Inference of causal and pleiotropic effects with multiple weak genetic instruments: application to effect of adiponectin on type 2 diabetes

Paul M McKeigue [1], Andrii Iakovliev [2], Athina Spiliopoulou [1], Helen M Colhoun [2]

1 Usher Institute, College of Medicine and Veterinary Medicine, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, Scotland.

2 Institute of Genetics and Cancer, College of Medicine and Veterinary Medicine, University of Edinburgh, Western General Hospital Campus, Crewe Road, Edinburgh EH4 2XUC, Scotland.

Abstract

Current methods for Mendelian randomization studies have several limitations: to construct unlinked genetic instruments they can use only one SNP from each clump of exposure-associated SNPs, they require that weak instruments are excluded, and they rely on makeshift procedures for downweighting outliers to allow inference of causality in the presence of pleiotropic effects. This paper describes methods that overcome these limitations. A scalar instrument is constructed from all exposure-associated SNPs in each clump, and inference of causality is based on marginalizing over the distribution of pleiotropic effects.

To demonstrate the approach, we tested the effect of circulating levels of adiponectin, encoded by *ADIPOQ*, on the risk of type 2 diabetes. Genotypic instruments were constructed from 24 unlinked *trans*-pQTLs detected in Icelanders using the Somalogic platform and 43 detected in the UK Biobank study using the Olink platform. These instruments were tested for association with type 2 diabetes in a non-overlapping subset of the UK Biobank cohort. In contrast to the results of earlier Mendelian randomization studies of adiponectin that used only a *cis*-pQTL, with multiple *trans*-pQTLs there was clear evidence for a causal effect: standardized log odds ratio -0.38 (95% CI -0.5 to -0.25) using DeCODE instruments and -0.33 (95% CI -0.43 to -0.23) using UK Biobank instruments.

Guidelines for the design of Mendelian randomization studies that recommend exclusion of weak instruments, or restricting the instruments to *cis*-acting variants where the exposure under study is a gene transcript or gene product, should be reassessed.

Introduction

Instrumental variable analysis with genetic instruments ("Mendelian randomization") has been widely used to infer causal effects of exposures (broadly defined to include behavioural traits, biomarkers and gene expression levels) on diseases. The biggest methodological challenge is how to infer causality when some of the genetic instruments have direct (pleiotropic) effects on the outcome that are not mediated through the exposure under study. These pleiotropic effects are usually not directly observed, and their distribution over the instruments is unknown. The most widely-used methods use summary statistics for the coefficients for the effect of each SNP on the outcome and on the exposure from non-overlapping samples.[1] A typical workflow is to select, from each genomic region in which genetic associations with the exposure have beeen detected, a single SNP that is reliably associated with the exposure. "Estimators" for the causal effect are constructed from the ratios of SNP-outcome to SNP-exposure coefficients. These methods have several limitations:

- To ensure that the instruments are unlinked and their effects can be modelled as independent, only one variant in each genomic region is used to construct each instrument. This does not use all the available information about the effects of variants in the region on the exposure, and the results may depend upon which variants are selected.
- Because the estimators are based on the sampling distribution of ratios of coefficient estimates, "weak instruments" with small or uncertain effects on the exposure have to be excluded because the sampling distribution of a ratio is badly-behaved when the denominator (in this case the SNP-exposure coefficient) is small and is estimated with uncertainty. This exclusion of "weak instruments" limits the number of unlinked instruments that can be included in the analysis. Where the exposure under study is the expression of a gene or the level of a circulating protein, this is likely to exclude many *trans*-eQTLs and *trans*-pQTLs which typically have weak effects.[2]
- Inference of causality in the presence of pleiotropic effects relies on makeshift procedures for the construction of estimators that downweight or exclude instruments that are outliers, for instance by calculating a weighted median. The results depend on which estimator is used: guidelines suggest that "investigators should pick a sensible range of methods to assess the sensitivity of their findings."[2]

This paper describes methods to overcome these limitations, and demonstrates their application to infer a causal effect of the protein adiponectin, encoded by *ADIPOQ*, on type 2 diabetes.

Methods

Statistical model

For a Mendelian randomization study with J unlinked genetic instruments, we specify a model with three parameters:

- $\boldsymbol{\alpha}$ vector of coefficients of effects of the instruments on exposure X.
- $\boldsymbol{\beta}$ vector of coefficients of direct (pleiotropic) effects of the instruments on outcome Y
- θ causal effect of X on Y

From summary statistics we have estimates $\hat{\boldsymbol{\alpha}}$ of the effects $\boldsymbol{\alpha}$ of the instruments on the exposure, with corresponding standard errors $\boldsymbol{s}_{\boldsymbol{\alpha}}$. As the instruments are unlinked, we can model the estimates $\hat{\boldsymbol{\alpha}}$ as independent Gaussian variables conditional on the true values $\boldsymbol{\alpha}$:

$$\hat{\alpha}_j \sim \mathcal{N}\left(\alpha_j, s^2_{\alpha(j)}\right)$$

With individual-level data on genotypic instruments $\boldsymbol{Z}$, covariates $\boldsymbol{X_u}$ and outcome Y in N individuals, we can model the dependence of Y on Z and $\boldsymbol{X_u}$ with a generalized linear model:

$$g\left(\mathbb{E}\langle Y\rangle\right) = \beta_0 + \boldsymbol{X_u} \cdot \boldsymbol{\beta_u} + \boldsymbol{Z} \cdot \boldsymbol{\beta} + \theta \boldsymbol{Z} \cdot \boldsymbol{\alpha}$$

where $g\left(\right)$ is a link function (logit for a logistic regression). With priors on $\beta_0, \boldsymbol{\beta_u}, \boldsymbol{\beta}, \theta$ we can compute the joint posterior distribution of these parameters given $\hat{\boldsymbol{\alpha}}, \boldsymbol{s_\alpha}, \boldsymbol{Z}, \boldsymbol{X_u}, Y$.

Where we have only summary-level data on the dependence of outcome Y on genotypic instruments Z (or to reduce the computational burden of an individual-level analysis), we can model the observed coefficient estimates $\hat{\boldsymbol{\gamma}}$ of the regression of Y on Z as independent Gaussian variables conditional on the true values $\boldsymbol{\gamma}$, using the standard errors $\boldsymbol{s_\gamma}$ of the estimates to set the prior standard deviations.

$$\hat{\gamma}_j \sim \mathcal{N}\left(\gamma_j, s^2_{\gamma(j)}\right)$$

The crude effect of each instrument on the outcome is the sum of the direct effect and the causal effect.

$$\gamma_j = \beta_j + \theta \alpha_j$$

With priors on $\boldsymbol{\beta}, \theta$ we can compute the joint posterior distribution of these parameters given $\hat{\boldsymbol{\gamma}}, \boldsymbol{s_\gamma}, \hat{\boldsymbol{\alpha}}, \boldsymbol{s_\alpha}$.

As the form of the distribution of pleiotropic effects over loci is unknown, any realistic statistical model has to specify a prior on these effects that encompassses a broad family of symmetric distributions ranging from a spike-and-slab to a Gaussian. One such prior is the horseshoe prior,[3] which has been applied to inference of causal effects from Mendelian randomization using individual-level data[4] or summary-level data.[5] A limitation of the original horseshoe prior is that there is no way of specifying a realistic prior on the size of the nonzero effects in the "slab" component. This is especially relevant to genetics, as effects of common genetic variants on a complex trait are usually small. The regularized horseshoe prior, known as the "Finnish horseshoe", overcomes this problem.[6] We name our method "MR-Hevo", using the Finnish word for a horse to distinguish it from the recently-described "MR-Horse" method which uses an unregularized horseshoe prior.[5] The statistical model is described in Supplementary Information.

To generate the posterior distribution given the model and the data, we use the program `Stan` to sample the posterior distribution. For the large datasets in this example, we have used summary-level data to keep the computation time short. To obtain the likelihood as a function of the parameter of interest, we divide the posterior by the prior on that parameter. In large samples, the log-likelihood is asymptotically quadratic and the maximum likelihood estimate of θ, its confidence interval and a test of the null hypothesis $\theta = 0$ can be obtained from this quadratic approximation. We estimated the likelihood as a function of the causal effect parameter θ by fitting a kernel density to the posterior samples of θ, weighting each observation by the inverse of the prior. A quadratic function was fitted by least-squares to the logarithm of this likelihood function.

Constructing scalar instruments from multiple SNPs

From summary statistics we have for each clump of exposure-associated SNPs a vector $\hat{\boldsymbol{\alpha}}_{\boldsymbol{u}}$ of univariate coefficient estimates for the effect of the SNPs on the exposure. Multivariable coefficient estimates $\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}$ are calculated by premultiplying the univariate coefficients by the inverse of the correlation matrix $\boldsymbol{\Sigma_G}$ between the SNP genotypes.

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}} = \boldsymbol{\Sigma}_{\boldsymbol{G}}^{-1} \hat{\boldsymbol{\alpha}}_{\boldsymbol{u}}$$

The correlation matrix $\boldsymbol{\Sigma}_{\boldsymbol{G}}$ is obtained from a reference panel such as 1000 Genomes. A shrinkage penalty (equivalent to ridge regression) can be imposed by adding a penalty factor to the diagonal elements of $\boldsymbol{\Sigma}_{\boldsymbol{G}}$. Where the correlation matrix is singular or ill-conditioned, a pseudo-inverse can be used to calculate the multivariable coefficients. For an individual with genotypes $\boldsymbol{G}$ at the exposure-associated SNPs, a locus-specific score S predicting the exposure is calculated as $\boldsymbol{G} \cdot \hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}$.

Because the score S is calculated from the genotypes and the genotype-exposure coefficients, we cannot use it as a genotypic instrument in a model of the relationship of the genotype-outcome coefficients to the genotype-exposure coefficients. We can factor the dot product $\boldsymbol{G} \cdot \hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}$ as the product of two scalars: the magnitude of the multivariable coefficient vector $\|\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}\|$ and a pseudo-genotype $\|\boldsymbol{G}\| \rho_{G,\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}}$, where $\rho_{G,\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}}$ is the correlation between $\boldsymbol{G}$ and $\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}$, geometrically equivalent to the cosine of the angles between these vectors. We can then substitute $\|\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}\|$ for the scalar coefficient estimate $\hat{\alpha}$ and $S/\|\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}\|$ for the scalar instrument Z as a pseudo-genotype in the statistical model. The derived coefficient for the estimated effect of the instrument on the exposure is always positive; flipping the coding of the alleles would flip the sign of $\rho_{G,\alpha}$. We can calculate the standard error of $\hat{\alpha}$ as $1/\sqrt{\mathcal{I}_\alpha}$ where $\mathcal{I}_\alpha$ is the Fisher information on the slope of the linear regression of X on Z, given by

$$\mathcal{I}_\alpha = N_\alpha \frac{\mathrm{Var}\,(Z)}{\sigma_X^2 - \hat{\alpha}^2 \mathrm{Var}\,(Z)} \quad ,$$

where N_α is the sample size of the study from which the estimated univariate coefficients $\hat{\boldsymbol{\alpha}}_{\boldsymbol{u}}$ were obtained, $\mathrm{Var}\,(Z)$ is the variance of the scalar instrument Z estimated in a reference panel, and σ_X^2 is the variance of the exposure X. If an estimate of σ_X^2 is not given, it can be calculated from the allele frequency a_j of the jth SNP, the standard error $s_{G(j)}$ of the jth univariate coefficient estimate $\hat{\alpha}_{u(j)}$, and the sample size N_α as

$$\sigma_X^2 = 2a_j\,(1 - a_j)\left(N_\alpha s_{G(j)}^2 + \alpha_{u(j)}^2\right) \quad .$$

An estimate $\hat{\gamma}$ of the coefficient for the effect of the scalar instrument Z on the outcome Y and its standard error can be obtained by fitting a regression of Y on Z. A method for calculating the instrument-outcome coefficient $\hat{\gamma}$ and its standard error from summary-level data for SNP-outcome coefficients is described in the Appendix, but not used in this paper.

For each instrument we standardize the pseudo-genotypes Z to unit variance and scale the coefficient estimates accordingly, so that the variance explained by the jth linear predictor $Z_j \hat{\alpha}_j$, which can be interpreted as the "strength of the instrument", is proportional to the square of the scalar coefficient $\hat{\alpha}_j$. The prior on the scale of the pleiotropic effects $\boldsymbol{\beta}$ of the instruments Z is thus independent of the strength of the instrument.

Example: effect of adiponectin on type 2 diabetes

As the methods described above overcome some of the limitations of standard methods for design and analysis of Mendelian randomization studies, they may help to resolve contradictory results obtained in earlier studies. As an example, we examine the effect of adiponectin, encoded by the *ADIPOQ* gene, on the risk of type 2 diabetes. Low adiponectin levels are strongly associated with type 2 diabetes. Of four previous studies using Mendelian randomization to investigate whether low adiponectin levels cause type 2 diabetes, two reported no evidence of a causal effect[7,8] and two reported evidence for a causal effect.[9,10] We obtained coefficient estimates for the effects of SNPs on circulating levels of adiponectin from two studies:

- The DeCODE study of 35559 Icelanders in which 4719 proteins on the SomaLogic v4 panel were measured in plasma.[11] 2207 aptamers on this platform that appeared to cross-react with complement factor H were excluded. The criteria for identifying these aptamers were: a *trans*-pQTL at the *CFH* locus and no *cis*-pQTL for the protein that the aptamer was designed to detect and association of the *trans*- score with age-related macular degeneration.

- The UK Biobank proteomics study of 54306 individuals in which 2924 proteins on the Olink Explore panel were measured in plasma[12]

The calculated standard deviation of adiponectin levels was 1.14 in the DeCODE dataset and 0.83 in the UK Biobank dataset. In each genomic region where there was at least one SNP-exposure association with $p < 10^{-6}$, all SNPs with $p < 10^{-6}$ separated by no more than 1 Mb were included in the calculation of the multivariable coefficient vector. This yielded 26 unlinked genotypic instruments from the DeCODE study, calcuated from 567 SNPs, and 44 instruments from the UK Biobank proteomics study, calculated from 945 SNPs. Each instrument was annotated with a list of nearby genes, and the genes on each list were matched to a list of 256 genes for which SNP associations with type 2 diabetes have been reported at $p < 5 \times 10^{-8}$, extracted from GWAS Catalog.

The target dataset with genotypes and type 2 diabetes status was the UK Biobank cohort. Ethical approval for the UK Biobank study was granted in 2011 by the North West Multi-centre Research Ethics Committee (11/NW/0382), and renewed every five years since then. Informed consent was obtained for all participants in UK Biobank. The work described herein was approved by the UK Biobank under application number 23652.

Separate analyses were undertaken using the instruments constructed from these two proteomics studies. When using instruments obtained from the UK Biobank proteomics study, the 54306 participants who were in the proteomics study were excluded from the target dataset.

Because the individuals in the proteomics studies were predominantly of European ancestry, the target dataset was restricted to individuals of European ancestry. Of 451447 unrelated individuals with nonmissing phenotype and genotype data, 37981 were classified as cases of type 2 diabetes and the all others were classified as noncases.

For each instrument, the pseudo-genotypes were standardized to unit variance and a logistic regression model was fitted with type 2 diabetes as outcome variable. The other covariates were sex and the first five principal components of the SNP relationship matrix. The coefficients for the effect of each scalar instrument on type 2 diabetes were used to fit the model. *Cis*-pQTLs were excluded. As described in the Appendix, the regularized horseshoe requires us to specify priors on two parameters: the scale s_{global} of the prior on the global shrinkage parameter, which encodes a guess of the fraction f of genetic instruments that have nonzero pleiotropic effects, and the scale s_{slab} of the prior on the regularization parameter, which encodes information about the likely size of these nonzero pleiotropic effects. For an initial analysis we set s_{global} based on a guess that $f = 0.2$, and set $s_{\text{slab}} = 0.25$. The sensitivity of the results to these prior settings was examined by repeating the model fitting with s_{slab} set based on a guess that $f = 0.8$, and $s_{\text{slab}} = 0.025$.

For comparison with inference based on the marginal likelihood of the causal effect parameter, three widely-used estimators for the causal effect were calculated from the coefficient ratios: the inverse-variance weighted mean, the weighted median, and the penalized weighted median.

Results

Table 1 shows summary statistics for the genetic instruments for adiponectin derived from each pQTL study. There is suprisingly little overlap between the 24 *trans*-pQTLs detected in

the DeCODE study and the 43 *trans*-pQTLs detected in UK Biobank. Only eleven *trans*-pQTLs – gene-poor regions at 219.4 Mb on chromosome 1, 34.2 Mb on chr 6, 139.5 Mb on chromosome 6, and regions containing *ALAS1*, *ADRB1*, *SPON1*, *PDE3A*, *ABCB9*, *CDH13*, *AKR1B1P7*, and *MIR6813* – were detected in both studies at a threshold of $p < 10^{-6}$. This may reflect the low power to detect weak *trans*-pQTLs even in these large datasets. In both studies there is a *cis*-QTL at the *ADIPOQ* transcription site. In the DeCODE study an extended *cis*-pQTL is detected 2 Mb upstream of the transcription site. Only two pQTLs – *SPON1* and *APOC1* – contained genes that were listed in GWAS Catalog as associated with type 2 diabetes.

Figure 1 shows that 22 of 24 *trans*-pQTL instruments from DeCODE and 39 of 43 *trans*-pQTL instruments from UKBB are inversely associated with type 2 diabetes. The *cis*-pQTL instruments are not associated with type 2 diabetes.

Table 2 shows the estimates of the causal effect parameter obtained by each method, as log odds ratios for unit change in adiponectin levels. The maximum likelihood estimates obtained by marginalizing over direct effects were -0.38 (95% CI -0.5 to -0.25) using DeCODE instruments and -0.33 (95% CI -0.43 to -0.23) using UK Biobank instruments. In comparison with the weighted median or penalized weighted median estimators, the confidence intervals obtained from the likelihood function were wider and the p-values were more conservative.

Supplementary Figure S1 shows the posterior distribution and log-likelihood of the causal effect parameter. As expected with this large sample size, the posterior distribution is approximately Gaussian and the log-likelihood is well approximated by a quadratic function. Supplementary Table S1 shows posterior summaries of the model parameters. The posterior mean of the effective fraction f of instruments that have nonzero pleiotropic effects was 0.48 (90% credible interval 0.39 to 0.58) in a model using the DeCODE instruments and 0.49 (90% credible interval 0.42 to 0.56) in a model using the UK Biobank instruments, somewhat higher than the prior guess that $f = 0.2$.

The standard deviation (evaluated as the square root of the posterior mean of the variance) of these nonzero pleiotropic effects was 0.019 for the DeCODE instruments and 0.023 for the UK Biobank instruments. These values are far smaller than the prior guess that the scaling factor $s_{\text{slab}} = 0.25$.

Although the model fitting infers that about half the instruments have nonzero pleiotropic effects (Figure 2), only a few instruments were clearly identifiable as having nonzero pleiotropic effects. Using an upper 90% credible limit of 0.25 for the shrinkage coefficient κ to identify those instruments that had mostly escaped shrinkage, only two instruments from the DeCODE study (*ABCB9*, *APOC1*) and two instruments from the UK Biobank study (*CDH13*, *CCND2*) were clearly pleiotropic.

When the model fitting was repeated with prior settings changed as described in the Methods section, the maximum likelihood estimates and confidence intervals for θ differed only slightly from those obtained with the initial prior settings: -0.36 (95% CI -0.51 to -0.21) using DeCODE instruments and -0.33 (95% CI -0.43 to -0.23) using UK Biobank instruments.

Discussion

Bayesian methods for marginalizing over the unobserved pleiotropic effects to infer the causal effect parameter in a Mendelian randomization study have been described previously.[4,5] The methods described in this paper extend this approach: (1) by allowing unlinked scalar genetic instruments to be constructed from all SNPs associated with the exposure without having to restrict to one SNP from each genomic region; (2) by using a regularized horseshoe prior which allows encoding of prior knowledge that genetic effects of a single genomic region on a complex trait are usually small; and (3) by constructing classical hypothesis tests based on the likelihood of the causal effect parameter.

The difficulties associated with using ratios of coefficients to construct estimators are eliminated in a Bayesian framework, where nothing is "estimated". Where the log-likelihood is asymptotically quadratic, statistical theory guarantees that the maximum-likelihood estimate has the sampling properties that are desirable for an estimator: consistency, minimum variance and unbiasedness. Where the log-likelihood is not approximately quadratic, it is preferable to base inference directly on the likelihood function rather than on the sampling properties of an estimator. Because Bayesian inference does not rely on the sampling properties of ratios, there is no need to exclude "weak instruments". This makes it possible to integrate Mendelian randomization into genome-wide aggregated *trans*- effects analysis, which attempts to identify core genes on which the polygenic effects of many variants coalesce via *trans*- effects to influence complex traits.[13] In this approach, the genetic instruments are clumps of SNPs with *trans*- effects on the expression of a gene as transcript or circulating protein.

Using instruments constructed from *trans*-pQTLs in two different studies that used different assays for the protein, there is clear support for a causal effect of low adiponectin levels on type 2 diabetes. This is apparent even without a formal statistical analysis, as the signs of the effects of the instruments are almost all negative. Earlier studies that failed to detect a causal effect were restricted to *cis*-acting SNPs.[7,8] Some of these *cis*- acting SNPs affect not only the measured levels but also the splicing of the gene product,[14] so their effects on measured adiponectin levels do not necessarily correspond to effects on the functional protein. The most direct human evidence for a causal effect of low adiponectin on type 2 diabetes comes from reports of families in which type 2 diabetes segregated with rare loss-of-function variants in *ADIPOQ*.[15] In the obese mouse model, overexpression of adiponectin in the liver protected against diabetes.[16] Though adiponectin analogues have been studied experimentally, none has reached the clinical stage of drug development. Misleading results from *cis*- Mendelian randomization studies may have discouraged development of a possible drug target.

Most of the SNP heritability of gene expression is attributable to *trans*- effects.[17] Where, as in this example, the exposure under study is a gene transcript or a circulating protein, study designs that use only *cis*- acting variants as genetic instruments have been advocated.[18] The rationale for thse "*cis*-MR" designs is that *cis*- acting variants are less likely to have pleiotropic effects on the outcome that are not mediated through the level of the circulating protein. This supposition is questionable: especially where *cis*-acting variants affect the splicing of the transcript, the relationship between measured levels and functional activity of the protein may be broken. Because *cis*- acting variants often have large effects on functional activity, *cis*- acting variants in core genes may be subject to purifying selection.[19]

Bayesian hypothesis testing is based on the Likelihood Principle that all information conveyed by observations that supports one model or one parameter value over another is contained in the ratio of the likelihoods of these models or parameter values. The likelihood-based approach to causal inference described in this paper is aligned with the not yet universally-accepted principle that causal inference is just a special case of statistical inference, requiring attention when making assumptions about exchangeability between observations of exposure and outcome and predictions of the effect on the outcome of perturbing the exposure.[20] In this decision-theoretic framework there is no need to invoke counterfactuals[21] or to require causal effects to be expressed as marginal rather than as conditional effects. In this example, the likelihood of a model with a causal effect parameter and only a few nonzero pleiotropic effects is higher than the likelihood of a model with no causal effect that requires more nonzero pleiotropic effect parameters to adjust to the data to obtain the same fit. Thus even though the causal effect parameter is not identifiable because for any setting of the causal effect parameter we can find a setting of pleiotropic effect parameters that has the same fit to the data, a formal hypothesis test can be obtained. With horseshoe priors, which allow the distribution of the pleiotropic effects to be learned from the

data, this model comparison is performed "under the hood,"[22] without the computational inconvenience of having to average over all possible partitions of the variables into two disjoint sets (spike-and-slab) as in the contamination mixture model[23] or over different settings of the number of nonzero effect parameters.[24] A corollary of inference based on model comparison is that reliable inference of causality in the presence of pleiotropy requires a relatively large number of unlinked genetic instruments to allow learning the distribution of pleiotropic effects.

The statistical model assumes that the effects of the instruments on the exposure and their direct effects on the outcome are independent in magnitude and direction: this has been denoted the InSIDE (Instrument Strength Independent of Direct Effect) assumption[25] though this term has also been used for the less stringent assumption that "the magnitude of the pleiotropic effects are independent of the SNP-exposure associations."[26] With the procedure described here for constructing scalar instruments, the estimates for the coefficients of the instrument-exposure effects can take only positive values (a negative effect would simply correspond to flipping of the allele coding), and the direct effects are assumed to be distributed with zero mean. This assumption that the instrument-exposure effects and the direct effects are uncoupled in direction is termed "balanced pleiotropy". It is possible in principle to infer a causal effect if the instrument-exposure effects and the direct effects are assumed to be uncoupled in magnitude even if they are coupled in direction[25] but it is not obvious why this would be a biologically realistic assumption.

If the statistical model is extended to allow the instrument-exposure effects and the direct effects to be coupled in magnitude and direction, it is not possible to infer causality without other information about pleiotropy because models with and without a causal effect can fit the data equally well with the same number of adjustable parameters. It may be more useful to formulate and test possible hypotheses about mechanisms that could give rise to coupling of instrument-exposure effects with direct effects on the outcome, as this may identify a shared aetiological pathway. For adiponectin and type 2 diabetes, for instance, coupling could arise if the pQTLs for adiponectin are QTLs for a trait such as obesity that causes low adiponectin levels and increased risk of type 2 diabetes. In this situation other circulating proteins such as leptin that are biomarkers for obesity would show a similar pattern of apparently causal pQTL effects on type 2 diabetes. More generally, pQTLs that have pleiotropic effects should be detectable on a heat map of correlations between genotypic scores for multiple proteins or transcripts.

Guidelines for the design of Mendelian randomization studies have recently been updated.[2] Based on the work described here, some suggestions for revisions to these guidelines are listed below.

- Study designs that rely on using variants from a single genomic region, for instance using only *cis*-acting variants to study the effects of a gene transcript or gene product on a disease or trait, are likely to give misleading results and should be deprecated. Because *cis*- acting variants are likely to affect the function of the gene through mechanisms (such as altered splicing) other than the measured levels of the transcript or gene product, it is preferable to report the effects of a *cis*- QTL separately.

- Summary statistics for SNP-exposure associations should be obtained from the largest available studies, to maximize the number of QTLs detected and thus the number of genetic instruments that can be constructed.

- Unless pleiotropic effects can be excluded (for instance where the biological effect of the genetic variant is well understood), the method of choice for inference of the causal effect is to marginalize over the distribution of pleiotropic effects, which using summary statistics with modern tools for Bayesian computation takes only a few seconds. There is no need for a plethora of "outlier-robust" methods, as sensitivity to prior guesses about the distribution of pleiotropic effects can be examined directly. The model-fitting

algorithms can easily be integrated into existing platforms for Mendelian randomization analysis that use only summary statistics.

Declarations

Data and code availability

A `Stan` script for fitting the Bayesian model used in this paper, together with the summary-level data used in this paper, is available at https://github.com/molepi-precmed/mrhevo.

Acknowledgements

No specific funding was received for this work. This research has been conducted using the UK Biobank Resource under application number [23652]. The development of the GENOSCORES platform was supported by a Springboard Award (SBF006/1109) to AS from the Academy of Medical Sciences, supported in turn by the Wellcome Trust, the UK Government Department of Business, Energy and Industrial Strategy, the British Heart Foundation, and Diabetes UK. AI was supported by the Medical Research Council Cross Disciplinary Fellowship (XDF) Programme (MC_FE_00035).

Declaration of interests

The authors declare no competing interests.

Web resources

GWAS Catalog: https://www.ebi.ac.uk/gwas/

References

1. Hemani, G., Zheng, J., Elsworth, B., Wade, K.H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., et al. (2018). The MR-Base platform supports systematic causal inference across the human phenome. eLife *7*, e34408. 10.7554/eLife.34408.

2. Burgess, S., Davey Smith, G., Davies, N.M., Dudbridge, F., Gill, D., Glymour, M.M., Hartwig, F.P., Kutalik, Z., Holmes, M.V., Minelli, C., et al. (2023). Guidelines for performing Mendelian randomization investigations: Update for summer 2023. Wellcome Open Res *4*, 186. 10.12688/wellcomeopenres.15555.3.

3. Carvalho, C.M., Polson, N.G., and Scott, J.G. (2009). Handling Sparsity via the Horseshoe. In Proceedings of the Twelth International Conference on Artificial Intelligence and Statistics (PMLR), pp. 73–80.

4. Berzuini, C., Guo, H., Burgess, S., and Bernardinelli, L. (2020). A Bayesian approach to Mendelian randomization with multiple pleiotropic variants. Biostatistics *21*, 86–101. 10.1093/biostatistics/kxy027.

5. Grant, A.J., and Burgess, S. (2023). A Bayesian approach to Mendelian randomization using summary statistics in the univariable and multivariable settings with correlated pleiotropy. 2023.05.30.542988. 10.1101/2023.05.30.542988.

6. Piironen, J., and Vehtari, A. (2017). Sparsity information and regularization in the horseshoe and other shrinkage priors. Electronic Journal of Statistics *11*, 5018–5051.

7. Yaghootkar, H., Lamina, C., Scott, R.A., Dastani, Z., Hivert, M.-F., Warren, L.L., Stancáková, A., Buxbaum, S.G., Lyytikäinen, L.-P., Henneman, P., et al. (2013). Mendelian randomization studies do not support a causal role for reduced circulating adiponectin levels in insulin resistance and type 2 diabetes. Diabetes *62*, 3589–3598. 10.2337/db13-0128.

8. Chen, Z., Bai, Y., Long, X., Luo, Q., Wen, Z., Li, Y., Huang, S., Yan, Y., and Mo, Z. (2020). Effects of Adiponectin on T2DM and Glucose Homeostasis: A Mendelian Randomization Study. Diabetes Metab Syndr Obes *13*, 1771–1784. 10.2147/DMSO.S248352.

9. Nielsen, M.B., Çolak, Y., Benn, M., and Nordestgaard, B.G. (2021). Low Plasma Adiponectin in Risk of Type 2 Diabetes: Observational Analysis and One- and Two-Sample Mendelian Randomization Analyses in 756,219 Individuals. Diabetes *70*, 2694–2705. 10.2337/db21-0131.

10. Zhang, M., Chen, X., Zhu, Y., Yin, L., Quan, Z., Ou, Y., and He, B. (2022). Causal associations of circulating adiponectin with cardiometabolic diseases and osteoporotic fracture. Sci Rep *12*, 6689. 10.1038/s41598-022-10586-1.

11. Ferkingstad, E., Sulem, P., Atlason, B.A., Sveinbjornsson, G., Magnusson, M.I., Styrmisdottir, E.L., Gunnarsdottir, K., Helgason, A., Oddsson, A., Halldorsson, B.V., et al. (2021). Large-scale integration of the plasma proteome with genetics and disease. Nat Genet *53*, 1712–1721. 10.1038/s41588-021-00978-w.

12. Sun, B.B., Chiou, J., Traylor, M., Benner, C., Hsu, Y.-H., Richardson, T.G., Surendran, P., Mahajan, A., Robins, C., Vasquez-Grinnell, S.G., et al. (2023). Plasma proteomic associations with genetics and health in the UK Biobank. Nature *622*, 329–338. 10.1038/s41586-023-06592-6.

13. Iakovliev, A., McGurnaghan, S.J., Hayward, C., Colombo, M., Lipschutz, D., Spiliopoulou, A., Colhoun, H.M., and McKeigue, P.M. (2023). Genome-wide aggregated trans-effects on risk of type 1 diabetes: A test of the "omnigenic" sparse effector hypothesis of complex trait genetics. Am J Hum Genet *110*, 913–926. 10.1016/j.ajhg.2023.04.003.

14. Lee, B.P., Lloyd-Laney, H.O., Locke, J.M., McCulloch, L.J., Knight, B., Yaghootkar, H., Cory, G., Kos, K., Frayling, T.M., and Harries, L.W. (2016). Functional characterisation of ADIPOQ variants using individuals recruited by genotype. Mol Cell Endocrinol *428*, 49–57. 10.1016/j.mce.2016.03.020.

15. Cook, J.R., and Semple, R.K. (2010). Hypoadiponectinemia–cause or consequence of human "insulin resistance"? J Clin Endocrinol Metab *95*, 1544–1554. 10.1210/jc.2009-2286.

16. Yamauchi, T., Kamon, J., Waki, H., Imai, Y., Shimozawa, N., Hioki, K., Uchida, S., Ito, Y., Takakuwa, K., Matsui, J., et al. (2003). Globular adiponectin protected ob/ob mice from diabetes and ApoE-deficient mice from atherosclerosis. J Biol Chem *278*, 2461–2468. 10.1074/jbc.M209033200.

17. Ouwens, K.G., Jansen, R., Nivard, M.G., van Dongen, J., Frieser, M.J., Hottenga, J.-J., Arindrarto, W., Claringbould, A., van Iterson, M., Mei, H., et al. (2020). A characterization of cis- and trans-heritability of RNA-Seq-based gene expression. Eur J Hum Genet *28*, PMID: 32591531 PMCID: PMC7320010. 3255.

18. Schmidt, A.F., Finan, C., Gordillo-Marañón, M., Asselbergs, F.W., Freitag, D.F., Patel, R.S., Tyl, B., Chopade, S., Faraway, R., Zwierzyna, M., et al. (2020). Genetic drug target validation using Mendelian randomisation. Nat Commun *11*, 3255. 10.1038/s41467-020-16969-0.

19. Liu, X., Li, Y.I., and Pritchard, J.K. (2019). Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. Cell *177*, 1022–1034.e6. 10.1016/j.cell.2019.04.014.

20. Rohde, D. (2022). Causal Inference, is just Inference: A beautifully simple idea that not everyone accepts. In I (Still) Can't Believe It's Not Better! Workshop at NeurIPS 2021 (PMLR), pp. 75–79.

21. Dawid, A.P. (2001). Causal Inference without Counterfactuals. In Foundations of Bayesianism Applied Logic Series., D. Corfield and J. Williamson, eds. (Springer Netherlands), pp. 37–74. 10.1007/978-94-017-1586-7_3.

22. Carvalho, C.M., Polson, N.G., and Scott, J.G. (2010). The horseshoe estimator for sparse signals. Biometrika *97*, 465–480.

23. Burgess, S., Foley, C.N., Allara, E., Staley, J.R., and Howson, J.M.M. (2020). A robust and efficient method for Mendelian randomization with hundreds of genetic variants. Nat Commun *11*, 376. 10.1038/s41467-019-14156-4.

24. Xue, H., Shen, X., and Pan, W. (2021). Constrained maximum likelihood-based Mendelian randomization robust to both correlated and uncorrelated pleiotropic effects. Am J Hum Genet *108*, 1251–1269. 10.1016/j.ajhg.2021.05.014.

25. Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. Int J Epidemiol *44*, 512–525. 10.1093/ije/dyv080.

26. Bowden, J., Burgess, S., and Smith, G.D. (2017). Difficulties in Testing the Instrument Strength Independent of Direct Effect Assumption in Mendelian Randomization. JAMA Cardiology *2*, 929–930. 10.1001/jamacardio.2017.1572.

Figures

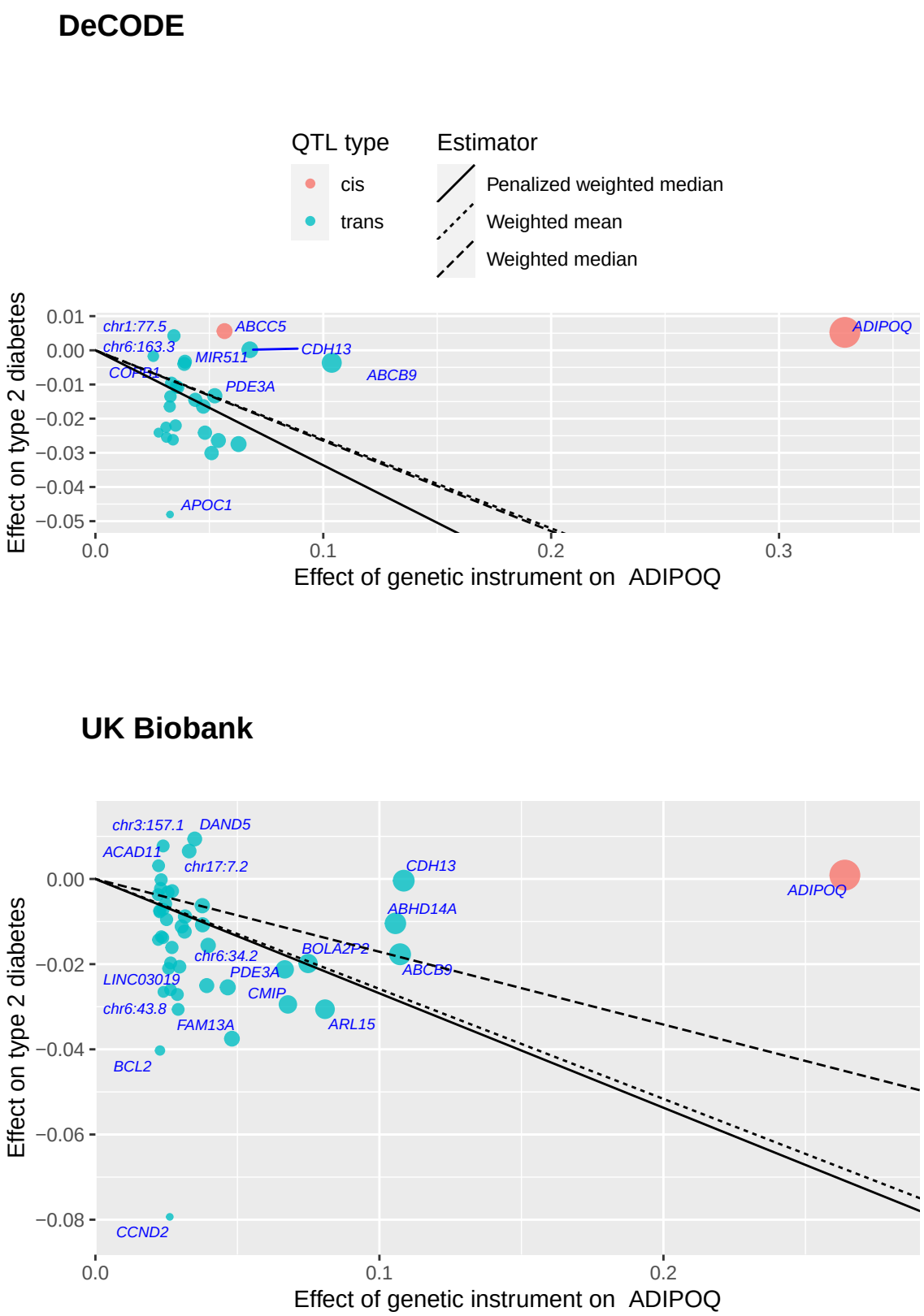

Fig 1. Plot of coefficients of regression of outcome on each instrument against coefficients of regression of exposure of each instrument. Size of each data point is inversely proportional to the standard error of the ratio estimate. The value of each estimator is shown as the slope of a line passing through the origin. Cis-QTLs are excluded from these estimates.

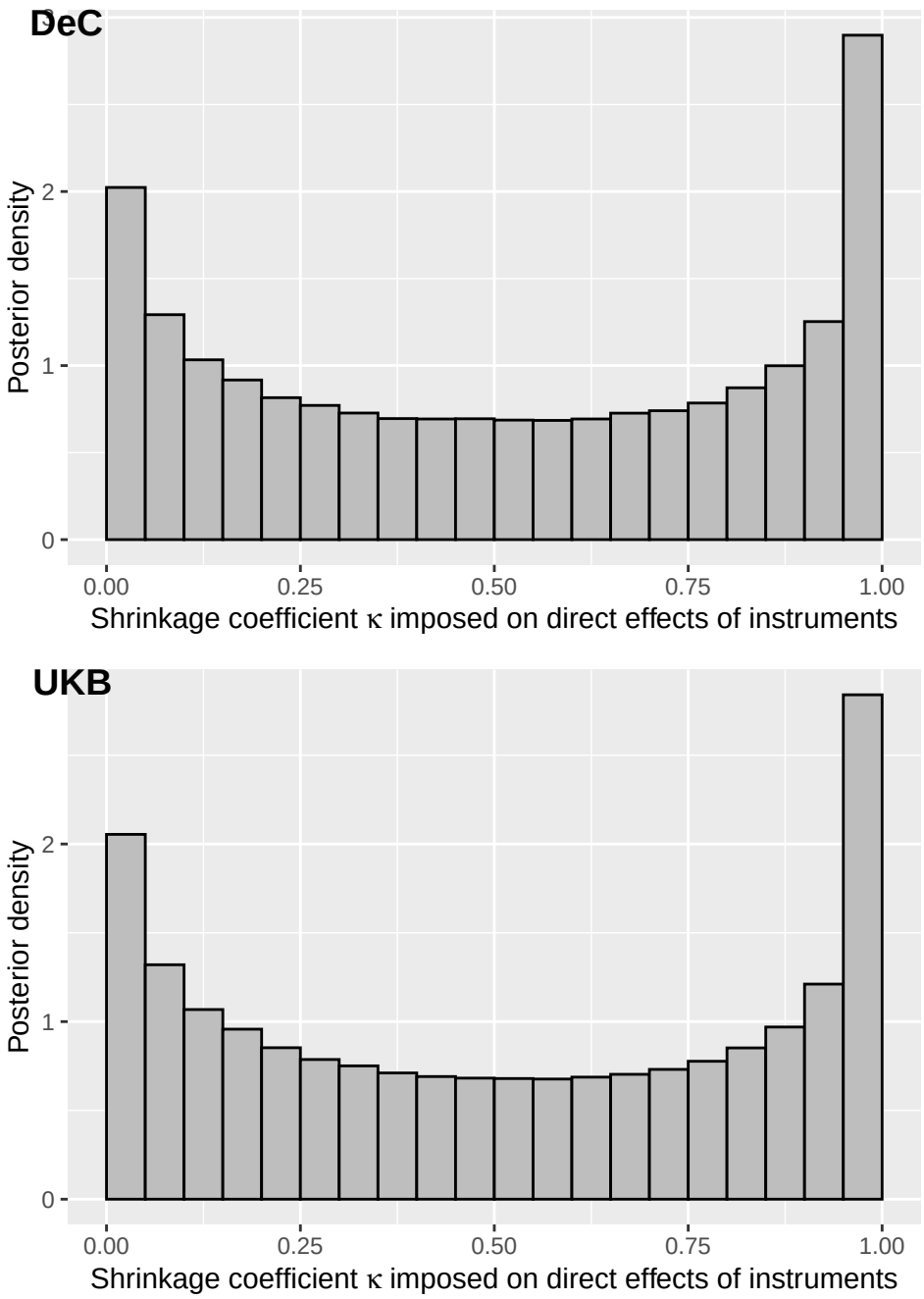

Fig 2. Posterior distribution of shrinkage coefficients κ imposed on pleiotropic (direct) effects of each instrument. A value of 0 represents no shrinkage, a value of 1 represents complete shrinkage

Tables

Table 1. Summary statistics for regression of outcome and exposure on genetic instruments

Study	Chr	Start (Mb)	End (Mb)	Nearby genes	$\hat{\gamma}$	SE($\hat{\gamma}$)	$\hat{\alpha}$	SE($\hat{\alpha}$)	$\hat{\theta}$
DeC	1	9.9	10.4	.	-0.011	0.005	0.0361	0.0061	-0.300
DeC	1	11.8	11.8	.	-0.014	0.005	0.0329	0.0061	-0.410
DeC	1	77.5	78.2	.	0.004	0.005	0.0345	0.0061	0.124
DeC	1	219.4	219.6	.	-0.030	0.005	0.0510	0.0061	-0.590
DeC	3	49.7	49.9	*AMIGO3, CDHR4, RNF123, TRAIP, UBA7*	-0.024	0.005	0.0277	0.0061	-0.870
DeC	3	52.2	53.2	*ALAS1, ALDOAP1, TWF2, TWF2-DT, WDR82*	-0.016	0.005	0.0474	0.0061	-0.347
DeC	3	183.6	184.5	*ABCC5, ABCC5-AS1, YEATS2, YEATS2-AS1, Y_RNA*	0.006	0.005	0.0567	0.0061	0.099
DeC	3	185.7	188.0	*ADIPOQ, ADIPOQ-AS1, ST6GAL1, TBCCD1, TRA2B*	0.005	0.005	0.3289	0.0058	0.016
DeC	4	88.8	88.8	*FAM13A*	-0.022	0.005	0.0310	0.0061	-0.726
DeC	5	53.8	54.1	*ARL15, ASS1P9, MIR581, RN7SL801P, RNU6-272P*	-0.024	0.005	0.0481	0.0061	-0.502
DeC	6	34.2	34.4	.	-0.022	0.005	0.0352	0.0061	-0.626
DeC	6	139.5	139.5	.	-0.025	0.005	0.0312	0.0061	-0.816
DeC	6	163.3	163.3	.	-0.002	0.005	0.0254	0.0061	-0.068
DeC	8	125.4	125.5	*TRIB1*	-0.014	0.005	0.0438	0.0061	-0.331
DeC	10	17.8	18.0	*MIR511, MRC1, SLC39A12*	-0.003	0.005	0.0393	0.0061	-0.085
DeC	10	114.0	114.1	*ADRB1*	-0.016	0.005	0.0326	0.0061	-0.504
DeC	11	14.1	14.9	*COPB1, CYP2R1,* ***PSMA1****,* ***RRAS2****,* ***SPON1****, SPON1-AS1*	-0.004	0.005	0.0390	0.0061	-0.102
DeC	12	20.2	20.4	*PDE3A, PDE3A-AS1*	-0.013	0.005	0.0524	0.0061	-0.254
DeC	12	106.7	106.9	*RFX4, RIC8B*	-0.011	0.005	0.0355	0.0061	-0.298
DeC	12	122.7	124.4	*ABCB9, ARL6IP4, VPS37B, Y_RNA, ZNF664*	-0.004	0.005	0.1037	0.0060	-0.035
DeC	16	81.5	81.5	*CMIP, PPIAP51*	-0.026	0.005	0.0540	0.0061	-0.489
DeC	16	82.5	83.0	*CDH13, MIR8058, RN7SL134P*	0.000	0.006	0.0678	0.0061	0.002
DeC	17	67.4	67.4	.	-0.010	0.005	0.0334	0.0061	-0.287
DeC	19	33.3	33.5	*AKR1B1P7, CEBPA, CEBPG, PEPD, RPS3AP50*	-0.027	0.005	0.0628	0.0061	-0.437
DeC	19	44.9	44.9	***APOC1****,* ***APOE****, NECTIN2,* ***TOMM40***	-0.048	0.005	0.0327	0.0061	-1.467
DeC	20	64.1	64.1	*LKAAEAR1, MIR6813, OPRL1, RGS19, TCEA2*	-0.026	0.005	0.0341	0.0061	-0.768
UKB	1	11.8	11.8	.	-0.007	0.006	0.0226	0.0045	-0.329
UKB	1	30.9	31.1	.	-0.009	0.006	0.0316	0.0045	-0.280
UKB	1	109.9	110.0	.	-0.012	0.006	0.0314	0.0045	-0.393
UKB	1	218.7	219.6	.	-0.025	0.006	0.0392	0.0045	-0.639
UKB	2	226.0	226.6		-0.037	0.006	0.0481	0.0045	-0.780
UKB	3	38.4	38.5	*ACVR2B, ACVR2B-AS1, RPL18AP7, XYLB, Y_RNA*	-0.003	0.006	0.0255	0.0045	-0.123
UKB	3	46.8	47.5	*BOLA2P2, CCDC12, SCAP, SETD2, SNORD13P3*	-0.020	0.006	0.0749	0.0045	-0.265
UKB	3	48.5	53.5	*ABHD14A, ABHD14A-ACY1, ZMYND10, ZMYND10-AS1, ZNF652P1*	-0.010	0.006	0.1056	0.0045	-0.099

UKB	3	132.5	132.7	*ACAD11, ACKR4, NPHP3, NPHP3-ACAD11, UBA5*	0.003	0.006	0.0223	0.0045	0.140
UKB	3	157.1	157.1	.	0.008	0.006	0.0238	0.0045	0.325
UKB	3	186.5	187.0	*ADIPOQ, ADIPOQ-AS1, SNORD2, ST6GAL1, TBCCD1*	0.001	0.006	0.2639	0.0043	0.003
UKB	4	55.4	55.7	*CLOCK, LINC02928, SRD5A3-AS1, TMEM165, Y_RNA*	-0.006	0.006	0.0247	0.0045	-0.237
UKB	4	88.7	88.8	*FAM13A, FAM13A-AS1, HERC3*	-0.031	0.006	0.0291	0.0045	-1.051
UKB	5	54.0	54.0	*ARL15*	-0.031	0.006	0.0808	0.0045	-0.378
UKB	5	158.5	158.6		-0.021	0.006	0.0258	0.0045	-0.816
UKB	6	34.2	34.4	.	-0.021	0.006	0.0296	0.0045	-0.696
UKB	6	43.8	43.8	.	-0.027	0.006	0.0240	0.0045	-1.106
UKB	6	139.5	139.5	.	-0.027	0.006	0.0288	0.0045	-0.940
UKB	6	153.1	153.2	.	-0.014	0.006	0.0238	0.0045	-0.581
UKB	6	163.7	163.7	.	-0.014	0.006	0.0221	0.0045	-0.645
UKB	7	17.8	18.0	*MRM3P2, SNX13*	-0.007	0.006	0.0238	0.0045	-0.301
UKB	7	77.5	78.0	*APTR, PHTF2, RSBN1L, TMEM60, Y_RNA*	-0.006	0.006	0.0376	0.0045	-0.167
UKB	7	150.5	150.7	*ALDH7A1P3, GIMAP4, GIMAP7, STRADBP1, TRPC6P3*	-0.010	0.006	0.0251	0.0045	-0.380
UKB	8	12.8	12.8	*LINC03019*	-0.026	0.006	0.0265	0.0045	-0.984
UKB	8	71.5	71.6	*EYA1*	-0.014	0.005	0.0232	0.0045	-0.583
UKB	9	129.1	129.2	*CRAT, DOLPP1, IER5L, IER5L-AS1, PTPA*	-0.011	0.006	0.0304	0.0045	-0.367
UKB	9	134.2	134.2	.	-0.004	0.006	0.0221	0.0045	-0.166
UKB	10	114.0	114.1	*ADRB1*	-0.016	0.006	0.0269	0.0045	-0.597
UKB	11	14.1	14.9	*COPB1, CYP2R1,* ***PSMA1, RRAS2, SPON1,*** *SPON1-AS1*	-0.016	0.006	0.0397	0.0045	-0.393
UKB	12	4.2	4.3	*CCND2, CCND2-AS1*	-0.079	0.007	0.0262	0.0045	-3.030
UKB	12	20.2	20.5	*PDE3A, PDE3A-AS1, UBE2L2*	-0.021	0.006	0.0667	0.0045	-0.318
UKB	12	56.5	57.3	*ATP5F1B, BAZ2A, STAT6, TAC3, ZBTB39*	-0.011	0.006	0.0377	0.0045	-0.285
UKB	12	121.9	124.2	*ABCB9, ARL6IP4, Y_RNA, ZCCHC8, ZNF664*	-0.018	0.006	0.1073	0.0045	-0.165
UKB	14	24.4	24.4	*NYNRIN*	0.000	0.006	0.0232	0.0045	-0.008
UKB	16	81.5	81.6	*CMIP, MIR7854, PPIAP51*	-0.029	0.006	0.0678	0.0045	-0.435
UKB	16	82.6	83.7	*CDH13, CDH13-AS1, MIR3182, MIR8058, RN7SL134P*	0.000	0.006	0.1086	0.0045	-0.004
UKB	17	7.2	7.3	.	0.007	0.006	0.0330	0.0045	0.198
UKB	17	47.9	48.0	.	-0.003	0.006	0.0270	0.0045	-0.105
UKB	18	63.2	63.2	*BCL2*	-0.040	0.006	0.0227	0.0045	-1.771
UKB	19	13.0	13.1	*DAND5, LYL1, NFIX*	0.009	0.006	0.0350	0.0045	0.268
UKB	19	33.3	33.6	*AKR1B1P7, CEBPA, CEBPG, PEPD, RPS3AP50*	-0.025	0.006	0.0466	0.0045	-0.546
UKB	20	7.3	7.4	*LINC01706, MIR8062*	-0.002	0.006	0.0229	0.0045	-0.096
UKB	20	52.6	52.6	*LINC01524*	-0.008	0.006	0.0227	0.0045	-0.341
UKB	20	64.1	64.1	*MIR6813, OPRL1, RGS19, TCEA2*	-0.020	0.006	0.0265	0.0045	-0.744

Genes in bold are those for which GWAS Catalog lists associations with type 2 diabetes detected at $p < 5 times 10^{-8}$.

Table 2. Comparison of parameter estimates obtained with different methods

QTL study	Estimator	Estimate	95% CI	pvalue
DeC	Weighted mean	-0.26	-0.31, -0.21	5×10^{-24}
DeC	Weighted median	-0.26	-0.37, -0.16	8×10^{-7}
DeC	Penalized weighted median	-0.34	-0.43, -0.24	6×10^{-12}
DeC	Marginalize over direct effects	-0.38	-0.5, -0.25	1×10^{-9}
UKB	Weighted mean	-0.26	-0.3, -0.22	1×10^{-36}
UKB	Weighted median	-0.17	-0.25, -0.1	9×10^{-6}
UKB	Penalized weighted median	-0.27	-0.35, -0.19	2×10^{-10}
UKB	Marginalize over direct effects	-0.33	-0.43, -0.23	4×10^{-10}

Supplementary information

Regularized horseshoe prior

The regression coefficients for the direct effects $\boldsymbol{\beta}$ are given a regularized horseshoe prior:[6]

$$\beta_j \sim \mathcal{N}\left(0, \tau^2 \tilde{\lambda}_j^2\right), \tilde{\lambda}_j^2 = \frac{\eta \lambda_j^2}{\eta + \tau^2 \lambda_j^2}$$

Half-Cauchy priors are specified on the unregularized local scale parameters λ_j and the global scale parameter τ.

$$\lambda_j \sim C^+(0, 1)$$

$$\tau \sim C^+(0, s_{\text{global}})$$

A weakly informative gamma prior is specified for the regularization parameter η:

$$\eta \sim \text{Gamma}\left(0.5\nu_{\text{slab}}, 0.5\frac{\nu_{\text{slab}}}{s_{\text{slab}}}\right)$$

For the jth instrument, the shrinkage coefficient κ_j is

$$\kappa_j = \frac{1}{1 + \tilde{\lambda}_j^2}$$

This takes values from 0 (no shrinkage) to 1 (complete shrinkage). The prior on the shrinkage coefficients has a horseshoe shape. The effective fraction f of nonzero coefficients is

$$f = \frac{1}{J}\sum_{j=1}^{J}(1 - \kappa_j)$$

The heavy tail of the half-Cauchy priors on λ_j allows some of the regression coefficients to escape the shrinkage imposed by small values of the global parameter τ. These nonzero coefficients are the slab component of the spike-and-slab distribution. Their scale is regularized by the parameter η. Even the largest coefficients will be regularized as a Gaussian with variance η.

The value of ν_{slab} controls the shape of the distribution of η. Piironen and Vehtari recommend setting $\nu_{\text{slab}} = 1$, but their applications are to models with high-dimensional predictors where there is enough information to learn reliably the value of η from the data. For Mendelian randomization analyses there is rather less information in the data about the distribution of nonzero pleiotropic effects, though we usually know in advance that genotype-outcome coefficients are unlikely to be large. To constrain η so that the sampling algorithm does not diverge, we set $\nu_{\text{slab}} = 2$. The scaling factor s_{slab} is specified based on prior information about the expected size of the largest direct effects.

Piironen and Vehtari[6] show that the prior expectations of the shrinkage coefficients κ_j and the effective fraction f of coefficients that are nonzero depend on $\tau\sqrt{\mathcal{I}_\beta}$, where $\mathcal{I}_\beta$ is the Fisher information on the effect parameters β_j. As sample size and information increase, a smaller value of τ is required to shrink small effects nearly to zero, rather than allowing them to escape into the slab component. To keep our prior expectation of f consistent, the scale of the prior on τ should vary inversely with $\sqrt{\mathcal{I}_\beta}$. To set the prior median on τ to the value that it has when the fraction of nonzero coefficients is equal to its prior expectation f, we specify the scale factor s_{global} as

$$s_{\text{global}} = \frac{f}{(1 - f)\sqrt{\mathcal{I}_\beta}}$$

For this analysis, we have used the information $\mathcal{I}_\gamma$ on the crude effects γ_j to set the prior median on τ based on a guess of f, as we cannot calculate in advance the information $\mathcal{I}_\beta$ on the unobserved pleiotropic effects β_j which will be somewhat less than $\mathcal{I}_\gamma$. As described in this paper, the sensitivity of the results to this prior setting can be examined.

Calculation of instrument-outcome coefficients from summary-level data

Where only summary-level data are available for the genotype-outcome associations, an estimate $\hat{\gamma}$ of the coefficient for the effect of the scalar instrument Z on the outcome and its standard error can be obtained using individual-level genotype data from a reference panel. First, the vector $\hat{\boldsymbol{\gamma}}_{\boldsymbol{m}}$ of estimated multivariable coefficients is calculated by premultiplying the estimates $\hat{\boldsymbol{\gamma}}_{\boldsymbol{u}}$ of the univariate coefficients by $\boldsymbol{\Sigma}_{\boldsymbol{G}}^{-\mathbf{1}}$. The coefficient $\hat{\gamma}$ for the scalar instrument Z is then estimated by minimizing the sum of the squared differences between the predictor calculated from Z and the linear predictor calculated from the vector of genotypes $\boldsymbol{G}$, where the sum is taken over the genotypes of all n individuals in the reference panel.

$$\sum_{i=1}^{n} \left(\gamma Z_i - \boldsymbol{G}_i \cdot \hat{\boldsymbol{\gamma}}_{\boldsymbol{m}}\right)^2$$

Equating to zero the derivative of this expression with respect to γ, substituting $\|\boldsymbol{G_i}\|\rho_{G,\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}}$ for Z_i and factoring the dot product $\boldsymbol{G}_i \cdot \hat{\boldsymbol{\gamma}}_{\boldsymbol{m}}$ as $\|\boldsymbol{G}_i\|\rho_{\boldsymbol{G}_i,\hat{\boldsymbol{\gamma}}_{\boldsymbol{m}}}\|\hat{\boldsymbol{\gamma}}_{\boldsymbol{m}}\|$ we obtain

$$\hat{\gamma} = \|\hat{\boldsymbol{\gamma}}_{\boldsymbol{m}}\| \frac{\sum_{i=1}^{n} \rho_{\boldsymbol{G}_i,\hat{\boldsymbol{\gamma}}_{\boldsymbol{m}}} \|\boldsymbol{G}_i\|^2 \rho_{G_i,\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}}}{\sum_{i=1}^{n} \|\boldsymbol{G_i}\|^2 \rho^2_{G_i,\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}}}$$

The ratio in this expression can be recognized as a weighted average of the ratio $\rho_{G_i,\hat{\boldsymbol{\gamma}}_{\boldsymbol{m}}}/\rho_{G_i,\hat{\boldsymbol{\alpha}}_{\boldsymbol{m}}}$, with weights Z_i^2. The standard error of $\hat{\gamma}$ is $1/\sqrt{\mathcal{I}_\gamma}$ where $\mathcal{I}_\gamma$ is the Fisher information on the slope of the regression of Y on Z, given (for a logistic regression model) by

$$\mathcal{I}_\gamma = N_\gamma \mathrm{Var}\left(Z\right) p\left(1-p\right)$$

where N_γ and p are respectively the total sample size and the the proportion of cases in the dataset from which the coefficient estimates $\hat{\boldsymbol{\gamma}}_{\boldsymbol{u}}$ were obtained.

Supplementary Figures

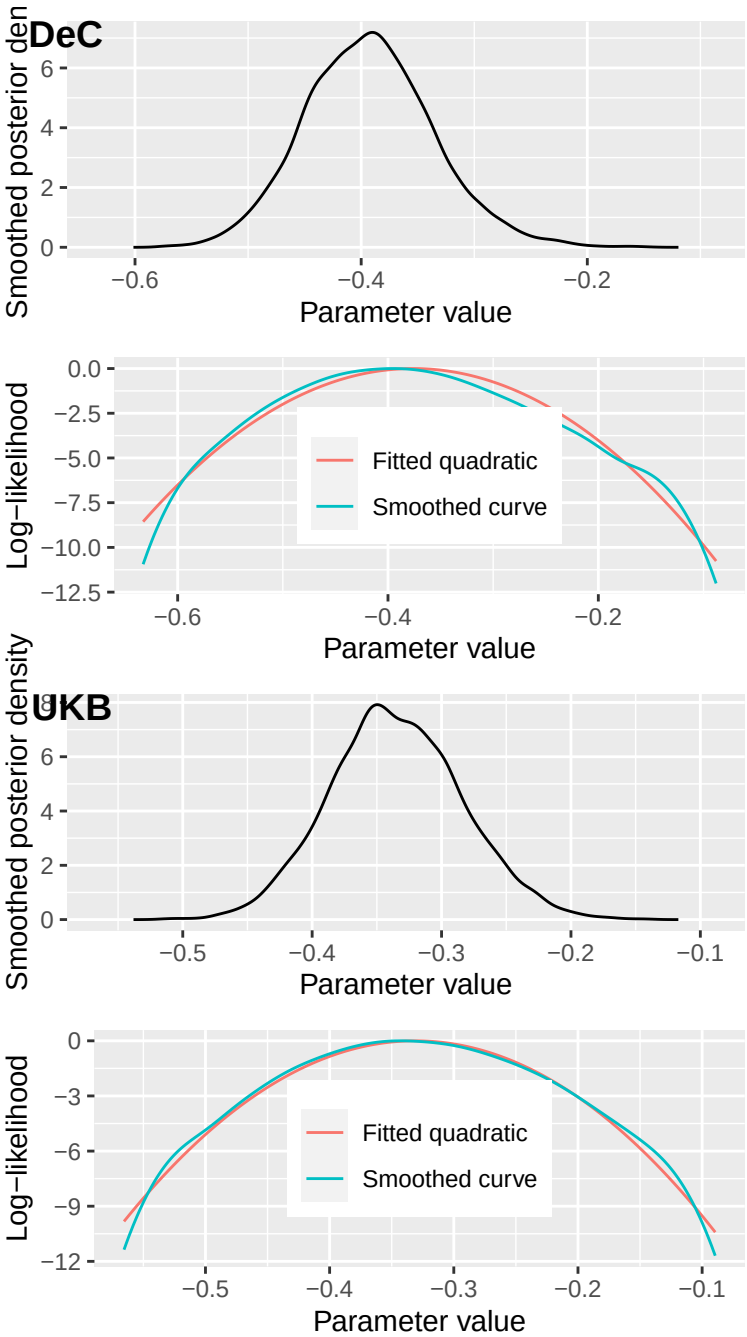

Fig S1. Posterior density and log-likelihood of causal effect parameter θ

Supplementary Tables

Table S1. Posterior summaries and sampling properties of model parameters

		Posterior distribution		
QTL study	Variable	Mean	10th centile	90th centile
DeC	θ	-0.39	-0.46	-0.32
DeC	$\log \eta$	-1.17	-1.84	-0.36
DeC	$\log \tau$	-4.92	-5.50	-4.35
DeC	f	0.48	0.39	0.58
DeC	chr1:9.9	0.00	0.00	0.01
DeC	chr1:11.8	0.00	0.00	0.00
DeC	chr1:77.5	0.01	0.00	0.02
DeC	chr1:219.4	-0.01	-0.01	0.00
DeC	*AMIGO3*	-0.01	-0.02	0.00
DeC	*ALAS1*	0.00	0.00	0.01
DeC	*FAM13A*	-0.01	-0.01	0.00
DeC	*ARL15*	0.00	-0.01	0.00
DeC	chr6:34.2	0.00	-0.01	0.00
DeC	chr6:139.5	-0.01	-0.02	0.00
DeC	chr6:163.3	0.00	0.00	0.01
DeC	*TRIB1*	0.00	0.00	0.01
DeC	*MIR511*	0.01	0.00	0.02
DeC	*ADRB1*	0.00	-0.01	0.00
DeC	*COPB1*	0.01	0.00	0.01
DeC	*PDE3A*	0.00	0.00	0.01
DeC	*RFX4*	0.00	0.00	0.01
DeC	*ABCB9*	0.04	0.02	0.05
DeC	*CMIP*	0.00	-0.01	0.00
DeC	*CDH13*	0.02	0.01	0.03
DeC	chr17:67.4	0.00	0.00	0.01
DeC	*AKR1B1P7*	0.00	-0.01	0.00
DeC	*APOC1*	-0.03	-0.04	-0.02
DeC	*LKAAEAR1*	-0.01	-0.02	0.00
UKB	θ	-0.34	-0.40	-0.27
UKB	$\log \eta$	-1.20	-1.88	-0.41
UKB	$\log \tau$	-4.87	-5.29	-4.46
UKB	f	0.49	0.42	0.56
UKB	chr1:11.8	0.00	0.00	0.00
UKB	chr1:30.9	0.00	0.00	0.01
UKB	chr1:109.9	0.00	-0.01	0.00
UKB	chr1:218.7	-0.01	-0.02	0.00
UKB		-0.02	-0.03	-0.01
UKB	*ACVR2B*	0.00	0.00	0.01
UKB	*BOLA2P2*	0.00	0.00	0.01
UKB	*ABHD14A*	0.02	0.01	0.03
UKB	*ACAD11*	0.01	0.00	0.01
UKB	chr3:157.1	0.01	0.00	0.02
UKB	*CLOCK*	0.00	0.00	0.01
UKB	*FAM13A*	-0.02	-0.03	-0.01
UKB	*ARL15*	0.00	-0.01	0.00

UKB		-0.01	-0.02	0.00
UKB	chr6:34.2	-0.01	-0.01	0.00
UKB	chr6:43.8	-0.01	-0.02	-0.01
UKB	chr6:139.5	-0.01	-0.02	0.00
UKB	chr6:153.1	0.00	-0.01	0.00
UKB	chr6:163.7	0.00	-0.01	0.00
UKB	*MRM3P2*	0.00	0.00	0.00
UKB	*APTR*	0.00	0.00	0.01
UKB	*ALDH7A1P3*	0.00	0.00	0.00
UKB	*LINC03019*	-0.01	-0.02	0.00
UKB	*EYA1*	0.00	-0.01	0.00
UKB	*CRAT*	0.00	-0.01	0.00
UKB	chr9:134.2	0.00	0.00	0.01
UKB	*ADRB1*	0.00	-0.01	0.00
UKB	*COPB1*	0.00	-0.01	0.00
UKB	*CCND2*	-0.07	-0.08	-0.06
UKB	*PDE3A*	0.00	0.00	0.01
UKB	*ATP5F1B*	0.00	0.00	0.01
UKB	*ABCB9*	0.01	0.00	0.03
UKB	*NYNRIN*	0.00	0.00	0.01
UKB	*CMIP*	0.00	-0.01	0.00
UKB	*CDH13*	0.03	0.02	0.04
UKB	chr17:7.2	0.01	0.00	0.02
UKB	chr17:47.9	0.00	0.00	0.01
UKB	*BCL2*	-0.03	-0.04	-0.02
UKB	*DAND5*	0.02	0.01	0.03
UKB	*AKR1B1P7*	-0.01	-0.01	0.00
UKB	*LINC01706*	0.00	0.00	0.01
UKB	*LINC01524*	0.00	0.00	0.00
UKB	*MIR6813*	-0.01	-0.01	0.00