies. *Dtsch Arztebl Int* 2022; **119**: 107-22.
147. Hughes RA, Davies NM, Davey Smith G, Tilling K. Selection Bias When Estimating Average Treatment Effects Using One-sample Instrumental Variable Analysis. *Epidemiology* 2019; **30**: 350-7.
148. Coscia C, Gill D, Benítez R, Pérez T, Malats N, Burgess S. Avoiding collider bias in Mendelian randomization when performing stratified analyses. *Eur J Epidemiol* 2022; **37**: 671-82.
149. Gkatzionis A, Burgess S. Contextualizing selection bias in Mendelian randomization: how bad is it likely to be? *Int J Epidemiol* 2019; **48**: 691-701.
150. Barry C, Liu J, Richmond R, et al. Exploiting collider bias to apply two-sample summary data Mendelian randomization methods to one-sample individual level data. *PLoS Genet* 2021; **17**: e1009703.
151. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol* 2018; **47**: 226-35.
152. Burgess S, Thompson SG. *Mendelian Randomization: Methods for Causal Inference Using Genetic Variants*. 2nd ed. Boca Raton (FL): Chapman and Hall/CRC; 2021.
153. Pierce BL, VanderWeele TJ. The effect of non-differential measurement error on bias, precision and power in Mendelian randomization studies. *Int J Epidemiol* 2012; **41**: 1383-93.

154. Vansteelandt S, Babanezhad M, Goetghebeur E. Correcting Instrumental Variables Estimators for Systematic Measurement Error. *Stat Sin* 2009; **19**: 1223-46.
155. Little MP, Patel A, Lee C, Hauptmann M, Berrington de Gonzalez A, Albert P. Impact of Reverse Causation on Estimates of Cancer Risk Associated With Radiation Exposure From Computerized Tomography: A Simulation Study Modeled on Brain Cancer. *Am J Epidemiol* 2022; **191**: 173-81.
156. Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur J Epidemiol* 2018; **33**: 947-52.
157. Li Y. Modern epigenetics methods in biological research. *Methods* 2021; **187**: 104-13.
158. Weinhold B. Epigenetics: the science of change. *Environ Health Perspect* 2006; **114**: A160-7.
159. Gayon J. From Mendel to epigenetics: History of genetics. *C R Biol* 2016; **339**: 225-30.
160. Burgess S. Statistical issues in Mendelian randomization: use of genetic instrumental variables for assessing causal associations [thèse en ligne]. Cambridge (GB): University de Cambridge; 2012 [cité le 6 déc 2022]. Disponible: <https://doi.org/10.17863/CAM.16219>.
161. He W, Castiblanco J, Walter EA, Okulicz JF, Ahuja SK. Mendelian randomization: potential use of genetics to enable causal inferences regarding HIV-associated biomarkers and outcomes. *Curr Opin HIV AIDS* 2010; **5**: 545-59.
162. Ogbuanu IU, Zhang H, Karmaus W. Can we apply the Mendelian randomization methodology without considering epigenetic effects? *Emerg Themes Epidemiol* 2009; **6**: 3.
163. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol* 2013; **42**: 1497-501.
164. Freeman G, Cowling BJ, Schooling CM. Power and sample size calculations for Mendelian randomization studies using one genetic instrument. *Int J Epidemiol* 2013; **42**: 1157-63.
165. Deng L, Zhang H, Yu K. Power calculation for the general two-sample Mendelian randomization analysis. *Genet Epidemiol* 2020; **44**: 290-9.
166. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* 2013; **14**: 483-95.

167. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I² statistic. *Int J Epidemiol* 2016; **45**: 1961-74.
168. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014; **23**: R89-R98.
169. Bennett DA, Holmes MV. Mendelian randomisation in cardiovascular research: an introduction for clinicians. *Heart* 2017; **103**: 1400-7.
170. Jordan DM, Verbanck M, Do R. HOPS: a quantitative score reveals pervasive horizontal pleiotropy in human genetic variation is driven by extreme polygenicity of human traits and diseases. *Genome Biol* 2019; **20**: 222.
171. Anreiter I, Sokolowski MB. Deciphering pleiotropy: How complex genes regulate behavior. *Commun Integr Biol* 2018; **11**: 1-4.
172. Cadby G, Melton PE, McCarthy NS, et al. Pleiotropy of cardiometabolic syndrome with obesity-related anthropometric traits determined using empirically derived kinships from the Busselton Health Study. *Hum Genet* 2018; **137**: 45-53.
173. Watanabe K, Stringer S, Frei O, et al. A global overview of pleiotropy and genetic architecture in complex traits. *Nat Genet* 2019; **51**: 1339-48.
174. Palmer TM, Ramsahai RR, Lawlor DA, Sheehan NA, Didelez V. Re: "credible mendelian randomization studies: Approaches for evaluating the instrumental variable assumptions". *Am J Epidemiol* 2012; **176**: 457-8.
175. Harrison S, Howe L, Davies AR. Making sense of Mendelian randomisation and its use in health research [En ligne]. Cardiff (GB): Public Health Wales NHS Trust & Bristol University; 2020 [cité le 20 jan 2023]. Disponible: <https://phw.nhs.wales/publications/publications1/making-sense-of-mendelian-randomisation-and-its-use-in-health-research/>.
176. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018; **7**: e34408.
177. Lawlor DA, Tilling K, Smith GD. Triangulation in aetiological epidemiology. *Int J Epidemiol* 2016; **45**: 1866-86.

178. Sung H, Siegel RL, Torre LA, et al. Global patterns in excess body weight and the associated cancer burden. *CA Cancer J Clin* 2019; **69**: 88-112.
179. Bendor CD, Bardugo A, Pinhas-Hamiel O, Afek A, Twig G. Cardiovascular morbidity, diabetes and cancer risk among children and adolescents with severe obesity. *Cardiovasc Diabetol* 2020; **19**: 79.
180. Statistique Canada. Embonpoint et obésité chez les adultes, 2018 [En ligne]. Feuilles d'information de la santé, No 82-625-X au catalogue de Statistique Canada; juin 2019 [cité le 20 juin 2022]. Disponible: <https://www150.statcan.gc.ca/n1/fr/pub/82-625-x/2019001/article/00005-fra.pdf?st=D10MTIjT>
181. Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat Rev Genet* 2007; **8**: 657-62.
182. Scuteri A, Sanna S, Chen WM, et al. Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet* 2007; **3**: e115.
183. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; **518**: 197-206.
184. Cho YS, Go MJ, Kim YJ, et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat Genet* 2009; **41**: 527-34.
185. Wen W, Cho YS, Zheng W, et al. Meta-analysis identifies common variants associated with body mass index in east Asians. *Nat Genet* 2012; **44**: 307-11.
186. Monda KL, Chen GK, Taylor KC, et al. A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. *Nat Genet* 2013; **45**: 690-6.
187. Yang Q, Xiao T, Guo J, Su Z. Complex Relationship between Obesity and the Fat Mass and Obesity Locus. *Int J Biol Sci* 2017; **13**: 615-29.
188. Laber S, Cox RD. Commentary: FTO obesity variant circuitry and adipocyte browning in humans. *Front Genet* 2015; **6**: 318.
189. Ganef IM, Bos MM, van Heemst D, Noordam R. BMI-associated gene variants in FTO and cardiometabolic and brain disease: obesity or pleiotropy? *Physiol Genomics* 2019; **51**: 311-22.

190. Rzehak P, Scherag A, Grallert H, et al. Associations between BMI and the FTO gene are age dependent: results from the GINI and LISA birth cohort studies up to age 6 years. *Obes Facts* 2010; **3**: 173-80.
191. Haworth CM, Carnell S, Meaburn EL, Davis OS, Plomin R, Wardle J. Increasing heritability of BMI and stronger associations with the FTO gene over childhood. *Obesity (Silver Spring)* 2008; **16**: 2663-8.
192. Zhao X, Yang Y, Sun B-F, Zhao Y-L, Yang Y-G. FTO and Obesity: Mechanisms of Association. *Curr Diab Rep* 2014; **14**: 486.
193. Gerken T, Girard CA, Tung YC, et al. The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* 2007; **318**: 1469-72.
194. Labayen I, Ruiz JR, Ortega FB, et al. Association between the FTO rs9939609 polymorphism and leptin in European adolescents: a possible link with energy balance control. The HELENA study. *Int J Obes* 2011; **35**: 66-71.
195. Obradovic M, Sudar-Milovanovic E, Soskic S, et al. Leptin and Obesity: Role and Clinical Implication. *Front Endocrinol (Lausanne)* 2021; **12**: 585887.
196. Speakman JR, Rance KA, Johnstone AM. Polymorphisms of the FTO gene are associated with variation in energy intake, but not energy expenditure. *Obesity (Silver Spring)* 2008; **16**: 1961-5.
197. Claussnitzer M, Dankel SN, Kim KH, et al. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N Engl J Med* 2015; **373**: 895-907.
198. Zhao X, Yang Y, Sun BF, et al. FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res* 2014; **24**: 1403-19.
199. Äijälä M, Ronkainen J, Huusko T, et al. The fat mass and obesity-associated (FTO) gene variant rs9939609 predicts long-term incidence of cardiovascular disease and related death independent of the traditional risk factors. *Ann Med* 2015; **47**: 655-63.
200. Liu C, Mou S, Pan C. The FTO gene rs9939609 polymorphism predicts risk of cardiovascular disease: a systematic review and meta-analysis. *PLoS One* 2013; **8**: e71901.
201. Pausova Z, Syme C, Abrahamowicz M, et al. A common variant of the FTO gene is associated with not only increased adiposity but also elevated blood pressure in French Canadians. *Circ Cardiovasc Genet* 2009; **2**: 260-9.

202. Hertel JK, Johansson S, Sonestedt E, et al. FTO, type 2 diabetes, and weight gain throughout adult life: a meta-analysis of 41,504 subjects from the Scandinavian HUNT, MDC, and MPP studies. *Diabetes* 2011; **60**: 1637-44.
203. Li H, Kilpeläinen TO, Liu C, et al. Association of genetic variation in FTO with risk of obesity and type 2 diabetes with data from 96,551 East and South Asians. *Diabetologia* 2012; **55**: 981-95.
204. Li T, Wu K, You L, et al. Common variant rs9939609 in gene FTO confers risk to polycystic ovary syndrome. *PLoS One* 2013; **8**: e66250.
205. Cronin RM, Field JR, Bradford Y, et al. Phenome-wide association studies demonstrating pleiotropy of genetic variants within FTO with and without adjustment for body mass index. *Front Genet* 2014; **5**: 250.
206. Iles MM, Law MH, Stacey SN, et al. A variant in FTO shows association with melanoma risk not due to BMI. *Nat Genet* 2013; **45**: 428-32.
207. Delahanty RJ, Beeghly-Fadiel A, Xiang YB, et al. Association of obesity-related genetic variants with endometrial cancer risk: a report from the Shanghai Endometrial Cancer Genetics Study. *Am J Epidemiol* 2011; **174**: 1115-26.
208. van den Broek N, Treur JL, Larsen JK, Verhagen M, Verweij KJH, Vink JM. Causal associations between body mass index and mental health: a Mendelian randomisation study. *J Epidemiol Community Health* 2018; **72**: 708-10.
209. Sun L, Ma L, Zhang H, et al. Fto deficiency reduces anxiety-and depression-like behaviors in mice via alterations in gut microbiota. *Theranostics* 2019; **9**: 721-33.
210. Li H, Kilpeläinen TO, Liu C, et al. Association of genetic variation in FTO with risk of obesity and type 2 diabetes with data from 96,551 East and South Asians. *Diabetologia* 2012; **55**: 981-95.
211. Davey Smith G. Commentary: Random allocation in observational data: How small but robust effects could facilitate hypothesis-free causal inference. *Epidemiology* 2011; **22**: 460-3.
212. Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *Am J Epidemiol* 2004; **159**: 702-6.

213. Chen W, Qian L, Shi J, Franklin M. Comparing performance between log-binomial and robust Poisson regression models for estimating risk ratios under model misspecification. *BMC Med Res Methodol* 2018; **18**: 63.
214. Wang J. Dose as Instrumental Variable in Exposure–Safety Analysis Using Count Models. *J Biopharm Stat* 2012; **22**: 565-81.
215. Raina P, Wolfson C, Kirkland S, et al. Cohort Profile: The Canadian Longitudinal Study on Aging (CLSA). *Int J Epidemiol* 2019; **48**: 1752-3j.
216. Forgetta V, Li R, Darmond-Zwaig C, et al. Cohort profile: genomic data for 26 622 individuals from the Canadian Longitudinal Study on Aging (CLSA). *BMJ Open* 2022; **12**: e059021.
217. Raina PS, Wolfson C, Kirkland S. Canadian Longitudinal Study on Aging (CLSA). Protocol (version 3). Hamilton, ON, Canada: CLSA; 2008.
218. Canadian Longitudinal Study on Aging. Pulse Rate and Blood Pressure Measurement [En ligne]. Hamilton (ON): CLSA ÉLCV; fév 2017 [cité le 15 août 2022]. Document no SOP_DCS_0008, version 2.4. Disponible: <https://www.clsa-elcv.ca/doc/2303>.
219. Rabi DM, McBrien KA, Sapir-Pichhadze R, et al. Hypertension Canada’s 2020 Comprehensive Guidelines for the Prevention, Diagnosis, Risk Assessment, and Treatment of Hypertension in Adults and Children. *Can J Cardiol* 2020; **36**: 596-624.
220. van de Vijver SJM, Oti SO, Agyemang C, Gomez GB, Kyobutungi C. Prevalence, awareness, treatment and control of hypertension among slum dwellers in Nairobi, Kenya. *J Hypertens* 2013; **31**: 1018-24.
221. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology* 2017; **28**: 30-42.
222. Jabre JF, Bland JDP. Body mass index changes: an assessment of the effects of age and gender using the e-norms method. *BMC Med Res Methodol* 2021; **21**: 40.
223. Karageorgi S, Alsmadi O, Behbehani K. A review of adult obesity prevalence, trends, risk factors, and epidemiologic methods in Kuwait. *J Obes* 2013; **2013**: 378650.
224. Haidar YM, Cosman BC. Obesity epidemiology. *Clin Colon Rectal Surg* 2011; **24**: 205-10.

225. Sobal J, Hanson KL. Marital Status, Marital History, Body Weight, and Obesity. *Marriage & Family Review* 2011; **47**: 474-504.
226. Wang Y, Chen J, Wang K, Edwards CL. Education as an important risk factor for the prevalence of hypertension and elevated blood pressure in Chinese men and women. *Journal of Human Hypertension* 2006; **20**: 898-900.
227. Kornitzer M, Dramaix M, De Backer G. Epidemiology of Risk Factors for Hypertension. *Drugs* 1999; **57**: 695-712.
228. Lipowicz A, Lopuszanska M. Marital differences in blood pressure and the risk of hypertension among Polish men. *European Journal of Epidemiology* 2005; **20**: 421-7.
229. Costanian C, Edgell H, Ardern CI, Tamim H. Hormone therapy use in the Canadian Longitudinal Study on Aging: a cross-sectional analysis. *Menopause* 2018; **25**: 46-53.
230. Griffith LE, Raina P, Kanters D, et al. Frailty differences across population characteristics associated with health inequality: a cross-sectional analysis of baseline data from the Canadian Longitudinal Study on Aging (CLSA). *BMJ Open* 2021; **11**: e047945.
231. Mian O, Anderson LN, Belsky DW, et al. Associations of Adverse Childhood Experiences with Frailty in Older Adults: A Cross-Sectional Analysis of Data from the Canadian Longitudinal Study on Aging. *Gerontology* 2021: 1-10.
232. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol* 1993; **46**: 153-62.
233. Logan SL, Gottlieb BH, Maitland SB, Meegan D, Spriet LL. The Physical Activity Scale for the Elderly (PASE) questionnaire; does it predict physical health? *Int J Environ Res Public Health* 2013; **10**: 3967-86.
234. Hagiwara A, Ito N, Sawai K, Kazuma K. Validity and reliability of the Physical Activity Scale for the Elderly (PASE) in Japanese elderly people. *Geriatrics & Gerontology International* 2008; **8**: 143-51.
235. Hardin JW. The Robust Variance Estimator for Two-stage Models. *Stata J* 2002; **2**: 253 - 66.
236. Davey Smith G. Random allocation in observational data: how small but robust effects could facilitate hypothesis-free causal inference. *Epidemiology* 2011; **22**: 460-3.
237. Burgess S, Bowden J, Dudbridge F, Thompson S. Assessing the effectiveness of robust instrumental variable methods using multiple candidate instruments with application to

Mendelian randomization. arXiv 1606.03729v2 [Prépublication]. 2018 [cité le 19 fév 2023]: [45 p.]. Disponible: <https://arxiv.org/pdf/1606.03729.pdf>.

238. Abbasi A, Deetman PE, Corpeleijn E, et al. Bilirubin as a potential causal factor in type 2 diabetes risk: a Mendelian randomization study. *Diabetes* 2015; **64**: 1459-69.
239. Kunutsor SK, Frysz M, Verweij N, Kieneker LM, Bakker SJL, Dullaart RPF. Circulating total bilirubin and risk of non-alcoholic fatty liver disease in the PREVEND study: observational findings and a Mendelian randomization study. *Eur J Epidemiol* 2020; **35**: 123-37.
240. Lee YS, Cho Y, Burgess S, et al. Serum gamma-glutamyl transferase and risk of type 2 diabetes in the general Korean population: a Mendelian randomization study. *Hum Mol Genet* 2016; **25**: 3877-86.
241. Barning F, Abarin T. Assessing the Causality Factors in the Association between (Abdominal) Obesity and Physical Activity among the Newfoundland Population-A Mendelian Randomization Analysis. *Genet Epigenet* 2016; **8**: 15-24.
242. Klovaite J, Benn M, Nordestgaard BG. Obesity as a causal risk factor for deep venous thrombosis: a Mendelian randomization study. *J Intern Med* 2015; **277**: 573-84.
243. Tan AG, Kifley A, Flood VM, et al. Evaluating the associations between obesity and age-related cataract: a Mendelian randomization study. *Am J Clin Nutr* 2019; **110**: 969-76.
244. Siedlinski M, Jozefczuk E, Xu X, et al. White Blood Cells and Blood Pressure: A Mendelian Randomization Study. *Circulation* 2020; **141**: 1307-17.
245. Yu Z, Coresh J, Qi G, et al. A bidirectional Mendelian randomization study supports causal effects of kidney function on blood pressure. *Kidney Int* 2020; **98**: 708-16.
246. Linneberg A, Jacobsen RK, Skaaby T, et al. Effect of Smoking on Blood Pressure and Resting Heart Rate: A Mendelian Randomization Meta-Analysis in the CARTA Consortium. *Circ Cardiovasc Genet* 2015; **8**: 832-41.
247. Lin C, Sun Z, Mei Z, et al. The causal associations of circulating amino acids with blood pressure: a Mendelian randomization study. *BMC Medicine* 2022; **20**: 414.
248. Ding M, Huang T, Bergholdt HK, et al. Dairy consumption, systolic blood pressure, and risk of hypertension: Mendelian randomization study. *BMJ* 2017; **356**: j1000.

249. Smith GD, Lawlor DA, Harbord R, et al. Association of C-reactive protein with blood pressure and hypertension: life course confounding and mendelian randomization tests of causality. *Arterioscler Thromb Vasc Biol* 2005; **25**: 1051-6.
250. Roa-Díaz ZM, Asllanaj E, Amin HA, et al. Age at Natural Menopause and Blood Pressure Traits: Mendelian Randomization Study. *Journal of clinical medicine* 2021; **10**: 4299.
251. Skaaby T, Husemoen LL, Martinussen T, et al. Vitamin D status, filaggrin genotype, and cardiovascular risk factors: a Mendelian randomization approach. *PLoS One* 2013; **8**: e57647.
252. Borges MC, Hartwig FP, Oliveira IO, Horta BL. Is there a causal role for homocysteine concentration in blood pressure? A Mendelian randomization study. *Am J Clin Nutr* 2016; **103**: 39-49.
253. Ehret GB, Munroe PB, Rice KM, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* 2011; **478**: 103-9.
254. Seidell JC, Halberstadt J. The global burden of obesity and the challenges of prevention. *Ann Nutr Metab* 2015; **66 Suppl 2**: 7-12.
255. Ference BA, Majeed F, Penumetcha R, Flack JM, Brook RD. Effect of naturally random allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGCR, or both: a 2 × 2 factorial Mendelian randomization study. *J Am Coll Cardiol* 2015; **65**: 1552-61.
256. Kamiza AB, Fatumo S, Singini MG, Yeh CC, Chikowore T. Hepatitis B infection is causally associated with extrahepatic cancers: A Mendelian randomization study. *EBioMedicine* 2022; **79**: 104003.
257. Wong JYY, Zhang H, Hsiung CA, et al. Tuberculosis infection and lung adenocarcinoma: Mendelian randomization and pathway analysis of genome-wide association study data from never-smoking Asian women. *Genomics* 2020; **112**: 1223-32.
258. Huang SY, Yang YX, Kuo K, et al. Herpesvirus infections and Alzheimer's disease: a Mendelian randomization study. *Alzheimers Res Ther* 2021; **13**: 158.
259. Zhu G, Zhou S, Xu Y, et al. Chickenpox and multiple sclerosis: A Mendelian randomization study. *J Med Virol* 2023; **95**: e28315.

260. Baranova A, Cao H, Zhang F. Causal effect of COVID-19 on Alzheimer's disease: A Mendelian randomization study. *J Med Virol* 2023; **95**: e28107.
261. Pommier J, Grimaud O. Les fonctions essentielles de santé publique : histoire, définition et applications possibles. *Sante Publique* 2007; **19**: 9-14.
262. Nardocci M, Polsky JY, Moubarac J-C. Consumption of ultra-processed foods is associated with obesity, diabetes and hypertension in Canadian adults. *Can J Public Health* 2021; **112**: 421-9.
263. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. *Nat Rev Nephrol* 2020; **16**: 223-37.
264. Drozd D, Alvarez-Pitti J, Wójcik M, et al. Obesity and Cardiometabolic Risk Factors: From Childhood to Adulthood. *Nutrients* 2021; **13**: 4176.
265. Statistique Canada [En ligne]. Ottawa (ON): Statistique Canada; 2020. Embonpoint et obésité selon l'indice de masse corporelle mesuré, selon le groupe d'âge et le sexe [modifié le 13 juillet 2023; cité le 13 juillet 2023]; [environ 4 écrans]. Disponible: <https://doi.org/10.25318/1310037301-fra>.
266. Falkner B. The Childhood Role in Development of Primary Hypertension. *Am J Hypertens* 2018; **31**: 762-9.
267. Frontiers in Genetics [En ligne]. Lausanne (CH): Frontiers; 2023. Mendelian Randomization: A Review of Methods for the Prevention, Assessment, and Discussion of Pleiotropy in Studies Using the Fat Mass and Obesity-Associated Gene as an Instrument for Adiposity [date de mise à jour inconnue; cité le 12 avril 2023]; [environ 2 écrans]. Disponible: <https://loop-impact.frontiersin.org/impact/article/803238#totalviews/views>.
268. Koh HB, Heo GY, Kim KW, et al. Trends in the association between body mass index and blood pressure among 19-year-old men in Korea from 2003 to 2017. *Sci Rep* 2022; **12**: 6767.
269. Ostchega Y, Hughes JP, Terry A, Fakhouri THI, Miller I. Abdominal Obesity, Body Mass Index, and Hypertension in US Adults: NHANES 2007–2010. *Am J Hypertens* 2012; **25**: 1271-8.
270. Hu G, Barengo NC, Tuomilehto J, Lakka TA, Nissinen A, Jousilahti P. Relationship of Physical Activity and Body Mass Index to the Risk of Hypertension: A Prospective Study in Finland. *Hypertension* 2004; **43**: 25-30.

271. Gleiberman L. Sodium, blood pressure, and ethnicity: what have we learned? *American journal of human biology : the official journal of the Human Biology Council* 2009; **21**: 679-86.
272. Fosgate GT. Non-differential measurement error does not always bias diagnostic likelihood ratios towards the null. *Emerging Themes in Epidemiology* 2006; **3**: 7.
273. Gonçalves VSS, Andrade KRC, Carvalho KMB, Silva MT, Pereira MG, Galvao TF. Accuracy of self-reported hypertension: a systematic review and meta-analysis. *J Hypertens* 2018; **36**: 970-8.
274. Bettcher DW, Sapirie S, Goon EH. Essential public health functions: results of the international Delphi study. *World Health Stat Q* 1998; **51**: 44-54.
275. Pan American Health Organization [En ligne]. Washington (DC): PAHO. Essential Public Health Functions [cité le 03 mars 2023]; [environ 5 écrans]. Disponible: <https://www.paho.org/en/topics/essential-public-health-functions>.
276. Brownson RC, Fielding JE, Maylahn CM. Evidence-based public health: a fundamental concept for public health practice. *Annu Rev Public Health* 2009; **30**: 175-201.

Annexes

Annexe 1. Certificats d'éthique



Comité d'éthique de la recherche du CHUM
Pavillon R, 900 rue St-Denis, 3^e étage
Montréal (Québec) H2X 0A9

Le 28 juillet 2020

Docteure Marie-Pierre Sylvestre
a/s: M. Fiston Ikwa Ndol Mbutiwi
courriel: [REDACTED]

Objet:	20.148 – Approbation initiale-FINALE (évaluation déléguée)
	Vérification de la pléiotropie en randomisation mendélienne : évaluation méthodologique et application à l'estimation de l'effet causal de l'adiposité sur la pression artérielle

Docteur,

Nous accusons réception, des documents soumis en vue de l'approbation du projet mentionné en rubrique. Votre projet a été évalué en évaluation déléguée compte tenu qu'il s'agit d'un projet considéré à risque minimal pour les participants. Nous avons examiné les documents suivants:

- Formulaire de demande d'évaluation éthique d'un projet (formulaire 11)
- Cadre de gestion des Biobanques (clsa access agreement template) [date : 1er octobre 2019, version 1]
- Cadre de gestion des Biobanques (privacy policy) [date : 1er octobre 2019, version 1]
- Curriculum Vitae (Marie-Pierre Sylvestre)
- Formulaire d'information et de consentement [date : 1er octobre 2019, version 1]
- Formulaire d'information et de consentement (Comprehensive Approval 2018 Feb Submission) [date : 1er octobre 2019, version 1]
- Formulaire d'information et de consentement (Tracking Approval 2018 Feb Submission) [date : 1er octobre 2019, version 1]
- Autres documents (Considérations éthiques) [date : 1er octobre 2019, version 1]
- Curriculum Vitae (Mbutiwi Fiston Ikwa Ndol)
- Formulaire Évaluation Budgétaire et Budget [date : 07 octobre 2019]
- Soutien Service/Département du CHUM
- Autres documents (offre bourse FRQS) [date : 8 octobre 2019, version 1]
- Autres documents (Accès aux données pour étudiants) [date : 8 octobre 2019, version 1]
- Protocole de recherche [date : 9 octobre 2019, version 2]
- Autres documents (protocol baseline) [date : 28 octobre 2019, version 1]

- Autres documents (Data access protocols) [date : 28 octobre 2019, version 1]
- Autres documents (Comprehensive Approval 2018 Feb Submission) [date : 28 octobre 2019, version 1]
- Autres documents (Tracking Approval 2018 Feb Submission) [date : 28 octobre 2019, version 1]
- Autres documents (Mendelian Randomization with a Continuous Outcome) [date : 22 juillet 2020, version 1]

Le CÉR du CHUM confirme également que vous avez déposé les documents requis pour établir que votre projet de recherche a fait l'objet d'un examen scientifique dont le résultat est positif.

Le tout étant jugé satisfaisant, il nous fait plaisir de vous informer que la présente constitue l'approbation finale de votre projet de recherche, **valide pour un an à compter du 28 juillet 2020**.

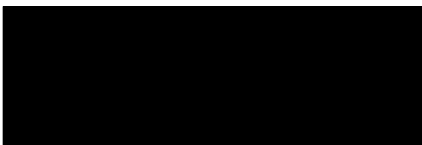
Vous devrez compléter le formulaire de renouvellement que nous vous ferons parvenir annuellement. De même, vous devrez soumettre pour approbation préalable, toute demande de modification ou document de suivi requis par le comité d'éthique conformément à ses Statuts et Règlements et ce via Nagano.

Veillez noter que le projet de recherche ne pourra débuter avant que vous n'ayez reçu la lettre de la personne mandatée pour autoriser cette recherche dans les murs de l'établissement. De même, lorsque cela s'applique à votre situation, le projet ne peut débuter tant que le contrat n'est pas finalisé et dûment signé.

Le comité d'éthique du CHUM est désigné par le gouvernement du Québec (MSSS) et adhère aux règles de constitution et de fonctionnement de l'Énoncé de Politique des trois Conseils (ÉPTC 2) et des Bonnes pratiques cliniques de la CIH.

Pour toute question relative à cette correspondance, veuillez communiquer avec la personne soussignée via NAGANO, ou avec le secrétariat du comité par téléphone ou courriel: ethique.recherche.chum@ssss.gouv.qc.ca – 514 890-8000, poste 14485, ou consulter le fichier «Questions-réponses» au bas de la page d'accueil Nagano.

Vous souhaitant la meilleure des chances dans la poursuite de vos travaux, nous vous prions d'accepter, Docteur, nos salutations distinguées.



M. Yves Poirier
Vice-président
Comité d'éthique de la recherche du CHUM



Comité d'éthique de la recherche du CHUM
Pavillon R, 900 rue St-Denis, 3^e étage
Montréal (Québec) H2X 0A9

Formulaire de demande de renouvellement annuel de l'approbation d'un projet de recherche

Titre du protocole : Vérification de la pléiotropie en randomisation mendélienne : évaluation méthodologique et application à l'estimation de l'effet causal de l'adiposité sur la pression artérielle

Chercheur principal (au CER Éval) : Marie-Pierre Sylvestre

Date de dépôt du formulaire : 2021-05-18 14:02

Déposé par : Mbutiwi, Fiston Ikwa Ndol

Date d'approbation du projet par le CER : 2020-07-28

Identifiant Nagano : Pléiotropie_en_RM

Numéro(s) de projet : 2021-8642, 20.148 - YP

Formulaire : F9-65985

Statut du formulaire : Approuvé

Suivi du BCER

1. **Statut de la demande:**

Demande approuvée

2. **La demande a été traitée par :**

Me Marie-Josée Bernardi

date de traitement:

2021-05-20

3. **Renouvellement accordé**

du 28 juillet 2021 au 28 juillet 2022

Déviations en cas de retard de renouvellement



Comité d'éthique de la recherche du CHUM
Pavillon R, 900 rue St-Denis, 3^e étage
Montréal (Québec) H2X 0A9

Formulaire de demande de renouvellement annuel de l'approbation d'un projet de recherche

Titre du protocole : Vérification de la pléiotropie en randomisation mendélienne : évaluation méthodologique et application à l'estimation de l'effet causal de l'adiposité sur la pression artérielle

Numéro(s) de projet : 2021-8642, 20.148 - YP

Formulaire : F9H-78863

Identifiant Nagano : Pléiotropie_en_RM

Date de dépôt initial du formulaire : 2022-06-06

Chercheur principal (au CER Éval) : Marie-Pierre Sylvestre

Date de dépôt final du formulaire : 2022-06-06

Date d'approbation du projet par le CER : 2020-07-28

Statut du formulaire : Formulaire approuvé

Suivi du BCER

1. **Statut de la demande:**

Demande approuvée

2. **La demande a été traitée par :**

Me Marie-Josée Bernardi

date de traitement:

2022-06-07

3. **Renouvellement accordé**

du 28 juillet 2022 au 28 juillet 2023

Déviations en cas de retard de renouvellement

Annexe 2. Autorisation de BMJ Publishing Group Ltd

09/04/2023 11:33

RightsLink Printable License

BMJ PUBLISHING GROUP LTD. LICENSE
TERMS AND CONDITIONS

Apr 09, 2023

This Agreement between University of Montreal -- Fiston Ikwa Ndol Mbutiwi ("You") and BMJ Publishing Group Ltd. ("BMJ Publishing Group Ltd.") consists of your license details and the terms and conditions provided by BMJ Publishing Group Ltd. and Copyright Clearance Center.

License Number	5403900669622
License date	Oct 07, 2022
Licensed Content Publisher	BMJ Publishing Group Ltd.
Licensed Content Publication	The BMJ
Licensed Content Title	What can mendelian randomisation tell us about modifiable behavioural and environmental exposures?
Licensed Content Author	George Davey Smith, Shah Ebrahim
Licensed Content Date	May 5, 2005
Licensed Content Volume	330
Type of Use	Dissertation/Thesis
Requestor type	Individual
Format	Print and electronic
Portion	Figure/table/extract

Number of figure/table/extracts	1
Description of figure/table/extracts	Figure 1 on page 1
Will you be translating?	No
Circulation/distribution	1
Title	Vérification de la pléiotropie en randomisation mendélienne : Évaluation méthodologique et application à l'estimation de l'effet causal de l'adiposité sur la pression artérielle
Institution name	University of Montreal
Expected presentation date	Dec 2022
Portions	Figure 1 on page 1
Requestor Location	University of Montreal Centre de recherche du CHUM Pavillon S, Bloc S03.706 850 Saint-Denis street Montreal, QC H2X 0A9 Canada Attn: Fiston Ikwa Ndol Mbutiwi
Publisher Tax ID	GB674738491
Total	0.00 CAD
Terms and Conditions	

Annexe 3. Auteurs des articles de la revue de littérature contactés

Identification de l'article	Personne contactée	Informations sollicitées
Dusingize JC, Olsen CM, An J, Pandeya N, Law MH, Thompson BS, Goldstein AM, Iles MM, Webb PM, Neale RE, Ong JS, MacGregor S, Whiteman DC. Body mass index and height and risk of cutaneous melanoma: Mendelian randomization analyses. <i>Int J Epidemiol.</i> 2020 Aug 1;49(4):1236-1245. doi: 10.1093/ije/dyaa009	Professor David Whiteman david.whiteman@qimrberghofer.edu.au	Les auteurs indiquent avoir retenu 730 SNPs pour l'IMC au final. Nous lister le(s) SNP(s) <i>FTO</i> .
Elgaeva EE, Tsepilov Y, Freidin MB, Williams FMK, Aulchenko Y, Suri P. ISSLS Prize in Clinical Science 2020. Examining causal effects of body mass index on back pain: a Mendelian randomization study. <i>Eur Spine J.</i> 2020 Apr;29(4):686-691. doi: 10.1007/s00586-019-06224-6.	Pradeep Suri Assistant professor and physician pradeep.suri@va.gov	Les auteurs indiquent avoir retenu 60 SNPs pour l'IMC qu'ils présentent dans le matériel supplémentaire. Cependant, l'image est une capture d'écran qui n'est pas lisible. Nous dire si parmi les 60 SNPs retenus il y avait des SNPs <i>FTO</i> . Si oui, lesquels ?
Fussey JM, Beaumont RN, Wood AR, Vaidya B, Smith J, Tyrrell J. Does Obesity Cause Thyroid Cancer? A Mendelian Randomization Study. <i>J Clin Endocrinol Metab.</i> 2020 Jul 1;105(7):e2398-407. doi: 10.1210/clinem/dgaa250	Jonathan Mark Fussey Department of Otolaryngology, Royal Devon and Exeter Hospital, Barrack Road, Exeter, UK. jfussey@nhs.net	Nous lister les SNPs <i>FTO</i> parmi les 73 SNPs retenus pour l'IMC.
Gharahkhani P, Ong JS, An J, Law MH, Whiteman DC, Neale RE, MacGregor S. Effect of increased body mass index on risk of diagnosis or death from cancer. <i>Br J Cancer.</i> 2019 Mar;120(5):565-570. doi: 10.1038/s41416-019-0386-9.	Puya Gharahkhani QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia. Puya.Gharahkhani@qimrberghofer.edu.au	Nous lister les SNPs <i>FTO</i> parmi les 520 SNPs retenus pour l'IMC.
Gharahkhani P, Ong JS, An J, Law MH, Whiteman DC, Neale RE, MacGregor S. Effect of increased body mass index on risk of diagnosis or death from cancer. <i>Br J Cancer.</i> 2019 Mar;120(5):565-570. doi: 10.1038/s41416-019-0386-9.	Associate Professor Stuart Macgregor Statistical Genetics Email: stuart.macgregor@qimrberghofer.edu.au	Nous lister les SNPs <i>FTO</i> parmi les 520 SNPs retenus pour l'IMC.
Gharahkhani P, Ong JS, An J, Law MH, Whiteman DC, Neale RE, MacGregor S. Effect of increased body mass index on risk of diagnosis or death from cancer. <i>Br J Cancer.</i> 2019 Mar;120(5):565-570. doi: 10.1038/s41416-019-0386-9.	Jue-Sheng Ong JueSheng.Ong@qimrberghofer.edu.au	Nous lister les SNPs <i>FTO</i> parmi les 520 SNPs retenus pour l'IMC.

<p>Jacobs BM, Noyce AJ, Giovannoni G, Dobson R. BMI and low vitamin D are causal factors for multiple sclerosis: A Mendelian Randomization study. <i>Neurol Neuroimmunol Neuroinflamm.</i> 2020 Jan 14;7(2):e662. doi: 10.1212/NXI.0000000000000662</p>	<p>Ruth Dobson From the Preventive Neurology Unit (B.M.J., A.J.N., G.G., R.D.), Wolfson Institute of Preventive Medicine, Barts and Queen Mary University of London; and Royal London Hospital (B.M.J., A.J.N., G.G., R.D.), Barts Health NHS Trust. ruth.dobson@qmul.ac.uk</p>	<p>Nous lister les SNPs <i>FTO</i> parmi les 941 SNPs retenus pour l'IMC adulte.</p>
<p>Kurz CF, Laxy M. Application of Mendelian Randomization to Investigate the Association of Body Mass Index with Health Care Costs. <i>Med Decis Making.</i> 2020 Feb;40(2):156-169. doi: 10.1177/0272989X20905809</p>	<p>Christoph F Kurz Institute of Health Economics and Health Care Management, Helmholtz Zentrum München, Neuherberg, Bayern, Germany christoph.kurz@helmholtz-muenchen.de</p>	<p>Nous lister les SNPs <i>FTO</i> parmi les 77 SNPs retenus pour l'IMC.</p>
<p>Speed MS, Jepsen OH, Børglum AD, Speed D, Østergaard SD. Investigating the association between body fat and depression via Mendelian randomization. <i>Transl Psychiatry.</i> 2019 Aug 5;9(1):184. doi: 10.1038/s41398-019-0516-4.</p>	<p>Maria Speed PhD Student maria@birc.au.dk</p>	<p>Nous lister les SNPs <i>FTO</i> parmi les 420 SNPs retenus pour l'IMC.</p>
<p>Speed MS, Jepsen OH, Børglum AD, Speed D, Østergaard SD. Investigating the association between body fat and depression via Mendelian randomization. <i>Transl Psychiatry.</i> 2019 Aug 5;9(1):184. doi: 10.1038/s41398-019-0516-4.</p>	<p>Søren D. Østergaard socoes@rm.dk</p>	<p>Nous lister les SNPs <i>FTO</i> parmi les 420 SNPs retenus pour l'IMC.</p>
<p>Martins-Silva T, Vaz JDS, Hutz MH, Salatino-Oliveira A, Genro JP, Hartwig FP, Moreira-Maia CR, Rohde LA, Borges MC, Tovo-Rodrigues L. Assessing causality in the association between attention-deficit/hyperactivity disorder and obesity: a Mendelian randomization study. <i>Int J Obes (Lond).</i> 2019 Dec;43(12):2500-2508. doi: 10.1038/s41366-019-0346-8.</p>	<p>Luciana Tovo-Rodrigues Postgraduate Program in Epidemiology, Universidade Federal de Pelotas, Pelotas, Brazil luciana.tovo@gmail.com</p>	<p>Dans les données de UK Biobank, combien de SNPs associés à l'IMC étaient exploités, et quels SNPs sur <i>FTO</i> étaient inclus ?</p>