RUNNING HEAD: BMI and type 2 diabetes on socioeconomic status

The effect of BMI and type 2 diabetes on socioeconomic status: a two-sample multivariable Mendelian randomization study

Sara Pedron, MSc[a, b, c, d], Christoph F Kurz, PhD[a, b, e], Lars Schwettmann, PD Dr.[a, f*], Michael Laxy, PhD[a, b, g, h*]

[a] *Institute of Health Economics and Health Care Management, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany*

[b] *German Center for Diabetes Research, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany*

[c] *Institute for Medical Information Processing, Biometry, and Epidemiology – IBE, Ludwig-Maximilians-University Munich, Marchioninistr. 15, 81377 Munich, Germany*

[d] *Pettenkofer School of Public Health, Munich, Germany*

[e] *Munich School of Management and Munich Center of Health Sciences, Ludwig-Maximilians-University Munich, Geschwister-Scholl-Platz 1, 80539 Munich, Germany*

[f] *Department of Economics, Martin Luther University Halle-Wittenberg, 06099 Halle (Saale), Germany*

[g] *Emory University, Rollins School of Public Health, 1518 Clifton Rd, Atlanta, GA 30322*

[h] *Technical University of Munich, Department of Sport and Health Sciences, Georg-Brauchle-Ring 60/62, 80992 Munich, Germany*

Corresponding author:
Sara Pedron
Institute of Health Economics and Health Care Management
Helmholtz Zentrum München, German Research Center for Environmental Health
Ingolstädter Landstr. 1, 85764 Neuherberg, Germany
Tel.: +49 890 3187 4603
Sara.pedron@helmholtz-muenchen.de

*Shared last authorship

Word count: 1333 (incl. abstract)
Tables: 1

ABSTRACT

OBJECTIVE

To assess the independent causal effect of BMI and type 2 diabetes (T2D) on socioeconomic outcomes applying two-sample Mendelian randomization (MR) analysis.

RESEARCH DESIGN AND METHODS

We carried out univariate and multivariate two-sample MR to jointly assess the effect BMI and T2D on socioeconomic outcomes. We used overlapping genome-wide significant single nucleotide polymorphisms (SNPs) for BMI and T2D as instrumental variables. Their causal impact on household income and regional deprivation was assessed using summary-level data from the UK Biobank.

RESULTS

In the univariate analysis, higher BMI was related with lower income (marginal effect of 1-SD increase in BMI [β=-0.092 (95% CI: -0.138; -0.047)] and higher deprivation [β=0.051 (95% CI: 0.022; 0.079)]. In the multivariate MR, the effect of BMI controlling for diabetes was slightly lower for income and deprivation. Diabetes was not associated with these outcomes.

CONCLUSIONS

High BMI, but not diabetes, shows a causal link with socioeconomic outcomes.

Previous evidence indicates that high BMI and type 2 diabetes (T2D) are associated with poorer labor market prospects, lower productivity and higher absenteeism (1-6). These disadvantages may accumulate over time and affect income and living circumstances, leading to a selection of individuals in more regionally deprived areas.

However, identifying the causal effect of BMI or diabetes on socioeconomic outcomes is challenging, mainly due to intrinsic problems of unmeasured confounding and reverse causation (1-3). Earlier approaches focused on the use of instrumental variable (IV) methods, exploiting the disease status of biological parents as IV (1-3). Recent studies have used genetic characteristics in one-sample Mendelian randomization (MR) approaches and showed an effect of BMI on socioeconomic status (4-6), while no effect of diabetes could be revealed (5).

This study aims at estimating the causal effect of BMI and T2D on household income and regional deprivation using a multivariate two-sample MR approach. This approach allows considering the shared genetic components of BMI and diabetes (7) to jointly estimate their causal effects on these socioeconomic outcomes (8).

RESEARCH DESIGN AND METHODS

Mendelian Randomization (MR)

The principle of MR roots in Mendel´s laws of inheritance, i.e. the individual genotype is largely independent of external factors and therefore independent of potential confounders. In MR techniques significant single nucleotide polymorphisms (SNPs) that are associated with the exposure are exploited as exogenous genetic variation in form of IVs (8, 9).

Genome-wide association studies (GWAS) have shown significant independent associations between several SNPs and BMI or T2D (10, 11), but also the presence of distinct signals influencing both conditions (7). While the relevance assumption and exclusion criteria are satisfied for our data (see Supplementary material 1), this overlap could lead to horizontal pleiotropy that violates the exchangeability assumption, i.e. the same SNP independently influences multiple phenotypes, resulting in biased estimates (9). Horizontal pleiotropy can be overcome by using multivariable MR methods, i.e. by considering the overlapping instruments directly in the estimation (8).

Data

For the associations between SNPs and socioeconomic outcomes, we used publicly available summary-level data from a GWAS of UK Biobank data (12, 13), including 464,708 individuals of European ancestry. Our outcomes were household income, defined as average total household income before tax, and regional deprivation, defined using the Townsend deprivation index (14) (Supplementary material 1).

Regarding the exposures, we utilized summary-level data on the associations between SNPs and BMI or T2D from published meta-analyses of GWAS (10, 11), excluding UK Biobank participants, as independency of data of the SNP-exposure and SNP-outcome association is a key prerequisite for the validity of the two-sample MR approach (9) (Supplementary materials 1 and 2).

Statistical analysis

First, we carried out a univariate MR analysis, testing the single effects of BMI and diabetes on the outcomes (8). Second, we estimated two-sample multivariate MR analysis of the effects of BMI and diabetes on the outcomes, using the set of overlapping SNPs as instruments (10, 11).

We estimated the effects using the *inverse-variance weighted (IVW)* method (9). Furthermore, we tested their robustness against other estimation methods, including *median based*, *MR Egger* and *MR-Robust Adjusted Profile Score* (RAPS) methods (Supplementary material 1 and 3). Moreover, we tested the sensitivity of the results by excluding other potentially pleiotropic SNPs (Supplementary material 4).

In both the univariate and the multivariate analyses, we tested the effect of two exposures on two outcomes. We therefore assumed a conservative Bonferroni-corrected p-value for statistical significance of 0.05/4=0.0125.

RESULTS

In total, we included 69 SNPs for BMI and 42 SNPs for T2D, which overlapped at two distinct loci: FTO and TCF7L2 (Supplementary Table S.2).

Results of the univariate MR analysis indicated that a higher BMI was associated with a lower household income [β= -0.092; 95% CI:-0.138; -0.047] and with a higher regional deprivation [β=0.051; 95% CI: 0.022; 0.079] (Table 1). Diabetes did not have any effect on the socioeconomic outcomes considered.

All analyses, except for BMI on income, presented low to middle levels of heterogeneity (I^2=0%-57%), indicating good validity of the instruments. The difference between MR Egger and IVW

estimates and a significant MR Egger intercept indicated the presence of horizontal pleiotropy, highlighting the need for multivariate MR analysis. The resulting effects from the multivariable MR analysis (Table 1) revealed that the direct effect of BMI controlling for diabetes was lower than in the univariate setting but still significant for both household income [β=-0.089; 95% CI: -0.13; -0.048] and regional deprivation [β=0.049; 95% CI: 0.023; 0.075]. Again, no effect of diabetes on socioeconomic outcomes could be observed.

The results from the MR Egger regression were almost identical to the estimates resulting from the IVW regression, indicating that the multivariable approach successfully accounted for the bias resulting from horizontal pleiotropy in the univariate setting.

[Table 1]

All results were robust to the use of alternative estimation methods (Supplementary material 3) and to the exclusion of other potentially pleiotropic SNPs (Supplementary material 4).

CONCLUSIONS

In this study, we estimated the independent effects of BMI and T2D on household income and regional deprivation using a novel multivariable MR technique (8). Our results indicate negative effects of BMI, but no effect of diabetes.

These findings strengthen the evidence of the deleterious role of BMI on income and regional deprivation, reported in previous observational and one-sample MR studies (1, 4-6). The potential underlying mechanisms include a lower ability-to-work, higher absenteeism, higher probability of musculoskeletal injuries and higher discrimination, which may lead to poorer career prospects,

decreasing labor market participation and lower income (1). A lower income could in turn affect living standards, leading individuals to self-select into more deprived areas with more affordable housing and food options.

Similar to a previous one-sample MR study (5), our results did not show any significant effect of T2D on household income or regional deprivation. In contrast, other studies that did not use a multivariate two-sample MR approach showed a negative effect of diabetes on socioeconomic outcomes (2, 3). This result should be object of further studies, aiming at establishing if this null effect can be replicated or if it is mainly due to methodological shortcomings in our study.

In fact, this paper entails some methodological limitations. First, despite being genetic characteristics largely independent of possible confounders, high BMI or diabetes genetic risk of parents might be an unmeasured confounder, causing a "dynastic bias" (15). Second, although the relevance assumption of our IVs is satisfied, the explanatory power of the set of SNPs used in the analysis for both the exposures and the outcomes is limited (10, 11). Finally, since the UK Biobank population is a selected one (13), our results might suffer from selection bias.

In conclusion, the present study provides evidence of a negative causal effect of higher BMI on income and regional deprivation, controlling for diabetes. In contrast, T2D does not have an effect on these two socioeconomic outcomes. Further studies should investigate this result, using new generations of GWAS with a higher explanatory power and including a more representative population. Furthermore, applied research may help to improve the understanding for the underlying mechanisms and to create targeted strategies to break the negative connection between BMI and socioeconomic outcomes.

Acknowledgements

Author contributions

S.P. formulated the research question, analyzed the data and wrote the manuscript. C.K. formulated the research question, provided support in the data analysis and reviewed/edited the manuscript. M.L. and L.S. contributed to the discussion and reviewed/edited the manuscript.

Acknowledgements

The authors thank Prof. Reiner Leidl, Ludwig-Maximilians-University of Munich und Helmholtz Zentrum München – German Research Center for Environmental Health, for important contributions and discussions regarding the research question, and Prof. Andrew Morris, University of Liverpool, for important support in the first stages of the analysis.

Guarantor´s name

Sara Pedron takes responsibility for the contents of the article.

Financial support

None.

Conflict of interest

All authors declare no conflicts of interest.

REFERENCES

1. Cawley J. An economy of scales: A selective review of obesity's economic causes, consequences, and solutions. *Journal of health economics*. 2015;43:244-68.
2. Pedron S, Emmert-Fees K, Laxy M, Schwettmann L. The impact of diabetes on labour market participation: a systematic review of results and methods. *BMC public health*. 2019;19(1):25.
3. Seuring T, Archangelidi O, Suhrcke M. The economic costs of type 2 diabetes: a global systematic review. *Pharmacoeconomics*. 2015;33(8):811-31.
4. Böckerman P, Cawley J, Viinikainen J, Lehtimäki T, Rovio S, Seppälä I, et al. The effect of weight on labor market outcomes: An application of genetic instrumental variables. *Health economics*. 2019;28(1):65-77.
5. Harrison S, Davies AR, Dickson M, Tyrrell J, Green MJ, Katikireddi SV, et al. Estimated effects of health conditions and risk factors on social and socioeconomic outcomes: mendelian randomisation of UK Biobank data. *The Lancet*. 2019;394:S49.
6. Tyrrell J, Jones SE, Beaumont R, Astley CM, Lovell R, Yaghootkar H, et al. Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. *BMJ*. 2016;352:i582.
7. Goodarzi MO. Genetics of obesity: what genetic association studies have taught us about the biology of obesity and its complications. *The Lancet Diabetes Endocrinology*. 2018;6(3):223-36.
8. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *American journal of epidemiology*. 2015;181(4):251-60.
9. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human molecular genetics*. 2014;23(R1):R89-R98.
10. 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(7538):197-206.
11. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An expanded genome-wide association study of type 2 diabetes in Europeans. *Diabetes*. 2017;66(11):2888-902.
12. Mitchell R, Elsworth B, Mitchell R, Raistrick C, Paternoster L, Hemani G, et al. MRC IEU UK Biobank GWAS pipeline version 2. 2019.
13. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *American journal of epidemiology*. 2017;186(9):1026-34.
14. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North: Routledge; 1988.
15. Fletcher JM. The promise and pitfalls of combining genetic and economic research. *Health economics*. 2011;20(8):889-92.

Tables

Table 1: MR results for the outcomes household income and regional deprivation.

	HH income (SD)						Deprivation (SD)					
method	**β**	**95% CI**	**pval**	**Q pval**	**I^2**	**Int pval**	**β**	**95% CI**	**pval**	**Q pval**	**I^2**	**Int pval**
Univariate MR												
BMI												
IVW	-0.092	[-0.138; -0.047]	0.000	0.000	71%		0.051	[0.022; 0.079]	0.001	0.000	57%	
MR Egger	-0.045	[-0.11; 0.019]	0.080	0.000	71%	0.359	0.013	[-0.036; 0.057]	0.282	0.000	55%	0.045
Type 2 diabetes												
IVW	-0.002	[-0.005; 0.008]	0.793	0.000	50%		0.002	[-0.005; 0.008]	0.634	0.524	0%	
MR Egger	-0.005	[-0.014; 0.02]	0.312	0.000	49%	0.702	0.003	[-0.012; 0.019]	0.346	0.516	0%	0.375
Multivariate MR												
BMI												
IVW	-0.089	[-0.13; -0.048]	0.000				0.049	[0.023; 0.075]	0.000			
MR Egger	-0.089	[-0.131; -0.048]	0.000				0.049	[0.023; 0.076]	0.000			
Type 2 diabetes												
IVW	-0.001	[-0.016; 0.013]	0.854				0.0004	[-0.009; 0.01]	0.940			
MR Egger	-0.001	[-0.016; 0.013]	0.866				0.0003	[-0.009; 0.01]	0.958			

Notes: β: marginal effect; CI: confidence interval; SD: standard deviation; HH income: household income; Q pval: p-value of the Cochrane´s Q statistic; IVW: inverse variance weighted estimator; MR: Mendelian randomization. Int pval: intercept p-value. The estimates for BMI indicate the change in the outcomes for a 1-SD increase in the genetically predicted level of BMI. The estimates for type 2 diabetes can be interpreted as the change in outcomes in response to a one unit increase in the loge odds of genetic risk of diabetes. All changes in outcome are expressed in SD units.