

## MODULATION OF CAUSAL EFFECTS

1           Widespread environment-specific causal effects detected in the UK Biobank  
2                    Leona Knüsel, Alice Man, Guillaume Paré, Zoltán Kutalik  
3           Department of Computational Biology, University of Lausanne, Switzerland  
4                    Swiss Bioinformatics Institute, Switzerland  
5           University Center for Primary Care and Public Health, 1005, Lausanne, Switzerland  
6                    Population Health Research Institute, McMaster University, Canada  
7

8                                    Leona.knuesel@gmail.com

9

10

|    |  |    |
|----|--|----|
| 11 | Table of Contents  |    |
| 12 | Abstract.....  | 4  |
| 13 | Introduction.....  | 5  |
| 14 | Causal Effects and the role of Mendelian Randomization.....                      | 5  |
| 15 | Interactions in MR .....   | 7  |
| 16 | Interactions in presence of non-linear causal effects .....                      | 8  |
| 17 | Methods.....   | 10 |
| 18 | Simulations .....  | 10 |
| 19 | Correction approach: 2SLS-I-corr .....   | 12 |
| 20 | Power analysis .....   | 13 |
| 21 | Application to UK biobank.....   | 14 |
| 22 | Cohort description.....  | 14 |
| 23 | Phenotype selection .....  | 14 |
| 24 | SNP selection and GRS calculation .....  | 15 |
| 25 | Inclusion of extended GRS (GRS <sub>ext</sub> ) to boost statistical power ..... | 16 |
| 26 | Mendelian Randomization Models.....  | 16 |
| 27 | Correction of the interaction .....  | 17 |
| 28 | Interaction tier score .....   | 17 |
| 29 | Sensitivity analyses.....  | 18 |
| 30 | Results.....   | 20 |

|    |  |    |
|----|--|----|
| 31 | Simulations .....  | 20 |
| 32 | Biased interaction effects for the uninstrumented regression model .....               | 20 |
| 33 | Biased interaction estimates in presence of quadratic effect of X on Y and effect of E |    |
| 34 | on X.....  | 21 |
| 35 | Accuracy of 2SLS-I-corr.....   | 22 |
| 36 | Power analysis of 2SLS-I.....  | 25 |
| 37 | Application.....   | 27 |
| 38 | Discussion.....  | 34 |
| 39 | References.....  | 42 |
| 40 | Supplementary Material.....  | 49 |
| 41 | Sensitivity Analyses .....   | 49 |
| 42 | Level 1 interactions.....  | 50 |
| 43 | Exposure on environment effects.....   | 51 |
| 44 | Inverse rank normal transformed outcome phenotypes .....                               | 52 |
| 45 | Replication analysis .....   | 55 |
| 46 |  |    |

47 Abstract

48 Background: Mendelian Randomization (MR) is a widely used tool to infer causal  
49 relationships. Yet, little research has been conducted on the elucidation of environment specific  
50 causal effects, despite mounting evidence for the relevance of causal effect modifying  
51 environmental variables.

52 Methods: To investigate potential modifications of causal effects, we extended two-stage-  
53 least-squares MR to investigate interaction effects (2SLS-I). We first tested 2SLS-I in a wide  
54 range of realistic simulation settings including quadratic and environment-dependent causal  
55 effects. Next, we applied 2SLS-I to investigate how environmental variables such as age,  
56 socioeconomic deprivation, and smoking modulate causal effects between a range of  
57 epidemiologically relevant exposure (such as systolic blood pressure, education, and body fat  
58 percentage) - outcome (e.g. forced expiratory volume (FEV1), CRP, and LDL cholesterol) pairs  
59 (in up to 337'392 individuals of the UK biobank).

60 Results: In simulations, 2SLS-I yielded unbiased interaction estimates, even in presence of non-  
61 linear causal effects. Applied to real data, 2SLS-I allowed for the detection of 182 interactions  
62 ( $P < 0.001$ ), with age, socioeconomic deprivation, and smoking being identified as important  
63 modifiers of many clinically relevant causal effects. For example, the positive causal effect of  
64 Triglycerides on systolic blood pressure was significantly attenuated in the elderly whilst the  
65 positive causal effect of Gamma-glutamyl transferase on CRP was intensified in smokers.

66 Conclusion: We present 2SLS-I, a method to simultaneously investigate environment-specific  
67 and non-linear causal effects. Our results highlight the importance of environmental variables in  
68 modifying well-established causal effects.

69 *Keywords:* [Mendelian Randomization, environmental modification of causal effects, age, SES]

## 70 **Introduction**

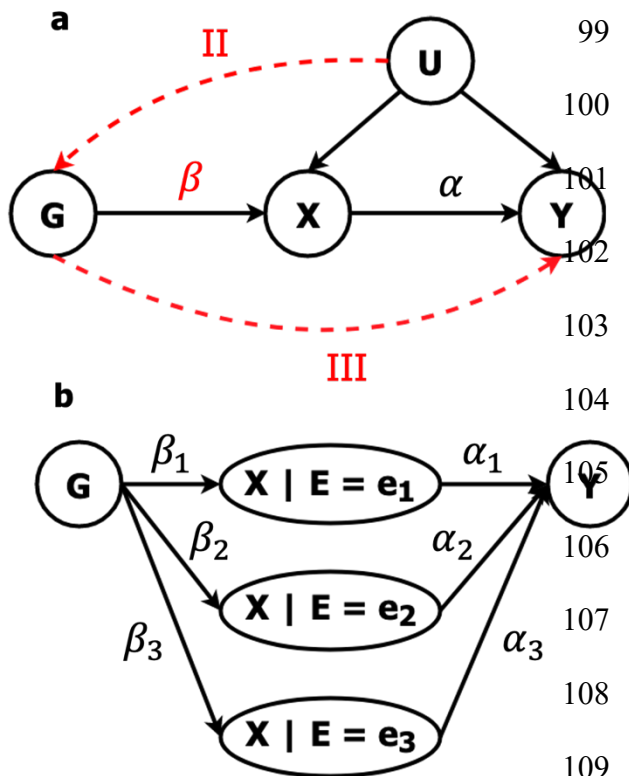
### 71 **Causal Effects and the role of Mendelian Randomization**

72 The distinction between associations and causal relationships is a central challenge in  
73 epidemiology, as understanding causal effects – in contrast to mere associations - allows for the  
74 design of effective interventions.

75 To characterize causal effects, randomized controlled trials (RCTs) are considered as the  
76 gold standard. In RCTs, individuals are allocated to a treatment or a placebo arm by chance.  
77 Therefore, observed differences between the groups result solely due to the treatment. Yet, RCTs  
78 are time consuming, resource-intensive, and at times impossible to conduct due to ethical  
79 considerations (Lawlor et al., 2008).

80 To address these shortcomings, Mendelian Randomization (MR) was developed. MR is a  
81 causal inference method that uses genetic variants as instrumental variables (IVs) to investigate  
82 the causal effect of an exposure on an outcome. As genetic variants are inherited randomly at  
83 birth, Mendelian Randomization is considered a natural genetic counterpart to randomized  
84 controlled trials. Like RCTs, MR is less prone to reverse causation and confounding than  
85 observational studies. In contrast to RCTs, MR requires no intervention and can be performed  
86 even on cross-sectional epidemiological data, allowing for a more diverse study sample than the  
87 typically restricted populations in RCTs (Lawlor et al., 2008). Furthermore, MR is cost- and  
88 time- efficient, in contrast to RCTs which require longitudinal data to investigate causal effects,  
89 which becomes particularly challenging to estimate life-time causal effects (e.g., effect of diet on  
90 a long-term outcome such as cancer) (Lawlor et al., 2008). Yet, it is important to note that MR  
91 estimates are prone to other sources of bias. To yield accurate causal estimates, MR relies on  
92 three main assumptions (*Figure 1a*): First, the relevance assumption: the genetic variants are

93 robustly associated with the outcome. Second, the exchangeability assumption: the genetic  
 94 variants are not associated with any confounder of the exposure-outcome relationship. Third,  
 95 exclusion restriction: the genetic variants are only associated with the outcome through the  
 96 exposure, i.e., the relationship between the genetic variants and the outcome is fully mediated  
 97 through the exposure (Lawlor et al., 2008) (Figure 1a). Whilst these are the standard  
 98 assumptions, MR has other implicit assumptions, such as gene-environment equivalence (a



99 change in the exposure should have the same  
 100 effect on the outcome, regardless of whether it  
 101 results from genetic or environmental  
 102 variation (Sanderson et al., 2022)), causal  
 103 effect linearity (the causal effect of an  
 104 exposure on an outcome is the same for all  
 105 levels of the exposure), and effect  
 106 homogeneity (the obtained causal effect is the  
 107 same for everyone (Sanderson et al., 2022)).  
 108 In this work, we aim to address one plausible  
 109 cause for the violation of the effect

homogeneity assumption, by considering how  
 environmental variables may modulate causal  
 effects (even in presence of non-linear  
 exposure-outcome relationships).

## 114 **Interactions in MR**

115           Recent studies indicate that environmental factors (E) such as age may modulate the  
116 observed causal estimates of an exposure (X) on an outcome (Y). For example, Richardson and  
117 colleagues (2020) investigated if the effect of body size on different health outcomes is age  
118 dependent. After obtaining distinct genetic instruments for adult and childhood body size, they  
119 fitted both uni- and multivariable MR models on a range of outcomes. If the effect of early life  
120 body size is significant in the univariable MR framework, but not in the multivariable MR  
121 framework, it indicates that the association is likely mediated through adult body size (and vice  
122 versa). Whilst the effect of early life body size on coronary heart disease was fully mediated  
123 through adult body size, they observed a strong protective effect of larger childhood body size on  
124 breast cancer risk, independent of adult body size. Richardson's approach allows for  
125 investigating the effect of the same exposure (e.g., body size) at different levels of the  
126 environment (e.g., age) on an outcome Y (e.g., coronary heart disease) (*Figure 1b*).

127           For Richardson's approach, it is necessary to obtain IVs to separately instrument the  
128 exposure at different levels of the environment. This requirement implies two major challenges.  
129 First, the levels of the environment are defined arbitrarily. Richardson and colleagues considered  
130 childhood body size and adult body size because data was available for these two time points in  
131 the UK Biobank (UKBB). Depending on the available data, artificial stratification of the  
132 environmental variable E is necessary, which likely would impact the results. Second, it is  
133 necessary to obtain IVs that differ in their association with the exposure depending on the  
134 environment E. For the ranges of age, socioeconomic status, smoking, air pollution, TV time,  
135 and physical activity that are available in the UKBB, we only found evidence for at least one

136 genome-wide significant gene-environment interaction effect for 22 out of 228 environment –  
137 exposure pairs. Therefore, the applicability of this approach is very limited.

### 138 **Interactions in presence of non-linear causal effects**

139 It has long been known that interactions can arise from unspecified nonlinear effects of  
140 correlated variables (Cortina, 1993; Matuschek & Kliegl, 2018). In simpler terms, if the exposure  
141 term  $X$  and the effect modifying variable  $E$  are correlated and  $X^2$  is associated with the outcome  
142  $Y$ , but the quadratic effect is not included in the regression model, the  $X \cdot E$  interaction term may  
143 be overestimated. Whilst the reverse is also true (i.e., in presence of a true interaction, if only the  
144 quadratic effect is included in the model, but not the interaction, the quadratic effect may be  
145 overestimated), the present work focuses on obtaining accurate interaction estimates and only  
146 considers nonlinear effect estimates as a nuisance parameter. Considering non-linear effects is of  
147 relevance as they play a role in a wide range of causal exposure-outcome relationships. A  
148 common example is the J-shaped association between BMI and all-cause mortality (Berrington  
149 de Gonzalez et al., 2010), which has not only been reported in observational, but also in MR  
150 studies (Sun et al., 2019). Furthermore, Sulc and colleagues (2021) estimated the causal effects  
151 of four anthropometric traits (such as body fat percentage) on seven health biomarkers (such as  
152 systolic blood pressure) and found significant evidence for nonlinearity for most of the  
153 investigated causal effects (84%).

154 Thus, to address the abovementioned limitations, we developed 2SLS-I, an approach to  
155 estimate exposure interactions with a non-instrumentable environmental variable  $E$  (e.g., age,  
156 socioeconomic status, physical activity), while accounting for potential nonlinear causal effects.  
157 We suggest investigating interactions by instrumenting  $X \cdot E$  with  $G \cdot E$ , where  $G$  is a genetic  
158 instrument for the exposure  $X$ . The advantage of this approach is that it does not require the



159 existence of genetic instruments for X that have different effects on X depending on the value of  
160 E and it allows for accounting for non-linear effects. In extensive simulation settings, we were  
161 able to obtain accurate main- and interaction estimates even in presence of sources of bias, such  
162 as linear and quadratic confounding, a quadratic exposure-outcome relationship, and a causal  
163 effect of the environment on the exposure. Application to a wide range of exposure (e.g. body fat  
164 percentage, education, and Vitamin D), environment (e.g. age, socioeconomic deprivation,  
165 physical inactivity), and outcome (e.g. fluid intelligence score, forced expiratory volume within 1  
166 s, CRP) combinations in the UK Biobank revealed multiple relevant settings where the  
167 environmental variable significantly modulated causal effects. For example, we found that  
168 multiple causal effects are attenuated with older age, whilst socioeconomic deprivation and  
169 smoking, in tendency, exacerbated detrimental causal effects.

170

171

## Methods

172 All simulations and analyses were performed using R version 4.2.1 “Funny looking kid”

173 (R Core Team, 2022). For the application of 2SLS-I, Snakemake version 7.25.3 (Mölder et al.,

174 2021) was used as a workflow manager.

### 175 Simulations

176 All simulations were performed with a sample size of 10’000 and 500 repetitions. We

177 simulated the data as follows.

178 *Equation 1*

$$179 \quad X = \beta_{E \rightarrow X} \cdot E + \beta_{GRS_X \rightarrow X} \cdot GRS_X + \beta_{U \rightarrow X} \cdot U + \varepsilon_X$$

180 *Equation 2*

$$181 \quad Y = \beta_{X \rightarrow Y} \cdot X + \beta_{X^2 \rightarrow Y} \cdot X^2 + \beta_{X \cdot E \rightarrow Y} \cdot X \cdot E + \beta_{U \rightarrow Y} \cdot U + \beta_{U^2 \rightarrow Y} \cdot U^2 + \beta_{U \cdot E \rightarrow Y} \cdot U \cdot E + \varepsilon_Y$$

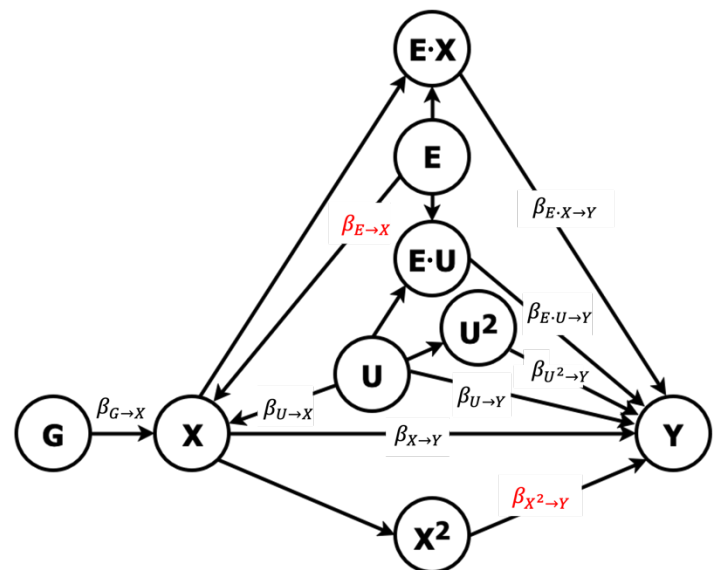
182 Where both E and U were drawn separately from a standard normal distribution. Thereby,

183 E represents the non-instrumentable effect modifying environment (e.g. age), whilst U represents

184 an unmeasured confounder. The GRS, representing a polygenic score for X, was drawn from a

*Table 1. Overview of the varied and constant parameters in simulations*

| Parameter                         | Values          |
|-----------------------------------|-----------------|
| $\beta_{E \rightarrow X}$         | {0, 0.1, 0.3}   |
| $\beta_{GRS_X \rightarrow X}$     | {1}             |
| $\beta_{X \cdot E \rightarrow Y}$ | {0, 0.1, 0.3}   |
| $\beta_{X \rightarrow Y}$         | {0.2}           |
| $\beta_{X^2 \rightarrow Y}$       | {0, 0.05, 0.15} |
| $\beta_{U \rightarrow X}$         | {0.3}           |
| $\beta_{U \rightarrow Y}$         | {0, 0.1, 0.3}   |
| $\beta_{U^2 \rightarrow Y}$       | {0, 0.1}        |
| $\beta_{U \cdot E \rightarrow Y}$ | {0, 0.1, 0.3}   |
| $h_X^2$                           | {0.1}           |



*Figure 2. Data simulation settings. G = genetic instruments for exposure X, X = exposure, U = confounding variable, E = non-instrumentable environment, Y = outcome.*

185 normal distribution with mean  $(\mu) = 0$  and variance  $(\sigma^2) = h_X^2$ , i.e., the variance of the GRS  
186 was set to be equal to the proportion of phenotypic variance explained by the polygenic score,  
187 which by default was set to 10%. The residual errors  $\varepsilon_X$  and  $\varepsilon_Y$  were drawn from a normal  
188 distribution with mean 0 and their respective variances were set to ensure that the overall  
189 variance of X and Y is equal to 1. For an overview of the varied variables, see *Table 1* and *Figure*  
190 *2*.

191 For all MR models,  $\beta_{U \rightarrow Y}$  was kept at 0.3 and  $\beta_{U^2 \rightarrow Y}$  at 0.1, invoking linear and quadratic  
192 confounding. Four models were fitted to investigate the accuracy of the interaction estimate in  
193 different settings. An observational model, and three different MR models, where we once  
194 omitted the quadratic exposure term, once omitted the interaction term, and once fitted the full  
195 model (*Table 2*). In the main text, we present estimates from the full MR model.

196 For the MR models, we instrument  $X \cdot E$  with  $G \cdot E$ , where G represents a genetic score  
197 based on the genome wide significant independent genetic instruments for X. Evidently, the MR  
198 assumptions need to be extended to  $G \cdot E$  for it to be a valid instrument of  $X \cdot E$ .

199

200 *Equation 3*

$$201 \quad Y \sim \beta_{X \rightarrow Y} \cdot \hat{X} + \beta_{E \rightarrow Y} \cdot E + \beta_{X * E \rightarrow Y} \cdot \hat{X} \cdot E + \beta_{X^2 \rightarrow Y} \cdot \hat{X}^2$$

202 with

203 *Equation 4*

$$204 \quad \hat{X} = \hat{\beta}_{GRS \rightarrow X} \cdot GRS$$

205 Table 2. Models fitted in simulations.  $Y$  refers to the outcome variable,  $X$  to the observational exposure variable,  $E$  to the (non-  
206 instrumentable) environmental variable and  $\hat{X}$  to the exposure variable when the effect is obtained using the GRS for  $X$ .

| Model                  | Formula  |     |
|------------------------|--|-----|
| Observational          | $Y \sim X + E + X^2 + E^2 + X \cdot E$               | 207 |
| MR without interaction | $Y \sim \hat{X} + E + E^2 + \hat{X}^2$               | 208 |
| MR without quadratic   | $Y \sim \hat{X} + E + E^2 + \hat{X} \cdot E$         | 209 |
| Full MR model          | $Y \sim \hat{X} + E + E^2 + \hat{X}^2 + GRS \cdot E$ | 210 |

## 211 Correction approach: 2SLS-I-corr

212 In simulations, we found a systematic bias of the interaction estimate if the causal effect  
213 modifying variable  $E$  had an effect on the exposure  $X$  and  $X^2$  had an effect on the outcome  $Y$ .  
214 From Equation 1 we obtain Equation 5.

215 Equation 5

$$216 \quad X^2 = (\beta_{GRS_{X \rightarrow X}} \cdot GRS_X)^2 + 2 \cdot \beta_{GRS_{X \rightarrow X}} \cdot GRS_X \cdot \beta_{E \rightarrow X} \cdot E + (\beta_{E \rightarrow X} \cdot E)^2 + 2 \cdot \beta_{GRS_{X \rightarrow X}} \cdot GRS_X \cdot$$

$$217 \quad \varepsilon_X + 2 \cdot \beta_{E \rightarrow X} \cdot E \cdot \varepsilon_X + \varepsilon_X^2$$

218 In Equation 5, it becomes evident that the MR interaction term  $GRS \cdot E$  is also  
219 represented in  $X^2$ , if  $\beta_{E \rightarrow X} \neq 0$ . As a consequence, the raw  $GRS \cdot E$  estimate is biased if  $X$  is a  
220 function of  $E$  and  $X^2$  influences  $Y$ . From this, one can derive the following correction term.

221 Equation 6

$$222 \quad \hat{\beta}_{GRS_{X \cdot E \rightarrow Y} \text{ corrected}} = \hat{\beta}_{GRS_{X \cdot E \rightarrow Y}} - \left( \frac{\hat{\beta}_{GRS_{X^2 \rightarrow Y}}}{(\hat{\beta}_{GRS_{X \rightarrow X}})^2} \cdot 2 \cdot \hat{\beta}_{GRS_{X \rightarrow X}} \cdot \hat{\beta}_{E \rightarrow X} \right)$$

223 Where the term in parentheses represents the extent to which  $\hat{\beta}_{GRS_{X \cdot E \rightarrow Y}}$  deviates from the  
224 true  $\beta_{GRS_{X \cdot E \rightarrow Y}}$  if  $\beta_{X^2 \rightarrow Y}$  and  $\beta_{E \rightarrow X}$  are not equal to 0.  $\frac{\hat{\beta}_{GRS_{X^2 \rightarrow Y}}}{(\hat{\beta}_{GRS_{X \rightarrow X}})^2}$  translates to the causal estimate of  
225  $\hat{\beta}_{X^2 \rightarrow Y}$ . The remaining terms represent the extent to which the true interaction is overestimated in  
226 response to the quadratic effect of  $X$  on  $Y$  (i.e. the extent to which  $X^2$  is a function of  $GRS \cdot E$ ),  
227 as visible in Equation 5 (marked in red). The full correction term from Equation 6 can be  
228 simplified as follows.

229 Equation 7

$$230 \quad \hat{\beta}_{GRS_{X \cdot E \rightarrow Y} \text{ corrected}} = \hat{\beta}_{GRS_{X \cdot E \rightarrow Y}} - 2 \cdot \left( \frac{\hat{\beta}_{GRS_{X^2 \rightarrow Y}}}{\hat{\beta}_{GRS_{X \rightarrow X}}} \cdot \hat{\beta}_{E \rightarrow X} \right)$$

231 From Equation 7, using the first order Taylor expansion-based approximation of the  
 232 variance of ratios, we can derive the variance of the corrected interaction term as follows.

233 Equation 8

$$\begin{aligned}
 234 \quad \hat{\sigma}_{GRS_X \cdot E \rightarrow Y}^2 &= \hat{\sigma}_{GRS_X \cdot E \rightarrow Y}^2 \\
 235 \quad &+ \left( \frac{\hat{\beta}_{GRS_X^2 \rightarrow Y}^2}{\hat{\beta}_{GRS_X \rightarrow X}^2} \cdot \left( \frac{\hat{\sigma}_{GRS_X^2 \rightarrow Y}^2}{\hat{\beta}_{GRS_X^2 \rightarrow Y}^2} + \frac{\hat{\sigma}_{GRS_X \rightarrow X}^2}{\hat{\beta}_{GRS_X \rightarrow X}^2} \right) \cdot \hat{\sigma}_{E \rightarrow X}^2 + \frac{\hat{\beta}_{GRS_X^2 \rightarrow Y}^2}{\hat{\beta}_{GRS_X \rightarrow X}^2} \right. \\
 236 \quad &\cdot \left. \left( \frac{\hat{\sigma}_{GRS_X^2 \rightarrow Y}^2}{\hat{\beta}_{GRS_X^2 \rightarrow Y}^2} + \frac{\hat{\sigma}_{GRS_X \rightarrow X}^2}{\hat{\beta}_{GRS_X \rightarrow X}^2} \right) \cdot \hat{\beta}_{E \rightarrow X}^2 + \hat{\sigma}_{E \rightarrow X}^2 \cdot \left( \frac{\hat{\beta}_{GRS_X^2 \rightarrow Y}^2}{\hat{\beta}_{GRS_X \rightarrow X}^2} + \frac{\hat{\sigma}_{GRS_X}^2 \cdot \hat{\beta}_{GRS_X^2 \rightarrow Y}^2}{\hat{\beta}_{GRS_X \rightarrow X}^3} \right)^2 \right) \cdot 2^2
 \end{aligned}$$

237 By setting  $\hat{\sigma}_{E \rightarrow X}^2$  to 0 (i.e. ignoring the variance in the E-to-X effect estimation), Equation  
 238 8 can be approximated as follows.

239 Equation 9

$$240 \quad \hat{\sigma}_{GRS_X \cdot E \rightarrow Y}^2 = \hat{\sigma}_{GRS_X \cdot E \rightarrow Y}^2 + 4 \cdot \frac{\hat{\beta}_{GRS_X^2 \rightarrow Y}^2}{\hat{\beta}_{GRS_X \rightarrow X}^2} \cdot \left( \frac{\hat{\sigma}_{GRS_X^2 \rightarrow Y}^2}{\hat{\beta}_{GRS_X^2 \rightarrow Y}^2} + \frac{\hat{\sigma}_{GRS_X \rightarrow X}^2}{\hat{\beta}_{GRS_X \rightarrow X}^2} \right) \cdot \hat{\beta}_{E \rightarrow X}^2$$

## 241 Power analysis

242 Table 3. Simulation settings of the statistical power analysis

| Parameter                      | Range                          |     |   |
|--------------------------------|--------------------------------|-----|---|
| $\beta_{E \rightarrow X}$      | {0.2}                          | 243 | The accuracy and power of the corrected   |
| $\beta_{GRS_X \rightarrow X}$  | {1}                            | 244 | interaction estimate was assessed. For the power                                  |
| $\beta_{GRS^*E \rightarrow Y}$ | {0, 0.05, 0.1}                 | 245 | analysis, $\beta_{E \rightarrow X}$ was set to 0.2, whilst the amount of          |
| $\beta_{GRS^*U \rightarrow Y}$ | {0}                            | 246 | variance that the GRS explained in X ( $h_X^2$ ), the                             |
| $\beta_{X \rightarrow Y}$      | {0.2}                          | 247 | true interaction ( $\beta_{GRS \cdot E \rightarrow Y}$ ), the quadratic effect of |
| $\beta_{X^2 \rightarrow Y}$    | {0, 0.1}                       | 248 | X on Y ( $\beta_{X^2 \rightarrow Y}$ ), and the sample size were varied           |
| $\beta_{U \rightarrow X}$      | {0.3}                          | 249 | systematically. The settings for the power analysis                               |
| $\beta_{U \rightarrow Y}$      | {0.5}                          | 250 | are listed in Table 3.  |
| $h_X^2$                        | {0.05, 0.1, 0.2}               |     |   |
| Sample Size                    | { $10^4, 5 \cdot 10^4, 10^5$ } | 251 |   |

## 252 **Application to UK biobank**

253 We applied 2SLS-I to the UK biobank (UKBB) to investigate if interactions with  
254 environmental variables occur and how prone they are to bias induced by nonlinear effects of the  
255 exposure on the outcome.

## 256 **Cohort description**

257 UKBB is a volunteer-based biomedical cohort of ~500'000 individuals from the general UK  
258 population (Sudlow et al., 2015). The data was accessed through application number 16389. We  
259 selected unrelated, white British participants, for whom the inferred and reported sex aligned,  
260 leading to an initial sample size of  $n = 337'392$ .

## 261 **Phenotype selection**

262 We selected environmental variables which are difficult to instrument genetically but are  
263 likely to modulate causal relationships, namely age, current smoking, physical inactivity  
264 (obtained using accelerometer data, sedentary behaviour, "PA: sed"), air pollution (Nitrogen  
265 dioxide ("AP: NO<sub>2</sub>"), time spent watching TV ("TV time"), and socioeconomic deprivation  
266 (Townsend deprivation index, "TDI")). A range of outcome traits was selected to address  
267 different health-, and cognitive variables, namely: fluid intelligence score (FIS), reaction time  
268 (RT), systolic blood pressure (SBP), low-density lipoprotein cholesterol (LDL), hand grip  
269 strength (HGS), forced expiratory volume within 1 second (FEV1), and C-reactive protein  
270 (CRP). For each of the outcome traits, we searched the literature and the EpiGraphDB (Liu et al.,  
271 2021) to obtain exposures that might have a causal effect on any outcome of interest. Exposure  
272 traits were considered for inclusion if they were strictly numeric, conceptually different from the  
273 outcome of interest (e.g. we removed HDL cholesterol as exposure for LDL cholesterol as the  
274 two variables are tightly related) and had at least one genome wide significant variant. If a

275 numeric alternative for a categorical trait was available, the numeric alternative was considered  
276 (e.g. instead of *hypertension*, we used *systolic blood pressure*). If multiple potential exposures  
277 covered conceptually very similar traits, only one was selected. For an overview of the selected  
278 exposure-, effect modifying environment-, and outcome traits, and the according UKBB field  
279 IDs, see *Supplementary Tables 1-3*. Exposure and outcome phenotypes were corrected for age,  
280 age<sup>2</sup>, sex, sex\*age, and 40 genetic PCs and exposure phenotypes were inverse rank normal  
281 transformed (IRNT) (McCaw et al., 2020). Potentially effect modifying environmental  
282 phenotypes were corrected for age, age<sup>2</sup>, sex, and age\*sex, except age (only corrected for sex).  
283 Effect modifying environment and outcome phenotypes were standardized to have zero mean  
284 and unit variance.

#### 285 **SNP selection and GRS calculation**

286 For each exposure phenotype, we accessed summary statistics (Neale, 2017) and filtered  
287 for SNPs with a p-value < 10<sup>-4</sup>. Next, we clumped the selected SNPs using PLINK version 1.9 to  
288 obtain independent genetic variants for MR. For the remaining variants, we reassessed their  
289 association with the IRNT exposure phenotype of interest whilst correcting for age, sex, age<sup>2</sup>,  
290 and 40 PCs. For the GRS, we only kept SNPs which were genome wide significantly associated  
291 (p < 5\*10<sup>-8</sup>) with the exposure in our sample. The selected SNPs were used to obtain the GRS for  
292 each exposure of interest. The GRS was calculated as follows:  $GRS_j = \sum_i^m \beta_i \cdot dosage_{ij}$ , where  
293  $m$  is the number of genome wide significant SNPs,  $\beta_i$  is the obtained effect size of  $SNP_i$  on the  
294 exposure of interest and the  $dosage_{ij}$  is the number of copies of the coded allele for SNP  $i$  in  
295 individual  $j$  (Collister et al., 2022). The obtained GRS was scaled to have zero mean and unit  
296 variance. We further assessed if there was evidence for an interaction between the GRS and the  
297 environment E on the exposure X (level 1 interaction) at p < 0.05. We observed this in 77 out of

298 228 (33.8%) of the settings (of which 38 survive Bonferroni correction). In-depth analysis of the  
299 role of level 1 interactions revealed a negligible role for level 1 interactions (for details on  
300 simulations and application, see Supplementary Material *Level 1 interactions*).

### 301 **Inclusion of extended GRS ( $GRS_{ext}$ ) to boost statistical power**

302 As interaction analyses require strong effects and a large sample size to obtain sufficient  
303 power, and the application of the quadratic effect correction further reduces the power to detect  
304 true interactions, we additionally performed MR with an extended GRS ( $GRS_{ext}$ ) to boost for  
305 power.  $GRS_{ext}$  was calculated for all selected exposure phenotypes as described in Privé et al.  
306 (2022). All analyses were performed as for the strict GRS.

### 307 **Mendelian Randomization Models**

308 Next, the GRS was used to instrument the effect of each exposure on each outcome  
309 whilst considering the effect modifying environmental variables of interest. We fitted four  
310 different models for each setting, including the full model:  $Y \sim GRS_X + E + E^2 + GRS_X^2 + E * GRS_X + Cov$ ,  
311 a model without the quadratic term:  $Y \sim GRS_X + E + E^2 + E \cdot GRS_X + Cov$ , a  
312 model without the interaction term:  $Y \sim GRS_X + GRS_X^2 + Cov$ , and a minimal model:  
313  $Y \sim GRS_X + Cov$ , where Y denotes the outcome of interest, the GRS is the previously obtained  
314 GRS for X, E is the effect modifying environmental variable of interest, and Cov translates to the  
315 covariates in the model. For all settings where E was not age, the covariates were age, age<sup>2</sup>, sex,  
316 and relevant medication. If the environmental variable was age, the age-related terms were  
317 excluded from the covariates. Only relevant medications were corrected for, obtained using UK  
318 Biobank data fields 6153 and 6177 for women and men, respectively. For an overview of the  
319 applied medication correction, see *Supplementary Table 1*. We used the ratio estimates for the



320 causal effects (X-to-Y and X·E-to-Y), i.e.  $\hat{\beta}_{GRS \rightarrow Y}$  and  $\hat{\beta}_{GRS \cdot E \rightarrow Y}$  were divided by  $\hat{\beta}_{GRS \rightarrow X}$ , whilst  
321  $\hat{\beta}_{GRS^2 \rightarrow Y}$  was divided by  $\hat{\beta}_{GRS \rightarrow X}^2$ .

### 322 **Correction of the interaction**

323 For the correction of the interaction (2SLS-I-corr), separate models were fitted to obtain  
324  $\hat{\beta}_{GRS_X \rightarrow X}$  and  $\hat{\beta}_{E \rightarrow X}$  (and the according standard errors to obtain the variance). X was inverse-  
325 normal rank transformed and E was z-standardized, and the models were corrected for the same  
326 covariates as the previously discussed MR models (depending on the exposure X and the  
327 environment E: age, age<sup>2</sup>, sex, medication).  $\hat{\beta}_{GRS_X \cdot E \rightarrow Y}$  and  $\hat{\beta}_{GRS_X^2 \rightarrow Y}$  were obtained from the full  
328 MR model described above. 2SLS-I-corr was only applied as discussed in the section *Correction*  
329 *approach: 2SLS-I-corr* if  $\hat{\beta}_{GRS_X^2 \rightarrow Y}$  was nominally significant ( $p < 0.05$ ), otherwise the correction  
330 would just reduce power due to the increased estimator variance.

### 331 **Interaction tier score**

332 To gauge the reliability of the obtained interactions using both, the strict GRS and  
333 GRS<sub>ext</sub>, we developed a rating strategy. For each interaction, we assessed whether  $\hat{\beta}_{X \cdot E}$  differed  
334 significantly from  $\hat{\beta}_{X_{ext} \cdot E}$ . From this, tiers were computed as follows.

$$335 \quad Tier_{X \cdot E} = 3 \cdot (both_{sig} \cdot effect_{direction\_alignes}) + 2 \cdot (either_{sig} \cdot effect_{size\_alignes}) \\ 336 \quad \quad \quad + only\_GRS_{sig} + main_{sig}$$

337 Where all variables are binary.  $both_{sig}$  is equal to 1 if both, the GRS and the GRS<sub>ext</sub> yield  
338 significant (at  $p < 0.001$ ) interaction estimates, and 0 otherwise.  $effect_{direction\_alignes}$  is 1 if the  
339 sign of the effects obtained by the GRS and the GRS<sub>ext</sub> agrees, and 0 otherwise.  
340  $effect_{size\_alignes}$  is 1 if the effect sizes obtained by the GRS and the GRS<sub>ext</sub> do not differ  
341 significantly ( $p > 0.05$ ), and 0 otherwise.  $only\_GRS_{sig}$  is 1 if the GRS yielded significance at  $p <$

342 0.001, but the  $GRS_{ext}$  did not, and the effect size obtained by the GRS does not align with the  
343 effect size obtained by the  $GRS_{ext}$ .  $main_{sig}$  is 1 if there is Bonferroni corrected significant  
344 evidence for a causal effect of the exposure on the outcome. This approach gives the highest  
345 confidence for interactions which yield significance for both, the GRS and the  $GRS_{ext}$ , and if on  
346 top of that the sign of the obtained interaction effect aligns. If either the GRS or the  $GRS_{ext}$  yields  
347 a significant effect (at  $p < 0.001$ ), and the obtained effects do not differ significantly ( $p > 0.05$ ),  
348 the interaction would be classified as tier 2. If only the GRS yields a significant interaction (at  $p$   
349  $< 0.001$ ), but the  $GRS_{ext}$  estimate does not agree, the interaction is classified as tier 1. By this we  
350 aim to account for the higher reliability of the GRS in comparison to the  $GRS_{ext}$ . Finally, if an  
351 interaction yielded a tier score  $> 0$ , we increased the tier score by +1 if we found significant  
352 evidence ( $p_{bonferroni} < 0.05$ ) for a main effect of the exposure on the outcome.

### 353 **Sensitivity analyses**

354 To ensure the robustness of our findings, we performed extensive sensitivity analyses  
355 validating our results.

356 For example, we assessed whether there was evidence for a stronger effect of the  
357 exposure on the environment than vice versa and flagged up such interactions as this  
358 circumstance may lead to a (collider) bias in the interaction estimate. The according interactions  
359 are shaded out in the respective results plot. For details of the analysis, see Supplementary  
360 Material, *Exposure on environment effects*.

361 Furthermore, we assessed the robustness of our results to inverse rank normal  
362 transforming the outcome phenotypes. For details of the analysis, see Supplementary Material,  
363 *Inverse rank normal transformed outcome phenotypes*.

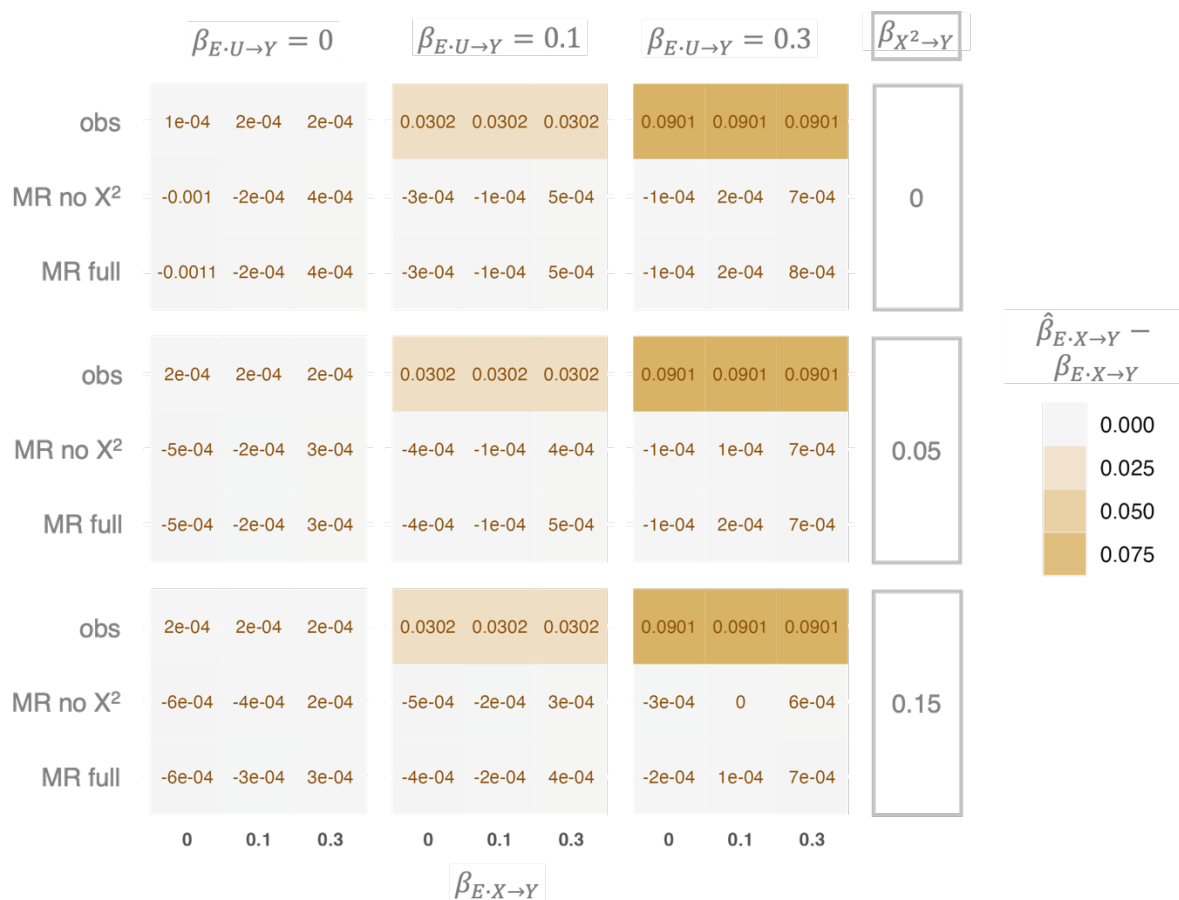
364           Finally, to investigate if some interactions are only picked up due to method specific  
365 features of 2SLS-I, we assessed the same settings in an extended doubly ranked stratification  
366 (Tian et al., 2023) framework (DRS). Whilst the doubly ranked stratification method was  
367 developed to investigate non-linear exposure-outcome relationships, it can be adapted to allow  
368 for investigating interactions by stratifying by an environmental variable (instead of the exposure  
369 value) adjusted by the exposure's instrument and regressing the obtained MR estimates on the  
370 predefined strata midpoint value. For a detailed effect comparison between the two methods, see  
371 Supplementary Material, *Replication analysis*.  
372

373

## Results

374 **Simulations**

375 **Biased interaction effects for the uninstrumented regression model**



376

377 *Figure 3. Deviation from the true simulated interaction effect for interaction estimates obtained from observational (obs) and two*  
 378 *different MR models (MR no X<sup>2</sup>: no quadratic effect of X on Y fitted, MR full: full MR model) in presence of an interaction*  
 379 *between the environment E and the confounder U and a quadratic effect of X on Y. The tile color indicates deviation of the*  
 380 *interaction estimate from the simulated interaction, with darker colors indicating stronger deviations.*

381 Before assessing the accuracy of interaction effects in a Mendelian Randomization (MR)

382 framework, we investigated under which circumstances MR may yield an advantage over

383 observational models. We found that interactions (in contrast to main effects ( $\hat{\beta}_{X \rightarrow Y}$ ) and

384 quadratic effects ( $\hat{\beta}_{X^2 \rightarrow Y}$ ), respectively) are robust against linear ( $\beta_{U \rightarrow X} = 0.3, \beta_{U \rightarrow Y} = 0.3$ ) and

385 quadratic ( $\beta_{U \rightarrow X} = 0.3, \beta_{U^2 \rightarrow Y} = 0.1$ ) confounding ( $M_{\text{diff}} = -0.00007, SD_{\text{diff}} = 0.009$ ). Yet, the

386 observational model yielded biased interaction estimates in presence of an interaction between

387 the confounder (U) and the environment (E) ( $\beta_{U \rightarrow X} = 0.3$ ,  $\beta_{U \cdot E \rightarrow Y} = 0.3$ ,  $M_{\text{diff}} = 0.9$ ,  $SD_{\text{diff}} =$   
 388 0.0099). In contrast, the interaction estimates obtained from MR were accurate even in the  
 389 strongest settings of confounding ( $\beta_{U \rightarrow X} = 0.3$ ,  $\beta_{U \rightarrow Y} = 0.3$ ,  $\beta_{U \cdot E \rightarrow Y} = 0.3$ ) ( $M_{\text{diff}} = -0.0003$ ,  
 390  $SD_{\text{diff}} = 0.036$ ).

391 **Biased interaction estimates in presence of quadratic effect of X on Y and effect of E**  
 392 **on X**

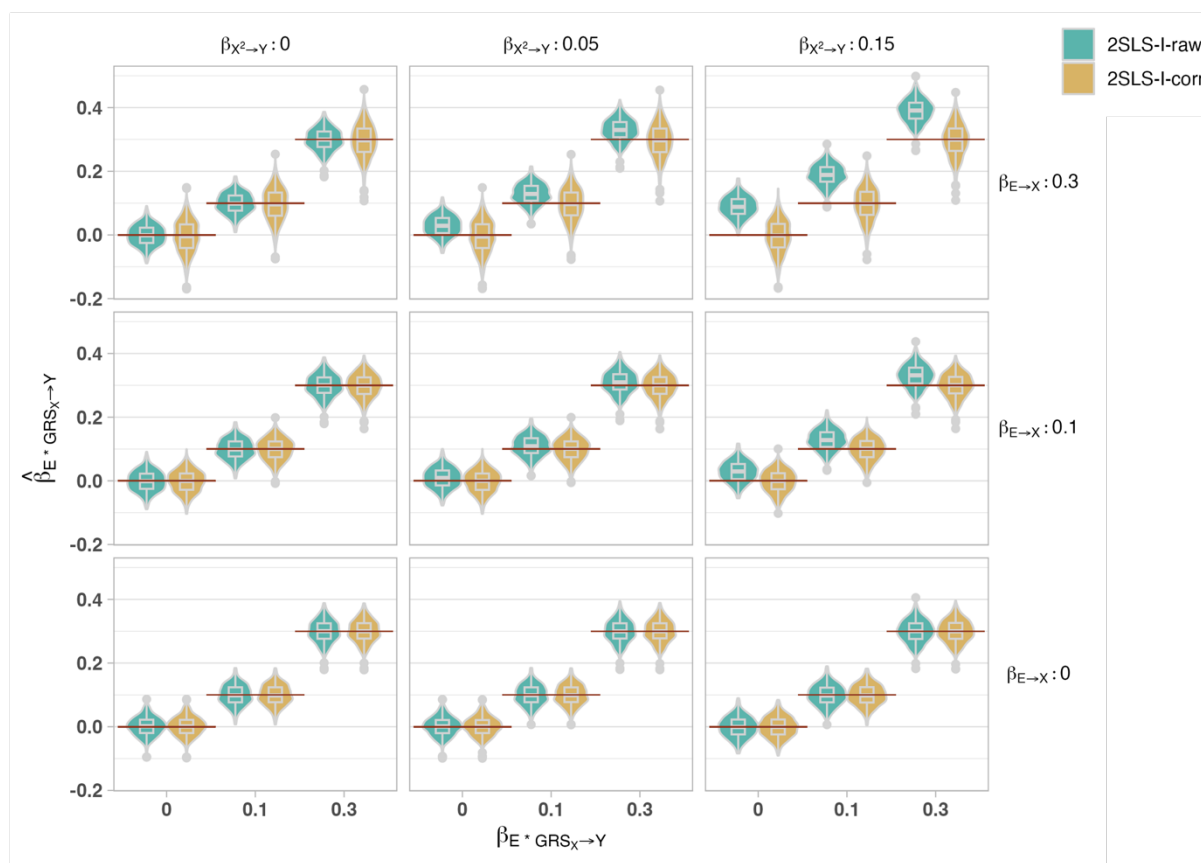


Figure 4. The obtained raw (2SLS-I-raw, blue) and corrected (2SLS-I-corr, yellow) interaction estimates for different levels of  $\beta_{E \rightarrow X}$ ,  $\beta_{X^2 \rightarrow Y}$ , and  $\beta_{E \cdot GRS_X \rightarrow Y}$ . The red lines indicate the different levels of the simulated interaction effect ( $\beta_{E \cdot GRS_X \rightarrow Y}$ ), also indicated on the x-axis. We observed accurate raw interaction estimates as long as either the environment-exposure or the exposure outcome quadratic effect is absent ( $\beta_{E \rightarrow X} = 0$  or  $\beta_{X^2 \rightarrow Y} = 0$ ) (corresponding to plots in the bottom row and left column). Deviation of both  $\beta_{E \rightarrow X}$  and  $\beta_{X^2 \rightarrow Y}$  from 0 led to an overestimation of the raw interaction (in blue, 2SLS-I-raw) between GRS and E on Y (top right). This bias was successfully attenuated in the corrected interaction estimates (in yellow, 2SLS-I-corr). The elements of the boxplot are as follows: center line: median, box limits (lower and upper hinges): quartiles (25<sup>th</sup> and 75<sup>th</sup> percentile), upper whisker: largest value no further than 1.5 times the interquartile range (IQR) from the hinge, lower whisker: smallest value no further than 1.5 times the IQR from the hinge. Data beyond the end of the whiskers are plotted individually as dots.

393 We assessed the accuracy of the interaction estimates in an MR framework using

394 simulations. Data was simulated with a constant effect of the exposure on the outcome ( $\beta_{X \rightarrow Y} =$

395 0.2), whilst the levels of the true interaction between the environment and the exposure ( $\beta_{E \cdot X \rightarrow Y}$ ),  
396 the quadratic effect of the exposure on the outcome ( $\beta_{X^2 \rightarrow Y}$ ), and the effect of the environment on  
397 the exposure ( $\beta_{E \rightarrow X}$ ) were varied systematically (for simulation settings, see *Table 1*).

398 We observed that if both the association between the environmental variable E and the  
399 exposure X ( $\beta_{E \rightarrow X}$ ) and the quadratic effect of the exposure X on Y ( $\beta_{X^2 \rightarrow Y}$ ) deviated from 0, the  
400 obtained raw interactions (2SLS-I-raw) between GRS and E on Y ( $\hat{\beta}_{E \cdot GRS \rightarrow Y}$ ), were  
401 systematically biased. For the most extreme simulation settings ( $\beta_{E \rightarrow X} = 0.3$  and  $\beta_{X^2 \rightarrow Y} = 0.15$ ),  
402  $\hat{\beta}_{E \cdot GRS \rightarrow Y}$  was overestimated by 0.089 (SD = 0.035) on average.

403 In absence of a quadratic effect of X on Y ( $\beta_{X^2 \rightarrow Y} = 0$ ), even a strong association  
404 between the environmental variable E and the exposure X ( $\beta_{E \rightarrow X} = 0.3$ ) did not lead to any bias  
405 using 2SLS-I-raw ( $M_{\text{diff}} = -0.0009$ ,  $SD_{\text{diff}} = 0.033$ ). Vice versa, if  $\beta_{E \rightarrow X}$  was set to 0, the obtained  
406 raw interaction estimates did not deviate from the true interaction, even in presence of a very  
407 strong quadratic effect of the exposure X on Y ( $\beta_{X^2 \rightarrow Y} = 0.15$ ) ( $M_{\text{diff}} = -0.0009$ ,  $SD_{\text{diff}} = 0.034$ )  
408 (*Figure 4*).

#### 409 **Accuracy of 2SLS-I-corr**

410 As described in the section *Correction approach: 2SLS-I-corr*, we investigated the source  
411 of bias and developed a method to correct for the bias observed if both the association between  
412 the environmental variable E and the exposure X ( $\beta_{E \rightarrow X}$ ) and the quadratic effect of the exposure  
413 X on Y ( $\beta_{X^2 \rightarrow Y}$ ) deviated from 0 (2SLS-I-corr). 2SLS-I-corr yielded unbiased interaction  
414 estimates independent of the simulated settings. Most importantly, 2SLS-I-corr allowed for  
415 accurate interaction estimates even in presence of both a strong association between E and X  
416 ( $\beta_{E \rightarrow X} = 0.3$ ) and a strong quadratic effect of X on Y ( $\beta_{X^2 \rightarrow Y} = 0.15$ ) ( $M_{\text{diff}} = -0.003$ ,  $SD_{\text{diff}} =$   
417 0.05). As expected, when both the quadratic effect of X on Y ( $\beta_{X^2 \rightarrow Y}$ ), and the association

418 between the effect modifying environment, E, and the exposure X ( $\beta_{E \rightarrow X}$ ) were set to 0, 2SLS-I-  
419 corr yielded accurate estimates of the interaction ( $M_{\text{diff}} = 0.0001$ ,  $SD_{\text{diff}} = 0.033$ ). In presence of a  
420 strong quadratic effect of X on Y ( $\beta_{X^2 \rightarrow Y} = 0.15$ ), and in absence of an association between E and  
421 X ( $\beta_{E \rightarrow X} = 0$ ), the estimates from 2SLS-I-corr did not deviate from the simulated interaction  
422 ( $M_{\text{diff}} = -0.0008$ ,  $SD_{\text{diff}} = 0.034$ ). It is worth noting that we observed an increase in the standard  
423 error (SE) for 2SLS-I-corr in comparison to the SE of the raw interactions (2SLS-I-raw) ( $M =$   
424  $0.0081$ ,  $SD = 0.0094$ ,  $p < 0.0001$ , 95% CI [0.0080, 0.0083]). We fitted a multiple linear  
425 regression to predict the difference in the SE between 2SLS-I-corr and 2SLS-I-raw based on  
426  $\beta_{E \rightarrow X}$ ,  $\beta_{X^2 \rightarrow Y}$ , and  $\beta_{E \cdot X \rightarrow Y}$ . This resulted in a significant model,  $F(3, 13496) = 87140$ ,  $p < 0.0001$ ,  
427  $R^2 = 0.95$ . The increased SE was positively associated with the strength of the association  
428 between the environmental variable, E, and the exposure X ( $\beta_{E \rightarrow X}$ ,  $\beta = 0.074$ ,  $SE = 0.0001$ ,  $p <$   
429  $0.0001$ ). Furthermore, the difference between the SE of the raw and corrected interaction was  
430 slightly reduced in presence of a stronger interaction between E and X ( $\beta_{E \cdot X \rightarrow Y}$ ,  $\beta = -0.0004$ ,  $SE$   
431  $= 0.0001$ ,  $p = 0.0058$ ). The strength of the quadratic effect of the exposure on the outcome  
432 ( $\beta_{X^2 \rightarrow Y}$ ) did not significantly contribute to the increase in variance observed in the corrected  
433 interaction estimates ( $\beta = 0.00005$ ,  $SE = 0.0003$ ,  $p = 0.87$ ).

434 To assess the accuracy of the estimator of the SE of 2SLS-I-corr, we compared our  
435 analytically derived SE with the empirical SE of the corrected interaction estimates. For  
436 comparison, we did the same for the raw interaction estimates, separately for each simulation  
437 setting. Across all settings, we found that the analytical SE agreed well with the empirical SE for  
438 both, the 2SLS-I-raw and 2SLS-I-corr. The raw model SE did not significantly differ from the  
439 raw empirical SE (mean difference = 0.002, and a similar observation was made for the corrected

440 model SE (mean difference = 0.00098). This confirms that our analytical formula for the  
441 variance of the corrected effect is unbiased.



442 **Power analysis of 2SLS-I**



Figure 5. Analysis of the type I error (TIE) and power of raw (2SLS-I-raw) and corrected (2SLS-I-corr) interaction estimates in absence and presence of bias. The first row illustrates the type I error, as the true interaction between E and X ( $\beta_{E \times X \rightarrow Y}$ ) was set to 0. In absence of any source of bias ( $\beta_{X^2 \rightarrow Y} = 0$ , first and second column), the TIE of both, 2SLS-I-raw and 2SLS-I-corr are close to 5%. In presence of a quadratic effect of X on Y ( $\beta_{X^2 \rightarrow Y} = 0.1$ , third and fourth column), the raw interaction estimates are nominally significant in up to 100% of the cases in absence of a true interaction. This effect is successfully attenuated in the corrected interaction estimates, where the TIE does not exceed 5%. In presence of a weak interaction (middle row), 2SLS-I-corr showed reduced power in comparison to 2SLS-I-raw in absence of bias. C: Lastly, the power of 2SLS-I-corr reached a ceiling effect in presence of a strong interaction ( $\beta_{E \times X \rightarrow Y} = 0.1$ ) with a sample size  $\geq 50,000$  and  $R^2_{GRS \rightarrow X} \geq 0.1$  (bottom column).

443 To investigate the impact of our correction approach in hypothesis testing, we compared  
 444 the type I error (TIE) and power of the 2SLS-I-corr with that of the 2SLS-I-raw. In these  
 445 simulations, the association between the environmental parameter E and the exposure X ( $\beta_{E \rightarrow X}$ )  
 446 was set to 0.2, whilst the true interaction ( $\beta_{E \times X \rightarrow Y}$ ), the quadratic effect of the exposure X on the  
 447 outcome Y ( $\beta_{X^2 \rightarrow Y}$ ), the amount of variance of X that is explained by the GRS ( $R^2_{GRS \rightarrow X}$ ), and the  
 448 sample size were varied systematically (for the simulation settings, see Table 3).

449 In absence of a true interaction ( $\beta_{E \cdot X \rightarrow Y} = 0$ ) and without the source of bias ( $\beta_{X^2 \rightarrow Y} = 0$ ),  
450 the (5% nominal level) T1E of 2SLS-I-corr (0.039 (SD = 0.01)) was comparable with the T1E of  
451 2SLS-I-raw (0.046 (SD = 0.0089)). In presence of the source of bias ( $\beta_{X^2 \rightarrow Y} = 0.1$ ), we observed  
452 a T1E of up to 1 for 2SLS-I-raw (on average 0.66 (SD = 0.33)). Importantly, even with a small  
453 sample size of 10.000 and a  $R^2_{GRS \rightarrow X}$  of 0.05, the T1E was 0.14 for the raw interaction in presence  
454 of bias. 2SLS-I-corr attenuated the T1E to a maximum of 0.05 (average of 0.035 across settings  
455 (SD = 0.0089)).

456 In presence of a weak interaction ( $\beta_{E \cdot X \rightarrow Y} = 0.05$ ) and in absence of bias, 2SLS-I-raw had  
457 a power of more than 0.7 in all settings where the sample size was larger than 10.000 (M = 0.75,  
458 SD = 0.3). In contrast, 2SLS-I-corr only exceeded the power of 0.7 if both the sample size was  
459 larger than 10.000 and the amount of variance in X explained by the GRS exceeded 0.05 (M =  
460 0.599, SD = 0.35). This reduction in power of the 2SLS-I-corr in comparison to the 2SLS-I-raw  
461 results from the increased variance in response to our additive correction in comparison to the  
462 raw interaction estimates.

463 In presence of a strong interaction, the power of 2SLS-I-raw exceeded 0.8 in all settings  
464 except if the  $R^2_{GRS \rightarrow X}$  and the sample size were set to 0.05 and 10.000, respectively (M = 0.94,  
465 SD = 0.14). For 2SLS-I-corr, the power exceeded 0.8 in all settings except if the  $R^2_{GRS \rightarrow X}$  was  
466 smaller than 0.2 and the sample size was limited to 10.000 (M = 0.86, SD = 0.25).

467 In summary, 2SLS-I-corr has a much better controlled type 1 error than the 2SLS-I-raw  
468 whilst maintaining considerable power to detect true interactions if  $R^2_{GRS \rightarrow X}$  exceeded 0.05 and  
469 for sample size > 10.000 participants.

470

## 471 **Application**

472 To elaborate the relevance of environmental moderators of causal effects, we investigated  
473 how a range of not genetically instrumentable environmental parameters (age, air pollution  
474 (NO<sub>2</sub>), sedentary behaviour, socioeconomic deprivation (Townsend Deprivation Index, TDI),  
475 smoking, and time spent watching TV (TV time)) modulate causal relationships between  
476 different health- and lifestyle- parameters. As simulations (see section *Biased interaction*  
477 *estimates in presence of quadratic effect of X on Y and effect of E on X*) indicated that non-linear  
478 effects of the exposure on the outcome may lead to spurious interaction effects, corrected  
479 interaction estimates (2SLS-I-corr) were considered in presence of a nominally significant  
480 quadratic effect of the exposure (215 out of 1274, 16.9%). For all other settings, raw interaction  
481 estimates (2SLS-I-raw) were considered, in order to maximise discovery power. To account for  
482 the limited power to detect true interactions, we replicated our analyses using an extended  
483 genetic risk score (GRS<sub>ext</sub>) in addition to the strict genetic risk score (where the GRS<sub>ext</sub> is based  
484 on all SNPs that contribute to an exposure whilst the strict GRS is only based on independent  
485 SNPs that are genome wide significantly associated to an exposure). Evidence for interactions  
486 were classified into tiers from 0 to 4 to account for different levels of confidence (see section  
487 *Interaction tier score*), considering significance (of both, the GRS and the GRS<sub>ext</sub>), robustness of  
488 the interaction effect, and presence of a main effect.

489 Of the effect modifying environmental parameters, age, TDI, and smoking modulated the  
490 most causal relationships ( $n_{\text{age}} = 60$  and  $n_{\text{TDI}} = 36$ ,  $n_{\text{smoking}} = 22$ , respectively). In particular, we  
491 found a reduction in strength for a range of causal effects with increasing age. For example, the  
492 causal effect of height on hand grip strength (HGS) ( $\beta = 0.22$ ,  $\text{SE} = 0.0037$ ,  $p < 10^{-323}$ ), was

493 significantly attenuated with higher age (Tier score = 4,  $\beta_{\text{GRS}^*E} = -0.019$ ,  $\text{SE}_{\text{GRS}^*E} = 0.0037$ ,  
494  $p_{\text{val}_{\text{GRS}^*E}} < 10^{-6}$ , *Figure 8a*).

495 For socioeconomic deprivation, we found that most interactions were associated with an  
496 exacerbation of the causal effects. For example, the negative effect of systolic blood pressure on  
497 forced expiratory volume within the first second (FEV1) ( $\beta = -0.099$ ,  $\text{SE} = 0.013$ ,  $p < 10^{-13}$ ), was  
498 intensified for people exposed to higher levels of socioeconomic deprivation (Tier score = 3,  
499  $\beta_{\text{GRS}^*E} = -0.021$ ,  $\text{SE}_{\text{GRS}^*E} = 0.013$ ,  $p_{\text{GRS}^*E} = 0.105$ ,  $\beta_{\text{GRS}_{\text{ext}}^*E} = -0.022$ ,  $\text{SE}_{\text{GRS}_{\text{ext}}^*E} = 0.0041$ ,  
500  $p_{\text{GRS}_{\text{ext}}^*E} < 10^{-7}$ , *Figure 8b*).

501 Furthermore, we found evidence that smoking modulates 22 causal effects. For example,  
502 the positive causal effect of Gamma-glutamyl transferase (Gamma GLT) on CRP ( $\beta = -0.107$ ,  $\text{SE}$   
503  $= 0.006$ ,  $p < 10^{-76}$ ) seems to be intensified in smokers (Tier score = 4,  $\beta_{\text{GRS}^*E} = 0.019$ ,  $\text{SE}_{\text{GRS}^*E} =$   
504  $0.006$ ,  $p_{\text{GRS}^*E} < 0.001$ ,  $\beta_{\text{GRS}_{\text{ext}}^*E} = 0.018$ ,  $\text{SE}_{\text{GRS}_{\text{ext}}^*E} = 0.004$ ,  $p_{\text{GRS}_{\text{ext}}^*E} < 10^{-5}$ ). Yet, it is worth  
505 noting that the interaction between smoking and Gamma GLT is scale dependent (i.e. the effect  
506 may be driven by individuals with very high CRP levels), as it yielded a tier score of 0 if the  
507 outcome was inverse-rank normal transformed.

508 Out of 19 causal effects that were modulated by air pollution ( $\text{NO}_2$ ), 8 affected the  
509 outcome CRP and 5 the outcome Forced Expiratory Volume (FEV1). For example, we observed  
510 that the effect of Education on FEV1 ( $\beta = 0.208$ ,  $\text{SE} = 0.016$ ,  $p < 10^{-38}$ ) was significantly  
511 intensified for people who live in areas with higher air pollution (Tier score = 4,  $\beta_{\text{GRS}^*E} = 0.060$ ,  
512  $\text{SE}_{\text{GRS}^*E} = 0.016$ ,  $p_{\text{GRS}^*E} = 0.0002$ , *Figure 8c*).

513 Sedentary behaviour, defined as time spent inactively, measured using an accelerometer  
514 device, mostly intensified causal effects on CRP (7 out of 8 interactions with a tier score  $> 0$ ).  
515 For three of these interactions, we found evidence that physical inactivity may act as a collider

516 between the exposures (BFP, education, and water mass) and the outcome (CRP), which likely  
517 leads to a bias in the interaction estimates. In addition, we observed that the effect of grip  
518 strength, cystatin C, Na in Urine, and SHB on CRP are significantly intensified in people who  
519 spend more time physically inactive. Finally, there was one setting where sedentary behaviour  
520 modified an effect on an outcome that is not CRP, namely the interaction with HbA1c on reaction  
521 time (tier = 1).

522 For time spent watching TV (TV time), we observed that the causal effect of Cystatin C  
523 on CRP ( $\beta = 0.032$ ,  $SE = 0.0061$ ,  $p < 10^{-7}$ ) is intensified in people who spend more time  
524 watching TV (Tier score = 3,  $\beta_{GRS*E} = 0.017$ ,  $SE_{GRS*E} = 0.0061$ ,  $p_{GRS*E} = 0.004$ ,  $\beta_{GRSext*E} = 0.02$ ,  
525  $SE_{GRSext*E} = 0.0034$ ,  $p_{GRSext*E} < 10^{-8}$ , *Figure 8d*). Noteworthy, the interaction between TV time  
526 and Cystatin C on CRP seems to be scale dependent (i.e. to some extent driven by individuals  
527 with very high levels of CRP), as the interaction on the IRN-transformed outcome phenotype  
528 yielded a tier score of 0.

529 For detailed results, see *Figure 6*, *Figure 7*, *Figure 8*, *Table 4*, and *Supplementary Table*  
530 *4*.

| X               | E       | Y    | GRS*E → Y |       |          | GRSext*E → Y |       |           | X → Y  |       |          | Tier<br>IRNT | Dir<br>IRNT |
|-----------------|---------|------|-----------|-------|----------|--------------|-------|-----------|--------|-------|----------|--------------|-------------|
|                 |         |      | β         | SE    | pval     | β            | SE    | pval      | β      | SE    | pval     |              |             |
| Body Fat %      | age     | LDL  | -0.086    | 0.008 | 9.14e-26 | -0.097       | 0.003 | 3.87e-191 | -0.067 | 0.008 | 3.05e-16 | 4            | agrees      |
| Body Fat %      | age     | SBP  | -0.036    | 0.008 | 1.25e-05 | -0.039       | 0.003 | 2.05e-32  | 0.124  | 0.008 | 1.67e-51 | 4            | agrees      |
| Water mass      | age     | FEV1 | -0.061    | 0.006 | 4.11e-26 | -0.040       | 0.003 | 1.03e-39  | 0.354  | 0.006 | <1e-323  | 4            | agrees      |
| Water mass      | age     | SBP  | -0.029    | 0.006 | 5.42e-07 | -0.039       | 0.003 | 7.77e-37  | -0.044 | 0.006 | 8.70e-15 | 4            | agrees      |
| Water mass      | age     | HGS  | -0.044    | 0.006 | 1.92e-15 | -0.039       | 0.003 | 3.34e-40  | 0.299  | 0.006 | <1e-323  | 4            | agrees      |
| ALT             | age     | LDL  | -0.061    | 0.008 | 1.62e-13 | -0.070       | 0.004 | 1.57e-57  | 0.077  | 0.008 | 3.01e-23 | 4            | agrees      |
| SHBG            | age     | FEV1 | -0.020    | 0.006 | 0.0008   | -0.013       | 0.004 | 0.0004    | 0.045  | 0.006 | 1.11e-14 | 3            | agrees      |
| Vit D           | age     | LDL  | 0.040     | 0.011 | 0.0004   | 0.055        | 0.006 | 3.63e-19  | -0.130 | 0.010 | 2.26e-40 | 4            | agrees      |
| Grip Strength   | age     | FEV1 | -0.049    | 0.014 | 0.0005   | -0.039       | 0.004 | 4.57e-22  | 0.564  | 0.014 | <1e-323  | 3            | agrees      |
| Height          | age     | FEV1 | -0.040    | 0.004 | 2.99e-22 | -0.040       | 0.003 | 3.25e-53  | 0.356  | 0.004 | <1e-323  | 4            | agrees      |
| Height          | age     | LDL  | 0.020     | 0.004 | 1.04e-07 | 0.026        | 0.003 | 2.06e-24  | -0.054 | 0.004 | 3.50e-46 | 4            | agrees      |
| Height          | age     | HGS  | -0.019    | 0.004 | 1.52e-07 | -0.013       | 0.002 | 1.89e-07  | 0.221  | 0.004 | <1e-323  | 4            | agrees      |
| Body Fat %      | AP: NO2 | FEV1 | -0.045    | 0.008 | 9.71e-08 | -0.017       | 0.003 | 6.77e-07  | -0.171 | 0.008 | 2.01e-92 | 4            | agrees      |
| Height          | AP: NO2 | CRP  | -0.013    | 0.004 | 0.0006   | -0.009       | 0.003 | 0.0007    | -0.034 | 0.004 | 3.65e-19 | 3            | agrees      |
| Educational (y) | AP: NO2 | FEV1 | 0.060     | 0.016 | 0.0002   | 0.043        | 0.006 | 1.10e-12  | 0.208  | 0.016 | 7.17e-39 | 4            | agrees      |
| FEV1            | TDI     | FIS  | 0.091     | 0.017 | 1.62e-07 | 0.042        | 0.007 | 2.21e-10  | 0.081  | 0.015 | 3.96e-08 | 3            | agrees      |
| FEV1            | TDI     | HGS  | 0.042     | 0.009 | 7.36e-06 | 0.017        | 0.004 | 1.95e-06  | 0.261  | 0.008 | 9.8e-212 | 3            | agrees      |
| FEV1            | smoking | FIS  | 0.073     | 0.018 | 3.56e-05 | 0.024        | 0.007 | 0.0004    | 0.081  | 0.015 | 3.96e-08 | 3            | agrees      |
| Gamma GLT       | smoking | CRP  | 0.019     | 0.006 | 0.0009   | 0.018        | 0.004 | 7.13e-06  | 0.107  | 0.006 | 4.44e-76 | 0            | agrees      |
| BR              | smoking | LDL  | -0.014    | 0.004 | 0.0009   | -0.010       | 0.003 | 0.0008    | -0.030 | 0.004 | 1.65e-12 | 4            | agrees      |
| Chol            | TV time | CRP  | -0.020    | 0.006 | 0.0004   | -0.023       | 0.004 | 4.58e-08  | -0.036 | 0.006 | 4.70e-10 | 0            | agrees      |

Table 4 Overview of the tier 4 interaction effects with no evidence for a strong exposure-to-environment effect. Tier IRNT indicates the tier score obtained when the outcome was inverse rank normal transformed. Dir IRNT agrees in all settings, meaning that both, the strict and the extended GRS yielded an interaction effect pointing in the same direction as the interaction effect obtained for the z-standardized outcome, when the outcome was inverse rank normal transformed. Yet, there are two settings where the strength of the effect was potentially driven by individuals on the extreme end of the scale for CRP, as these interactions were attenuated to a tier score of 0 when interaction analysis was performed using the IRN-transformed outcome phenotype. AP: NO2: Air Pollution, Nitrogen Dioxide, CRP: C-reactive protein, FIS: Fluid intelligence score, HGS: Grip strength, FEV1: forced expiratory volume within one second, LDL: low density lipoprotein, SBP: systolic blood pressure, TDI = Townsend deprivation index.

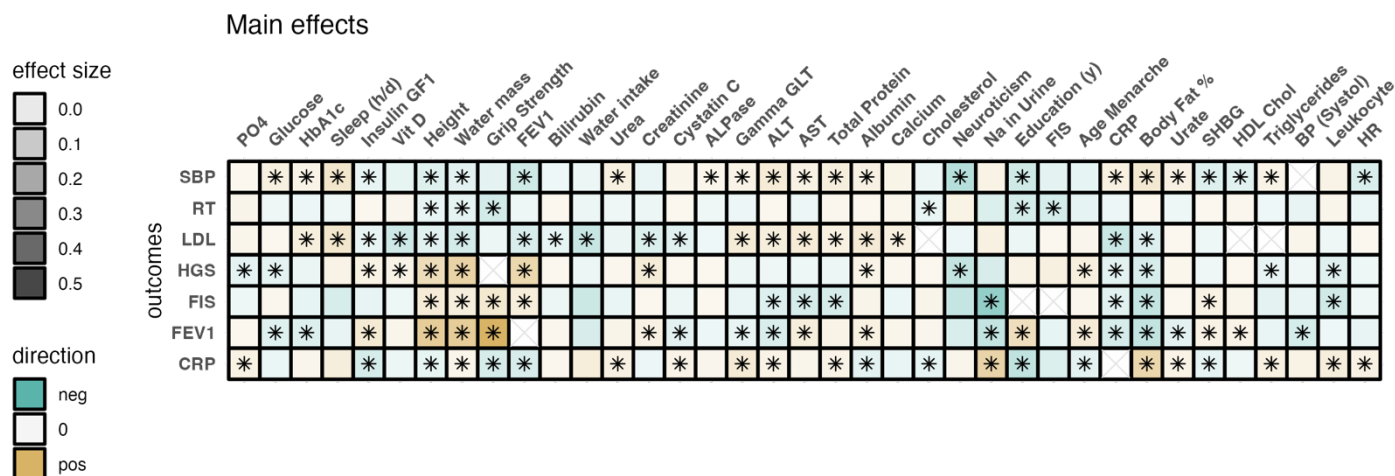
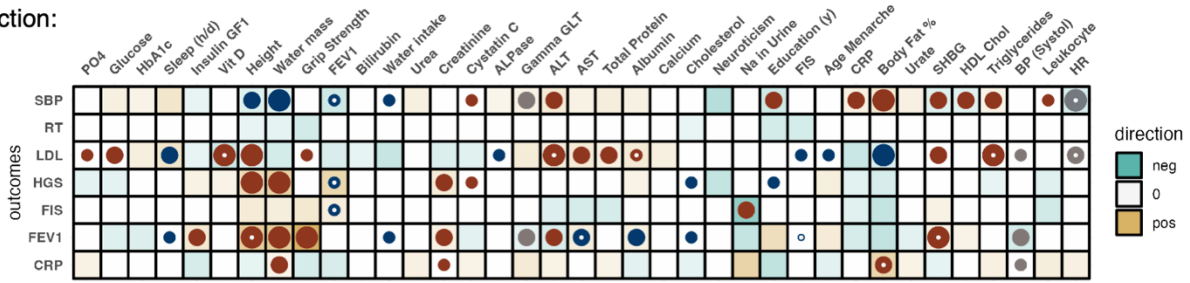
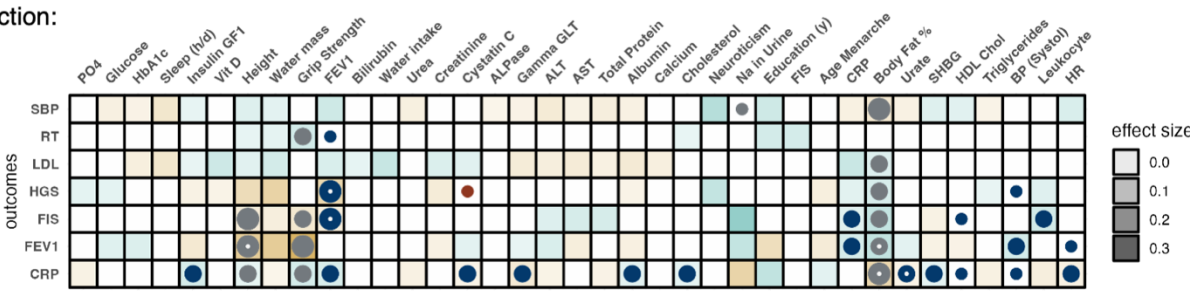


Figure 6. Causal effects of a range of exposures (x-axis) on a range of outcomes (y-axis), obtained from the minimal MR model. The color of the tiles indicates the direction of the main effect, whilst the opacity indicates the strength of the main effect. Bonferroni corrected significant effects are marked with a star. ALT: Alanine Aminotransferase, AST: Aspartate Aminotransferase, BP: Blood pressure, CRP: C-reactive Protein, FEV1: Forced Expiratory Volume within 1 second, FIS: Fluid Intelligence Score, HbA1c: Glycated hemoglobin, HDL: High density lipoprotein, HR: Heart rate, NO<sub>2</sub>: Nitrogen dioxide, SHBG: sex hormone binding globulin, PO4: phosphate, Vit D: Vitamin D

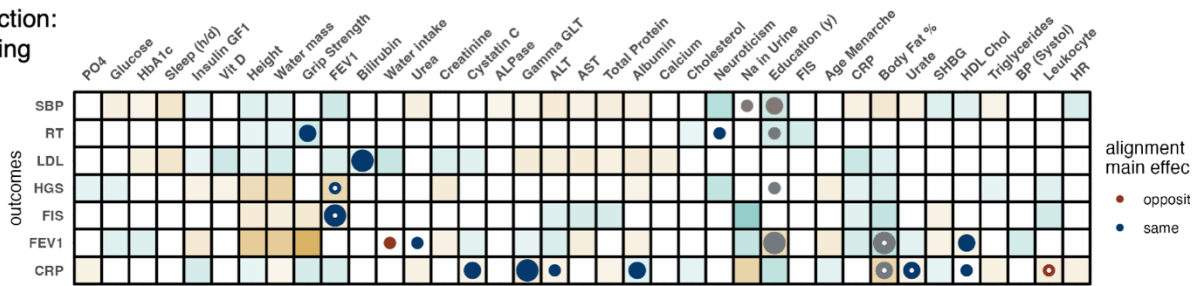
**Interaction:  
Age**



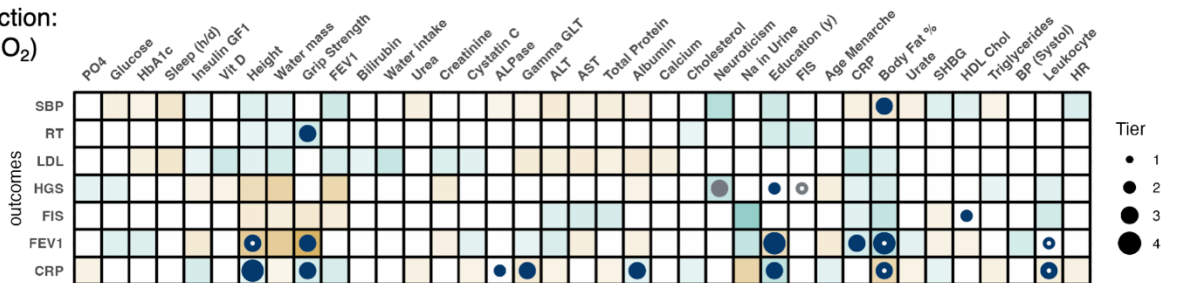
**Interaction:  
TDI**



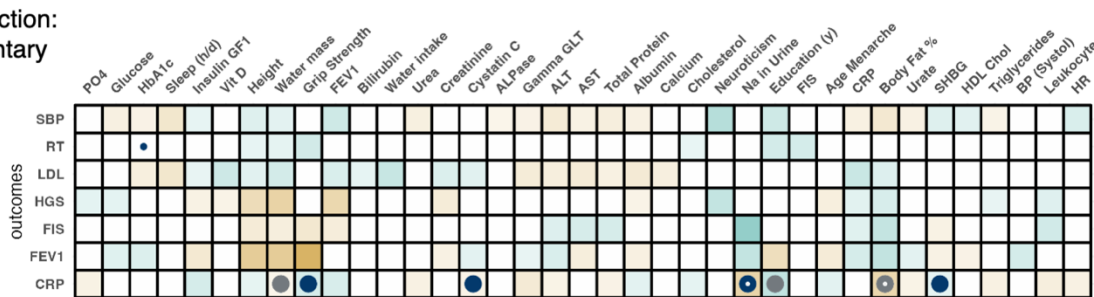
**Interaction:  
Smoking**



**Interaction:  
AP (NO<sub>2</sub>)**



**Interaction:  
Sedentary**



**Interaction:  
TV time**

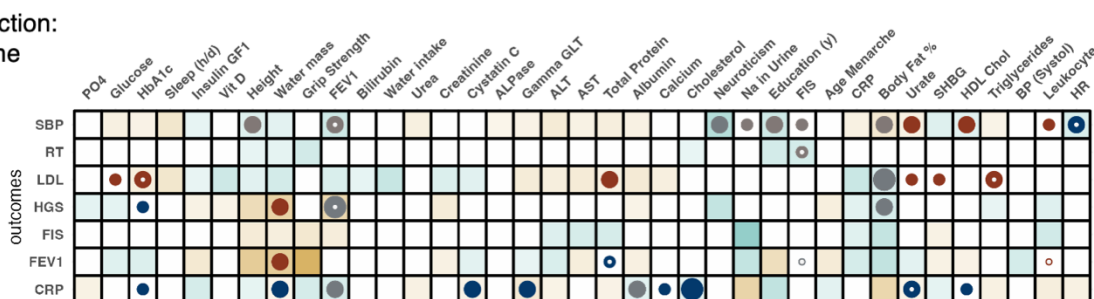




Figure 7. Results of the applied interaction analysis. Panels indicate interactions with each a different interaction parameter (age, socioeconomic deprivation (Townsend Deprivation Index, TDI), smoking, air pollution (NO<sub>2</sub>), sedentary behavior, and time spent watching TV (TV time). The dots indicate evidence for an interaction, with the size of the dots representing the tier score (1 = very little confidence, 2 = little confidence, 3 = some confidence, 4 = high confidence). The color of the dots indicates whether the interaction effect agrees with the direction of the main effect (dark blue) or not (red), i.e. whether the effect of the exposure on the outcome increases with higher levels of the environment (dark blue). The small white dots indicate when the corrected interaction estimate was considered due to the presence of a non-linear exposure-outcome relationship. Dots were shaded out if there was evidence for a strong causal effect of the exposure on the environment, as this may lead to biases in the interaction estimates. ALT: Alanine Aminotransferase, AST: Aspartate Aminotransferase, BP: Blood pressure, CRP: C-reactive Protein, FEV1: Forced Expiratory Volume within 1 second, FIS: Fluid Intelligence Score, HbA1c: Glycated hemoglobin, HDL: High density lipoprotein, HR: Heart rate, NO<sub>2</sub>: Nitrogen dioxide, SHBG: sex hormone binding globulin, PO4: phosphate, Vit D: Vitamin D

534

535

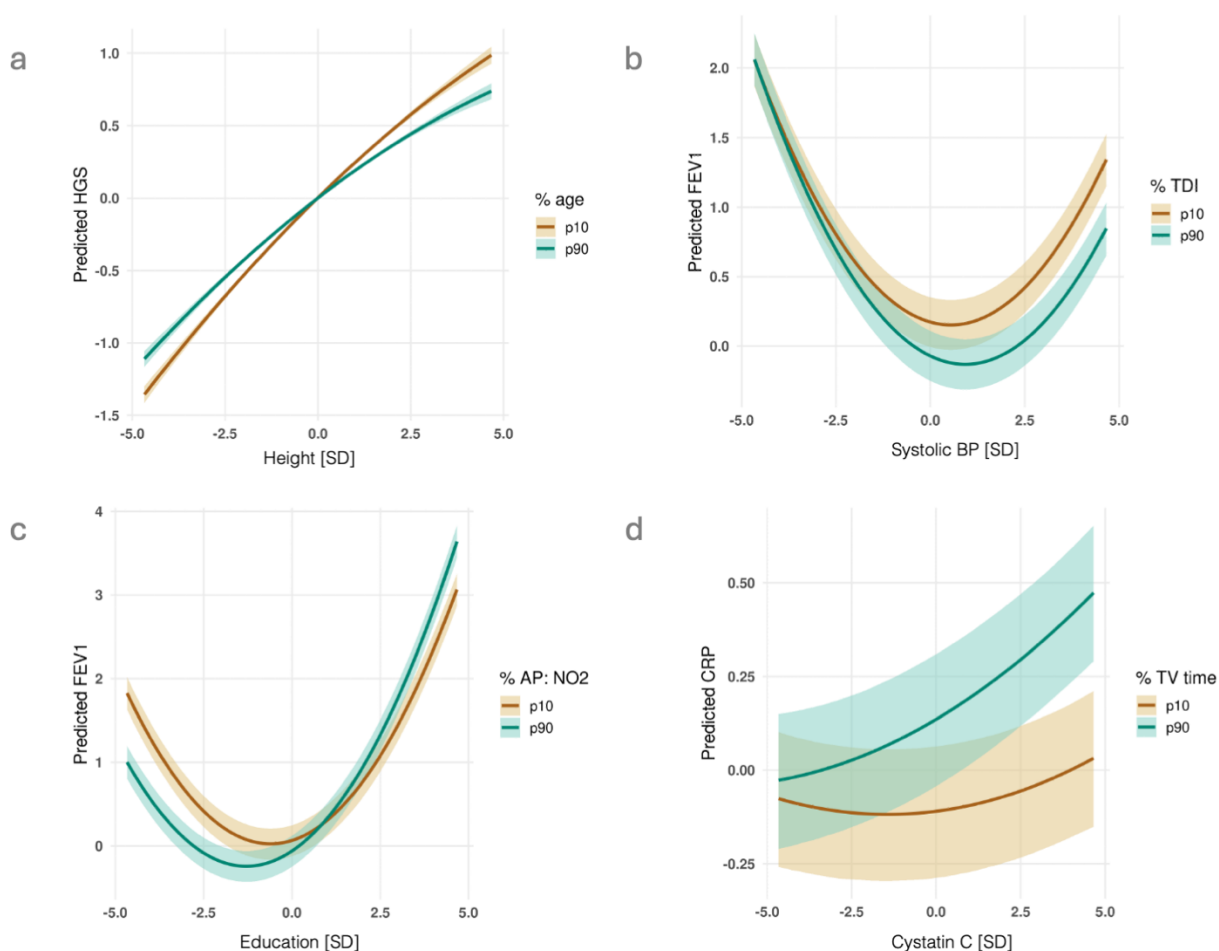


Figure 8. Plots of single interactions obtained using the full MR model. Selection of raw interactions obtained using the strict GRS. Units translate to the standard deviation (SD), and the confidence hulls indicate the 95% confidence interval. FEV1: Forced expiratory volume within 1 s, BP: blood pressure, p10: 10<sup>th</sup> percentile, p90: 90<sup>st</sup> percentile, TDI: Townsend Deprivation Index, AP: Air pollution, CRP: C-reactive protein.

536

## Discussion

537 We present 2SLS-I, an approach to investigate how environmental variables modulate  
538 causal effects. Extensive simulations revealed accurate interaction estimates for almost all  
539 settings. Yet, we observed biased interaction estimates in presence of a non-linear exposure-  
540 outcome relationship and dependent environment and exposure, for which we provide a method  
541 to correct if indicated. Application of 2SLS-I to a range of health-related exposure- and outcome  
542 phenotypes revealed that all considered environments (age, socioeconomic deprivation  
543 (Townsend Deprivation Index, TDI), smoking, air pollution (nitrogen dioxide, NO<sub>2</sub>), physical  
544 inactivity (sedentary behaviour, PA: sed) and time spent watching TV (TV time)) modulate some  
545 causal relationships. We found that the strength of causal relationships tends to be attenuated in  
546 higher age. For example, our results indicate that age modulates the well-known relationship  
547 between height and hand grip strength (HGS). Whilst it is known that not only height (Abaraogu  
548 et al., 2017), but also age (e.g. Frederiksen et al., 2006) predicts HGS, it seems that the effects of  
549 age and height on HGS are not only additive, but in fact multiplicative. The tendency for causal  
550 effects to be attenuated in older individuals may be the result of the accumulation of other health-  
551 related factors which become more relevant the older a person grows.

552 Interestingly, we observed the opposite pattern for socioeconomic deprivation (Townsend  
553 Deprivation Index, TDI), where (mostly detrimental) causal effects seem to be intensified for  
554 people living in more deprived areas. For example, we found that the detrimental effect of  
555 systolic blood pressure (SBP) on forced expiratory volume (FEV1) was significantly intensified  
556 for people with lower socioeconomic status. Whilst there is some evidence for an association  
557 between hypertension and FEV1 (Miele et al., 2018), others argue that the effect is reverse, in  
558 fact, with lower FEV1 increasing blood pressure (Engström et al., 2001). Furthermore, there is

559 also evidence that the negative association between blood pressure and FEV1 may result from  
560 confounding due to antihypertensive medication (Schnabel et al., 2011). Yet, causal inference  
561 methods point toward a negative effect of higher systolic blood pressure on FEV1 (Zekavat et al.,  
562 2021). Extending findings from Wheeler and colleagues (2005), who reported a positive  
563 association between socioeconomic status and FEV1, we provide evidence that the causal effect  
564 of SBP on FEV1 is exacerbated in response to socioeconomic deprivation.

565 For smoking, we observed that numerous causal effects were exacerbated in smokers in  
566 comparison to non-smokers. For example, we found that the causal effect of Gamma GLT on  
567 CRP is more pronounced in smokers than in non-smokers. Whilst it has long been known that  
568 Gamma GLT has a positive (i.e. increasing) effect on CRP (Lee et al., 2003), and studies found  
569 that smoking increases both, levels of Gamma GLT (Zhang et al., 2021) and CRP (O'Loughlin et  
570 al., 2008; Tracy et al., 1997), we provide evidence that the relationship between Gamma GLT  
571 and CRP is exacerbated in smokers in comparison to non-smokers.

572 Meanwhile, air pollution (NO<sub>2</sub>) mostly modulated effects on CRP and FEV1, whereby it  
573 exclusively intensified the causal effects. For example, the effect of education on FEV1 was  
574 intensified in response to living in an area with higher air pollution. Tabak and colleagues (2009)  
575 report a lower smoking adjusted FEV1 at baseline in people with a low educational level in  
576 comparison to people with a high educational level. Furthermore, it has long been known that  
577 short- and long-term exposure to air pollution (and NO<sub>2</sub> in particular) is negatively associated  
578 with lung function (both FVC and FEV1) (Ackermann-Lieblich et al., 1997; Strassmann et al.,  
579 2021). Interestingly, multiple studies show some modulating effect of (parental) socioeconomic  
580 status on the association between air pollution and lung capacity (Cakmak et al., 2016; Wheeler,  
581 2005). Yet, the effects seem to differ depending on participants' sex and the definition (and

582 potentially stratification) of the SES variable(s). A systematic comparison between varying  
583 definitions of education, potentially stratifying by sex, may contribute to a better understanding  
584 of the mechanisms through which education is protective against the detrimental effects of air  
585 pollution on lung function.

586         The effect modifying role of physical activity could only be assessed to a limited extent.  
587 As accelerometer data was only available in 67912 individuals after quality control, and physical  
588 inactivity seems to act as a collider in multiple relevant settings (e.g. the effect of body fat  
589 percentage on CRP), we only detected interactions with relatively low confidence (maximum tier  
590 score = 3). Importantly, physical inactivity mostly modulated effects on the outcome CRP (7 out  
591 of 8 interactions with a tier score > 0), of which all indicated that physical inactivity intensifies  
592 the (univariable) main effects. Given the well-known anti-inflammatory effect of physical  
593 activity, it is likely that physical inactivity exacerbates detrimental causal effects on CRP, which  
594 aligns with our observations. Nevertheless, it would be of great interest to investigate the effect  
595 modifying role of physical (in)activity more thoroughly in a larger sample.

596         The moderating effect of TV time turns out to be difficult to interpret. For example, we  
597 found that the positive effect of Cystatin C on CRP is significantly intensified in people who  
598 spend more time watching TV. Whilst there is evidence for an association between Cystatin C  
599 and CRP (Shlipak et al., 2005), it is worth noting that higher levels in Cystatin C have been  
600 found to be associated with higher age, higher triglycerides, lower HDL-Cholesterol and a range  
601 of inflammatory markers other than CRP, such as TNF $\alpha$  and Interleukin-6 (Luc et al., 2006).  
602 Meanwhile, TV time has been found to be positively associated with CRP even in children  
603 (Gabel et al., 2016) and after accounting for a range of relevant covariates such as waist  
604 circumference, physical activity, and dietary density (Gabel et al., 2016). As TV time is likely a

605 function of many factors, such as age, general health, and socioeconomic status, and cannot be  
606 instrumented genetically, it remains difficult to investigate to what extent an intervention on TV  
607 time would reduce the health consequences associated with it.

608 Collectively, these results illustrate the relevance of environmental parameters as  
609 modulators of causal effects. Nevertheless, it is important to keep in mind limitations of the  
610 current study.

611 Firstly, whilst we tested our methods in a wide range of settings and simulated many  
612 (potential) sources of bias, from non-linear causal effects to linear and non-linear confounding, to  
613 the presence of level 1 interactions, simulations are typically incapable of accounting for the  
614 complexities of the real world. Whilst we are confident that 2SLS-I yields accurate effects in a  
615 wide range of settings, it is likely that there are specific circumstances where 2SLS-I fails to  
616 provide accurate findings.

617 Second, the detection of interactions has low statistical power. To address this, we  
618 replicated our analysis using an extended GRS ( $GRS_{ext}$ ), which may violate MR assumptions  
619 (the weaker SNPs correlate with the exposure, the more likely that they act indirectly). To  
620 account for this circumstance, we ensured that the results from the  $GRS_{ext}$  only contributed to an  
621 increased tier score if the effect estimate agreed with the respective estimate of the strict GRS.  
622 Furthermore, we considered interactions as “significant” if they reached a p-value  $< 0.001$ , which  
623 allowed us to detect more potential interactions, but also increases the false discovery rate.  
624 Another way to increase power is by carefully choosing when to apply the corrected interaction  
625 estimates. Since the 2SLS-I-corr effects have larger variance than the uncorrected counterpart,  
626 they should primarily be used when there is evidence for a non-linear causal effect. The  
627 replication of the present findings in an additional cohort would be necessary to confidently

628 consider them as valid, in particular those with a tier score  $< 4$ . Such interactions, however, may  
629 be population-specific, which renders replication particularly difficult. Furthermore, despite  
630 making a considerable effort to account for the reduced power inherent to interaction analysis  
631 (and Mendelian Randomization), the absence of a significant interaction may still be the result of  
632 lacking power and does not translate to a proof of inexistence, in particular for phenotypes with  
633 relatively small sample size and high variance of noise, such as physical inactivity. Yet, thanks to  
634 the ever-increasing sample size of available biobanks, we are confident that the reduced power of  
635 interaction analyses will become a smaller challenge in the future. If this is the case, 2SLS-I  
636 would even allow for the investigation of environment-dependent non-linear effects (e.g.  $X^2 \cdot E$ ),  
637 which may be of relevance as we observed multiple settings with evidence for non-linear causal  
638 effects. Still, it has to be noted that we modelled only quadratic X-to-Y effects and our proposed  
639 correction is suboptimal for more complex non-linear X-Y relationships.

640 Third, due to the far-reaching effects of many environmental parameters, some  
641 interactions are difficult to interpret. A detected exposure-environment interaction may arise only  
642 due to a true interaction between the exposure and another (correlated) environment. Whilst  
643 understanding that an environmental variable modulates a causal effect is of great relevance, we  
644 note that the present analysis could (and should) be extended for almost all detected interactions  
645 to obtain a detailed understanding of the complex interplay between variables, in particular for  
646 those related to SES.

647 Fourth, detected interactions may be scale specific and not persist when modelling a  
648 transformed version of the outcome (e.g.  $\log(Y)$ ). This is a general weakness of all types of  
649 interaction analyses. Nevertheless, we replicated our analysis by inverse-rank normal  
650 transforming the outcome phenotypes and found that most interactions yield comparable effects

651 independent of the outcome standardization. Yet, in particular for CRP we observed multiple  
652 settings where the effect seems to depend on the standardization of the outcome phenotype due  
653 to the strong right-skewedness of CRP.

654 Fifth, one has to carefully exclude the possibility that a tested environment is a collider of  
655 the exposure-outcome relationship, because it can lead to biased interaction estimates. If such  
656 situation occurs, it is safer to regress out the exposure from the environmental variable and  
657 consider the residual trait as the new tested environment. Similarly, analyses should assess the  
658 presence and relevance of  $G \cdot E$  effects on X as they violate our assumption that the genetic  
659 effects on the exposure are environment-invariant.

660 Sixth, the full overlap between samples from which the SNP-exposure and SNP-outcome  
661 effects are estimated can introduce bias in the interaction estimates, which requires further work  
662 to account for. However, based on the impact of sample overlap on the causal main effect,  
663 (Mounier & Kutalik, 2023), it is likely to play a minor role for interactions too.

664 Furthermore, we did not aim at validating the non-linear effects observed using 2SLS-I.  
665 Although we obtained accurate non-linear estimates for all simulated settings, future projects  
666 could aim at extending the simulation settings to challenge the accuracy of 2SLS-I for non-linear  
667 effects or perform negative control experiments to investigate if the non-linear effects of 2SLS-I  
668 are robust across a wide range of settings.

669 Finally, the present analysis was restricted to individuals of White British ancestry. As  
670 many environmental parameters, such as SES, vary between different ancestries, it would be of  
671 great relevance to extend the present analyses to diverse ancestral groups.

672 In conclusion, we present 2SLS-I, a method to investigate how environmental variables  
673 modulate causal effects, even in presence of sources of bias, such as non-linear effects of the

674 exposure on the outcome. We demonstrate the reliability of 2SLS-I in a wide range of simulation  
675 settings. Finally, we provide evidence that it is relevant to consider the modulating effects of  
676 environmental variables such as age and SES when examining the causal effect of classical  
677 epidemiological risk factors, which is a step towards precision medicine.  
678



679 **Data availability**

680 The data that support the findings of this study are available from the UK Biobank  
681 (UKBB), but restrictions apply to the availability of these data, which were used under license  
682 for the current study. Access to the UKBB can be requested through a standard protocol  
683 (<https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>). Summary statistics were  
684 accessed from Neale and colleagues (2017), which are publicly available  
685 (<http://www.nealelab.is/uk-biobank>), just as the weights to calculate the extended genetic risk  
686 score as in Privé et al. (2022) (<https://www.pgscatalog.org/publication/PGP000263/>).

687 **Code availability**

688 Scripts used to perform the analyses are available at <https://github.com/leonakn/2SLS-I>.

689 **Additional information**

690 **Ethical approval**

691 The UK Biobank has approval from the North West Multi-centre Research Ethics  
692 Committee (MREC). The data was accessed through application number 16389.

693

694

## References

695 Abaraogu, U. O., Ezema, C. I., Ofodile, U. N., & Igwe, S. E. (2017). Association of grip strength

696 with anthropometric measures: Height, forearm diameter, and middle finger length in

697 young adults. *Polish Annals of Medicine*, 24(2), 153–157.

698 <https://doi.org/10.1016/j.poamed.2016.11.008>

699 Ackermann-Liebrich, U., Leuenberger, P., Schwartz, J., Schindler, C., Monn, C., Bolognini, G.,

700 Bongard, J. P., Brändli, O., Domenighetti, G., Elsasser, S., Grize, L., Karrer, W., Keller,

701 R., Keller-Wossidlo, H., Künzli, N., Martin, B. W., Medici, T. C., Perruchoud, A. P.,

702 Schöni, M. H., ... Zemp, E. (1997). Lung function and long term exposure to air

703 pollutants in Switzerland. Study on Air Pollution and Lung Diseases in Adults

704 (SAPALDIA) Team. *American Journal of Respiratory and Critical Care Medicine*,

705 155(1), 122–129. <https://doi.org/10.1164/ajrcm.155.1.9001300>

706 Berrington de Gonzalez, A., Hartge, P., Cerhan, J. R., Flint, A. J., Hannan, L., MacInnis, R. J.,

707 Moore, S. C., Tobias, G. S., Anton-Culver, H., Freeman, L. B., Beeson, W. L., Clipp, S.

708 L., English, D. R., Folsom, A. R., Freedman, D. M., Giles, G., Hakansson, N.,

709 Henderson, K. D., Hoffman-Bolton, J., ... Thun, M. J. (2010). Body-Mass Index and

710 Mortality among 1.46 Million White Adults. *New England Journal of Medicine*, 363(23),

711 2211–2219. <https://doi.org/10.1056/NEJMoa1000367>

712 Cakmak, S., Hebborn, C., Cakmak, J. D., & Vanos, J. (2016). The modifying effect of

713 socioeconomic status on the relationship between traffic, air pollution and respiratory

714 health in elementary schoolchildren. *Journal of Environmental Management*, 177, 1–8.

715 <https://doi.org/10.1016/j.jenvman.2016.03.051>

- 716 Collister, J. A., Liu, X., & Clifton, L. (2022). Calculating Polygenic Risk Scores (PRS) in UK  
717 Biobank: A Practical Guide for Epidemiologists. *Frontiers in Genetics, 13*, 818574.  
718 <https://doi.org/10.3389/fgene.2022.818574>
- 719 Cortina, J. M. (1993). *Interaction, Nonlinearity, and Multicollinearity: Implications for Multiple*  
720 *Regression. 19(4)*, 915–922.
- 721 Engström, G., Wollmer, P., Valind, S., Hedblad, B., & Janzon, L. (2001). Blood pressure increase  
722 between 55 and 68 years of age is inversely related to lung function: Longitudinal results  
723 from the cohort study ‘Men born in 1914’: *Journal of Hypertension, 19(7)*, 1203–1208.  
724 <https://doi.org/10.1097/00004872-200107000-00004>
- 725 Frederiksen, H., Hjelmberg, J., Mortensen, J., Mcgue, M., Vaupel, J., & Christensen, K. (2006).  
726 Age Trajectories of Grip Strength: Cross-Sectional and Longitudinal Data Among 8,342  
727 Danes Aged 46 to 102. *Annals of Epidemiology, 16(7)*, 554–562.  
728 <https://doi.org/10.1016/j.annepidem.2005.10.006>
- 729 Gabel, L., Ridgers, N. D., Della Gatta, P. A., Arundell, L., Cerin, E., Robinson, S., Daly, R. M.,  
730 Dunstan, D. W., & Salmon, J. (2016). Associations of sedentary time patterns and TV  
731 viewing time with inflammatory and endothelial function biomarkers in children.  
732 *Pediatric Obesity, 11(3)*, 194–201. <https://doi.org/10.1111/ijpo.12045>
- 733 Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N., & Davey Smith, G. (2008).  
734 Mendelian randomization: Using genes as instruments for making causal inferences in  
735 epidemiology. *Statistics in Medicine, 27(8)*, 1133–1163. <https://doi.org/10.1002/sim.3034>
- 736 Lee, D.-H., Jacobs, D. R., Gross, M., Kiefe, C. I., Roseman, J., Lewis, C. E., & Steffes, M.  
737 (2003). Glutaryltransferase Is a Predictor of Incident Diabetes and Hypertension: The

- 738 Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Clinical*  
739 *Chemistry*, 8.
- 740 Liu, Y., Elsworth, B., Erola, P., Haberland, V., Hemani, G., Lyon, M., Zheng, J., Lloyd, O.,  
741 Vabistsevits, M., & Gaunt, T. R. (2021). EpiGraphDB: A database and data mining  
742 platform for health data science. *Bioinformatics*, 37(9), 1304–1311.  
743 <https://doi.org/10.1093/bioinformatics/btaa961>
- 744 Luc, G., Bard, J.-M., Lesueur, C., Arveiler, D., Evans, A., Amouyel, P., Ferrieres, J., Juhan-  
745 Vague, I., Fruchart, J.-C., & Ducimetiere, P. (2006). Plasma cystatin-C and development  
746 of coronary heart disease: The PRIME Study. *Atherosclerosis*, 185(2), 375–380.  
747 <https://doi.org/10.1016/j.atherosclerosis.2005.06.017>
- 748 Matuschek, H., & Kliegl, R. (2018). On the ambiguity of interaction and nonlinear main effects  
749 in a regime of dependent covariates. *Behavior Research Methods*, 50(5), 1882–1894.  
750 <https://doi.org/10.3758/s13428-017-0956-9>
- 751 McCaw, Z. R., Lane, J. M., Saxena, R., Redline, S., & Lin, X. (2020). Operating characteristics  
752 of the rank-based inverse normal transformation for quantitative trait analysis in genome-  
753 wide association studies. *Biometrics*, 76(4), 1262–1272.  
754 <https://doi.org/10.1111/biom.13214>
- 755 Miele, C. H., Grigsby, M. R., Siddharthan, T., Gilman, R. H., Miranda, J. J., Bernabe-Ortiz, A.,  
756 Wise, R. A., & Checkley, W. (2018). Environmental exposures and systemic hypertension  
757 are risk factors for decline in lung function. *Thorax*, 73(12), 1120–1127.  
758 <https://doi.org/10.1136/thoraxjnl-2017-210477>
- 759 Mölder, F., Jablonski, K. P., Letcher, B., Hall, M. B., Tomkins-Tinch, C. H., Sochat, V., Forster,  
760 J., Lee, S., Twardziok, S. O., Kanitz, A., Wilm, A., Holtgrewe, M., Rahmann, S.,

- 761 Nahnsen, S., & Köster, J. (2021). Sustainable data analysis with Snakemake.  
762 *F1000Research*, 10, 33. <https://doi.org/10.12688/f1000research.29032.2>
- 763 Mounier, N., & Kutalik, Z. (2023). Bias correction for inverse variance weighting Mendelian  
764 randomization. *Genetic Epidemiology*, 47(4), 314–331.  
765 <https://doi.org/10.1002/gepi.22522>
- 766 Neale. (2017, September 20). *Rapid GWAS of thousands of phenotypes for 337,000 samples in*  
767 *the UK Biobank*. [http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-](http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank)  
768 [phenotypes-for-337000-samples-in-the-uk-biobank](http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank)
- 769 O’Loughlin, J., Lambert, M., Karp, I., McGrath, J., Gray-Donald, K., Barnett, T., Delvin, E.,  
770 Levy, E., & Paradis, G. (2008). Association between cigarette smoking and C-reactive  
771 protein in a representative, population-based sample of adolescents. *Nicotine & Tobacco*  
772 *Research*, 10(3), 525–532. <https://doi.org/10.1080/14622200801901997>
- 773 Privé, F., Aschard, H., Carmi, S., Folkersen, L., Hoggart, C., O’Reilly, P. F., & Vilhjálmsson, B.  
774 J. (2022). Portability of 245 polygenic scores when derived from the UK Biobank and  
775 applied to 9 ancestry groups from the same cohort. *The American Journal of Human*  
776 *Genetics*, 109(1), 12–23. <https://doi.org/10.1016/j.ajhg.2021.11.008>
- 777 R Core Team. (2022). *R: A language and environment for statistical computing*. (4.2.1)  
778 [Computer software]. R Foundation for Statistical Computing.
- 779 Richardson, T. G., Sanderson, E., Elsworth, B., Tilling, K., & Davey Smith, G. (2020). Use of  
780 genetic variation to separate the effects of early and later life adiposity on disease risk:  
781 Mendelian randomisation study. *BMJ*, m1203. <https://doi.org/10.1136/bmj.m1203>
- 782 Sanderson, E., Glymour, M. M., Holmes, M. V., Kang, H., Morrison, J., Munafò, M. R., Palmer,  
783 T., Schooling, C. M., Wallace, C., Zhao, Q., & Davey Smith, G. (2022). Mendelian

784 randomization. *Nature Reviews Methods Primers*, 2(1), 6.  
785 <https://doi.org/10.1038/s43586-021-00092-5>

786 Schnabel, E., Nowak, D., Brasche, S., Wichmann, H.-E., & Heinrich, J. (2011). Association  
787 between lung function, hypertension and blood pressure medication. *Respiratory*  
788 *Medicine*, 105(5), 727–733. <https://doi.org/10.1016/j.rmed.2010.12.023>

789 Shlipak, M. G., Katz, R., Cushman, M., Sarnak, M. J., Stehman-Breen, C., Psaty, B. M.,  
790 Siscovick, D., Tracy, R. P., Newman, A., & Fried, L. (2005). Cystatin-C and  
791 inflammatory markers in the ambulatory elderly. *The American Journal of Medicine*,  
792 118(12), 1416.e25-1416.e31. <https://doi.org/10.1016/j.amjmed.2005.07.060>

793 Strassmann, A., De Hoogh, K., Rösli, M., Haile, S. R., Turk, A., Bopp, M., Puhan, M. A., & for  
794 the Swiss National Cohort Study Group. (2021). NO<sub>2</sub> and PM<sub>2.5</sub> Exposures and Lung  
795 Function in Swiss Adults: Estimated Effects of Short-Term Exposures and Long-Term  
796 Exposures with and without Adjustment for Short-Term Deviations. *Environmental*  
797 *Health Perspectives*, 129(1), 017009. <https://doi.org/10.1289/EHP7529>

798 Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P.,  
799 Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A.,  
800 Sprosen, T., Peakman, T., & Collins, R. (2015). UK Biobank: An Open Access Resource  
801 for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age.  
802 *PLOS Medicine*, 12(3), e1001779. <https://doi.org/10.1371/journal.pmed.1001779>

803 Sulc, J., Sjaarda, J., & Kutalik, Z. (2021). *Polynomial Mendelian Randomization reveals*  
804 *widespread non-linear causal effects in the UK Biobank* [Preprint]. *Genetics*.  
805 <https://doi.org/10.1101/2021.12.08.471751>

- 806 Sun, Y.-Q., Burgess, S., Staley, J. R., Wood, A. M., Bell, S., Kaptoge, S. K., Guo, Q., Bolton, T.  
807 R., Mason, A. M., Butterworth, A. S., Di Angelantonio, E., Vie, G. Å., Bjørngaard, J. H.,  
808 Kinge, J. M., Chen, Y., & Mai, X.-M. (2019). Body mass index and all cause mortality in  
809 HUNT and UK Biobank studies: Linear and non-linear mendelian randomisation  
810 analyses. *BMJ*, 11042. <https://doi.org/10.1136/bmj.11042>
- 811 Tabak, C., Spijkerman, A. M. W., Verschuren, W. M. M., & Smit, H. A. (2009). Does educational  
812 level influence lung function decline (Doetinchem Cohort Study)? *European Respiratory*  
813 *Journal*, 34(4), 940–947. <https://doi.org/10.1183/09031936.00111608>
- 814 Tian, H., Mason, A. M., Liu, C., & Burgess, S. (2023). Relaxing parametric assumptions for non-  
815 linear Mendelian randomization using a doubly-ranked stratification method. *PLOS*  
816 *Genetics*, 19(6), e1010823. <https://doi.org/10.1371/journal.pgen.1010823>
- 817 Tracy, R. P., Psaty, B. M., Macy, E., Bovill, E. G., Cushman, M., Cornell, E. S., & Kuller, L. H.  
818 (1997). Lifetime Smoking Exposure Affects the Association of C-Reactive Protein with  
819 Cardiovascular Disease Risk Factors and Subclinical Disease in Healthy Elderly  
820 Subjects. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 17(10), 2167–2176.  
821 <https://doi.org/10.1161/01.ATV.17.10.2167>
- 822 Wheeler, B. W. (2005). Environmental equity, air quality, socioeconomic status, and respiratory  
823 health: A linkage analysis of routine data from the Health Survey for England. *Journal of*  
824 *Epidemiology & Community Health*, 59(11), 948–954.  
825 <https://doi.org/10.1136/jech.2005.036418>
- 826 Zekavat, S. M., Honigberg, M., Pirruccello, J. P., Kohli, P., Karlson, E. W., Newton-Cheh, C.,  
827 Zhao, H., & Natarajan, P. (2021). Elevated Blood Pressure Increases Pneumonia Risk:

828           Epidemiological Association and Mendelian Randomization in the UK Biobank. *Med,*  
829           2(2), 137-148.e4. <https://doi.org/10.1016/j.medj.2020.11.001>

830   Zhang, Z., Ma, L., Geng, H., & Bian, Y. (2021). Effects of Smoking, and Drinking on Serum  
831           Gamma-Glutamyl Transferase Levels Using Physical Examination Data: A Cross-  
832           Sectional Study in Northwest China. *International Journal of General Medicine, Volume*  
833           14, 1301–1309. <https://doi.org/10.2147/IJGM.S301900>

834

835



836

837

## **Supplementary Material**

### **Sensitivity Analyses**

839

Multiple sensitivity analyses were performed to account for potential biases in our

840

effects. Amongst them are the effect of level 1 interactions, meaning an interaction between the

841

GRS and the environment on the exposure X, the potential for the environment to act as a

842

collider between the exposure and the outcome, the risk for the interactions to depend on the

843

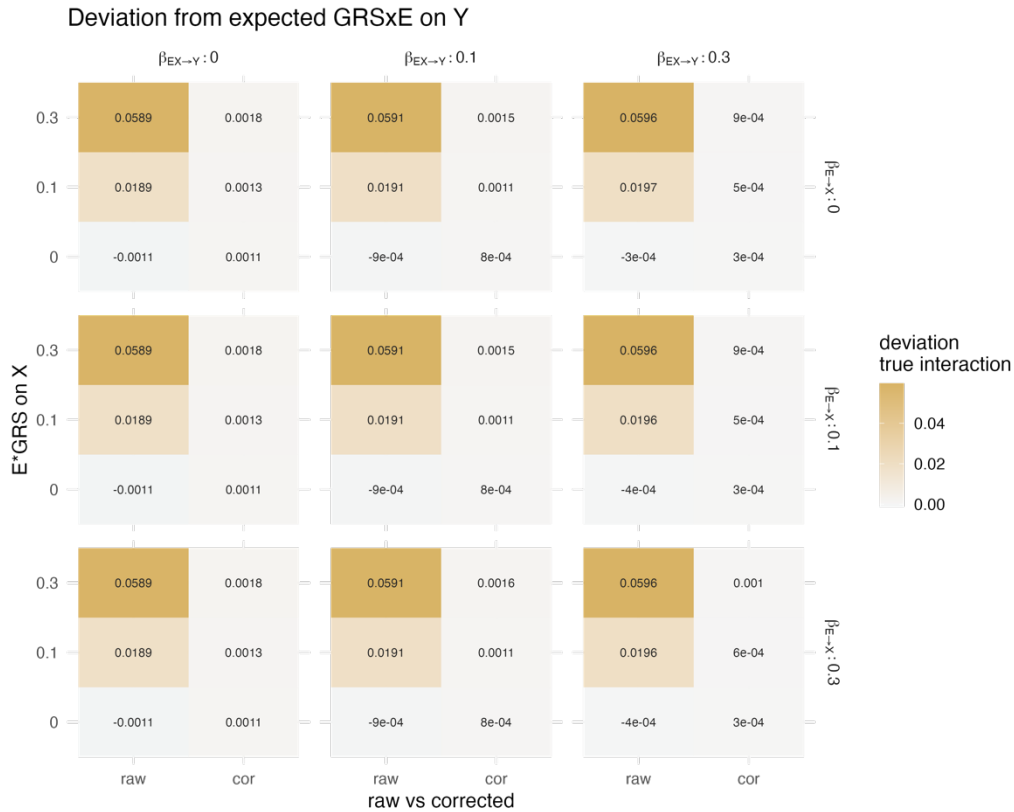
standardization or transformation of the outcome variable and the potential for the detected

844

interactions to be method-specific.

845

846 **Level 1 interactions**



Supplementary Figure 1: Bias in interaction effect in response to varying level 1 interactions if the simulated effect of the exposure on the outcome is 0.2. We observe that the bias translates to  $\beta_{GRS_X \cdot E \rightarrow X} \cdot \beta_{X \rightarrow Y}$  and can be corrected for by subtracting the according term from the interaction estimate, which attenuated the bias in interaction estimates to zero.

847 Assuming the exposure X is defined as follows:

848 
$$X = \beta_{E \rightarrow X} \cdot E + \beta_{GRS_X \rightarrow X} \cdot GRS_X + \beta_{GRS_X \cdot E \rightarrow X} \cdot GRS_X \cdot E + q_{U \rightarrow X} \cdot U + \varepsilon_X,$$

849 while the outcome Y is – amongst other things – a function of X, and  $\beta_{GRS_X \cdot E \rightarrow X}$  deviates

850 from zero, we will overestimate the effect of  $\beta_{GRS_X \cdot E \rightarrow Y}$  by  $\beta_{GRS_X \cdot E \rightarrow X} \cdot \beta_{X \rightarrow Y}$ . We validated this

851 in simulation settings (Supplementary Figure 1). We investigated to what extent the presence of

852 level 1 interactions may affect our interaction estimates from the application study. For 160

853 interactions with a tier score > 0, there was none where correction for level 1 interaction would

854 lead to a change in the sign of the interaction estimate. Furthermore, there were only 5 settings

855 where the absolute ratio between the obtained interaction estimate and the correction was smaller

856 than 5, namely the interaction between Body fat percentage and Nitrogen dioxide on CRP, the  
857 interactions between water mass and TV time on forced expiratory volume and hand grip  
858 strength, the interaction between CRP and age on systolic blood pressure and finally the  
859 interaction between height and socioeconomic deprivation on forced expiratory volume.

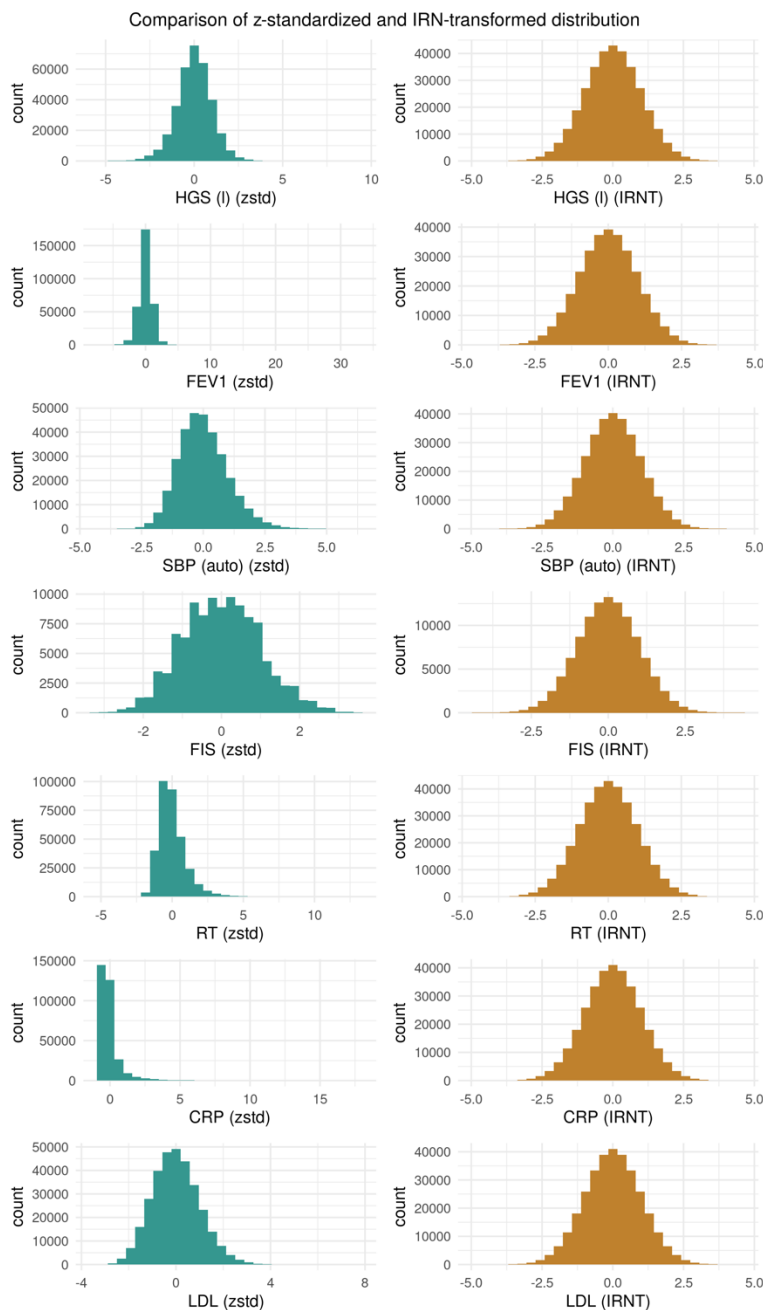
#### 860 **Exposure on environment effects**

861 For some settings, we observed that there was a causal effect of the exposure on the  
862 environment. For these settings, we obtained the (observational, as the environment often is not  
863 meaningfully genetically instrumentable) effect of the environment on the exposure. We obtained  
864 the ratio  $\frac{\beta_{X \rightarrow E}}{\beta_{E \rightarrow X}}$  and the according 95% confidence intervals. If the lower absolute confidence  
865 interval of the  $\frac{\beta_{GRS_{X \rightarrow E}}}{\beta_{E \rightarrow X}}$  ratio was larger than 0.5, we considered the interaction as potentially  
866 biased and shaded it out in the according plots. For example, for body fat percentage we found a  
867 relatively strong effect on SES, TV time, and sedentary behavior, relative to the reverse effect.  
868 Noteworthy, we also found evidence for some causal effects on age, namely of heart rate,  
869 Gamma GLT, and SBP. Those likely result from sample bias, meaning that people with  
870 genetically high systolic blood pressure do not feel well enough to participate in such a study at  
871 an age where their peers still are able to participate.

872

873 **Inverse rank normal transformed outcome phenotypes**

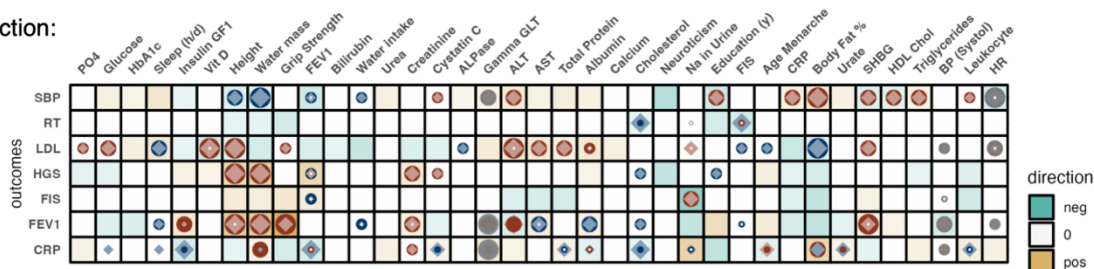
874 We replicated our analyses whereby we inverse rank normal transformed the outcome  
875 phenotypes (which by default were z-standardized). For the majority of settings (136 out of 182  
876 interactions with a tier score  $> 0$ , 74.7%), the IRNT outcome phenotypes yielded interaction  
877 results that agreed well (tier  $> 0$ , agreement in direction) with the z-standardized outcome  
878 phenotypes. It is worth noting, though, that we observed in 44 (24.2%) of all interactions with a  
879 tier score  $> 0$ , that the result could not be replicated when the outcome was IRN-transformed  
880 (tier score for IRN-transformed outcome = 0). There were two settings where both, the  
881 interaction on the IRN-transformed outcome and the interaction on the z-standardized outcome  
882 yielded a tier score  $> 0$  but the direction of the effect did not agree, namely the interaction  
883 between FEV1 and age on HGS and the interaction between Body fat % and age on CRP.  
884 Finally, there were 55 interactions which only yielded a tier score  $> 0$  when the outcome was  
885 IRN-transformed, but not when it was z standardized. The majority of interactions that deviated  
886 depending on the preprocessing of the outcome phenotype were interactions affecting CRP (64  
887 of 101 settings), RT (12 of 101 settings), and FEV1 (11 of 101 settings). This finding is little  
888 surprising given the heavy right skewing of the raw (and thus z-standardized) CRP phenotype,  
889 indicating that some of the interactions may be driven by the extreme values on the upper end of  
890 the scale. As these values may be of distinctive mechanistic relevance, it depends on the specific  
891 research question whether it is more accurate to consider the results obtained using the z-  
892 standardized or the IRN-transformed outcome phenotype.



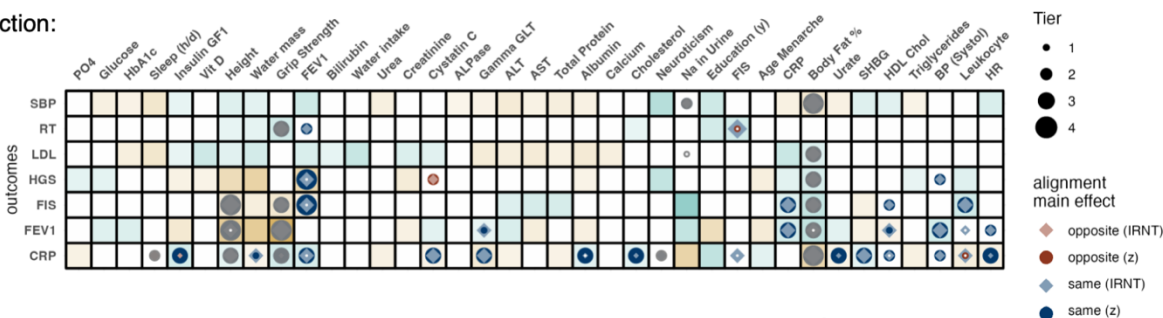
Supplementary Figure 2. Density plots of the z-standardized (zstd, blue, left) vs. inverse rank normal transformed (IRNT, right) outcome phenotypes. HGS (l): hand grip strength (left), FEV1: forced expiratory volume within 1 second, SBP (auto): systolic blood pressure, automated reading, FIS: fluid intelligence score, RT: reaction time at pattern matching task, CRP: C-reactive protein, LDL: Low density Lipoprotein.

Supplementary Figure 3. Replication of findings when outcome phenotype is not z-standardized but inverse-rank normal transformed (IRNT). The color of the tiles represents the direction of the main effect if there was evidence for a main effect. The round shapes represent the interactions with a tier score > 0 obtained when the outcome was z-standardized, whereby the color indicates the direction of the interaction relative to the main effect. The rhombuses represent the interactions with tier > 0 obtained when the outcome phenotype was inverse rank normal transformed, whereby the color indicates the direction of the interaction relative to the main effect. The size represents the tier score, with larger shapes indicating higher confidence. The small white dot indicates if an interaction was corrected due to evidence for a quadratic effect of the exposure on the outcome and an effect of the environment on the exposure. Finally, interactions are shaded out if there was evidence for a stronger effect of the exposure on the environment than vice versa, as this likely leads to biased interaction estimates. AP: Air pollution, ALT: Alanine Aminotransferase, AST: Aspartate Aminotransferase, BP: Blood pressure, CRP: C-reactive Protein, FEV1: Forced Expiratory Volume within 1 second, FIS: Fluid Intelligence Score, HbA1c: Glycated hemoglobin, HDL: High density lipoprotein, HR: Heart rate, NO<sub>2</sub>: Nitrogen dioxide, SHBG: sex hormone binding globulin, PO<sub>4</sub>: phosphate, TDI: Townsend deprivation index, Vit D: Vitamin D

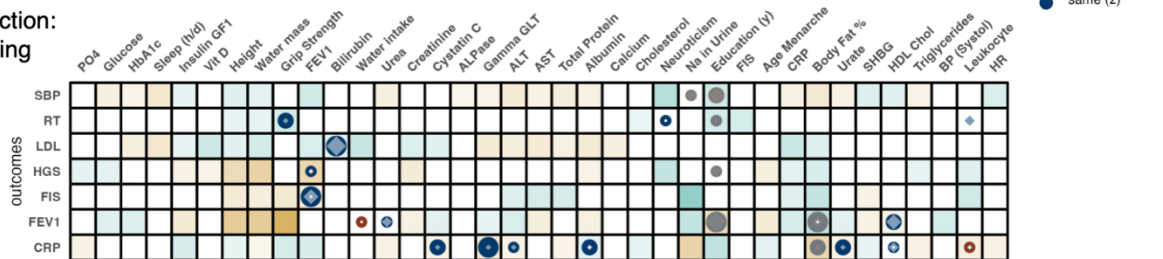
**Interaction:  
Age**



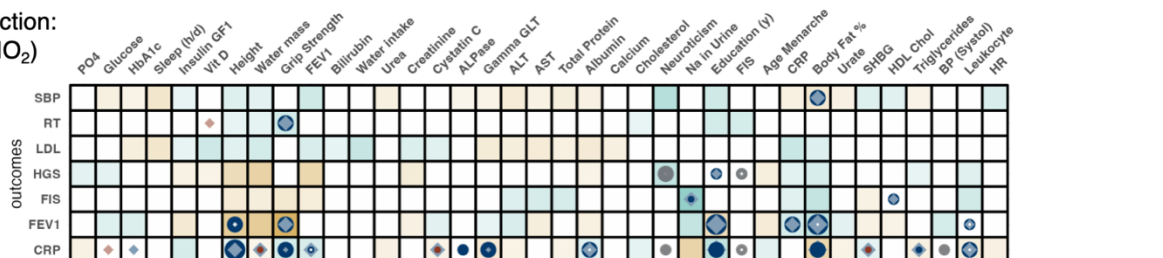
**Interaction:  
TDI**



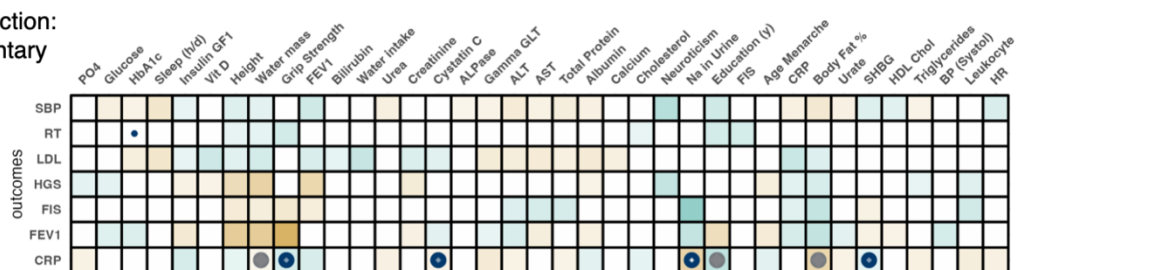
**Interaction:  
Smoking**



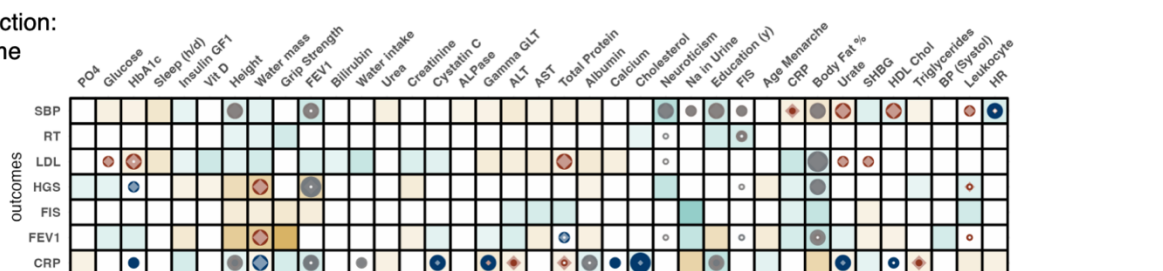
**Interaction:  
AP (NO<sub>2</sub>)**



**Interaction:  
Sedentary**



**Interaction:  
TV time**



895           **Replication analysis**

896           To validate the present interaction effects, we assessed the same settings in an extended  
897 doubly ranked stratification (Tian et al., 2023) framework (DRS). Whilst the doubly ranked  
898 stratification method was developed to investigate non-linear exposure-outcome relationships, it  
899 can be adapted to allow for investigating interactions by stratifying by an environmental variable  
900 (instead of the exposure value) adjusted by the exposure's instrument and regressing the obtained  
901 MR estimates on the predefined strata midpoint value. Applying DRS to the same phenotype  
902 combination as 2SLS-I, we found substantial agreement between the two methods. In 136 (85 %) of all  
903 interactions with a tier score  $> 0$ , the direction of DRS aligned. In 30 (18.75%) of all  
904 interactions with a tier score  $> 0$ , the doubly ranked method agreed in sign and significance. Vice  
905 versa, out of 50 settings where DRS yielded significance at  $p < 0.001$ , 2SLS-I yielded a tier score  
906  $> 0$  in 30 (60 %). There was no setting where both, the DRS (at  $p < 0.001$ ) and 2SLS-I (tier  $> 0$ )  
907 yielded significant effects which did not agree in sign. In summary, DRS and 2SLS-I yield  
908 comparable results (*Supplementary Figure 4*) Whilst this contributes to the confidence in the  
909 results obtained with 2SLS-I, it should be noted that there is a potential for both, 2SLS-I and the  
910 DRS yielding false-positive (e.g. if the environment acts as collider between the exposure and  
911 the outcome) or false-negative (e.g. due to lacking power) results.

912

*Supplementary Figure 4. All interaction effects from 2SLS-I were replicated using a doubly ranked stratification (DRS) method. The figure summarizes the effect agreement between 2SLS-I and DRS. Tiles are filled if all three estimates (strict GRS of 2SLS-I, extended GRS of 2SLS-I, and DRS) agree in the direction of the effect, with the fill color indicating the direction of the interaction. For each estimate that yielded significance at  $p < 0.05$ , a shape is plotted to indicate the level of significance: a circle ( $p < 0.05$ ), a rhombus ( $p < 0.01$ ) or a triangle ( $p < 0.001$ ). The position of the according shape within each tile indicates which method has yielded the according significance level, with the strict GRS of 2SLS-I being on the top left, the extended GRS of 2SLS-I on the top right and the DRS at the bottom. The color of each of these shapes indicates the direction of the effect. Finally, the number in the middle of each tile relates to the obtained tier score, i.e. the confidence rating obtained using 2SLS-I.*

