Published in final edited form as:

Hum Genet. 2018 March ; 137(3): 247–255. doi:10.1007/s00439-018-1876-1.

Inferring Causal Relationships between Phenotypes Using Summary Statistics from Genome-wide Association Studies

Xiang-He Meng^{1,†}, Hui Shen^{2,†}, Xiang-Ding Chen¹, Hong-Mei Xiao³, and Hong-Wen Deng^{1,2,3,*}

¹Laboratory of Molecular and Statistical Genetics, College of Life Sciences, Hunan Normal University, Changsha, Hunan 410081, China

²Center of Bioinformatics and Genomics, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA 70112, USA

³Institute of Reproduction and Stem Cell Engineering, School of Basic Medical Science, Central South University, Changsha, 410013, China

Abstract

Objective—Genome-wide association studies (GWAS) have successfully identified numerous genetic variants associated with diverse complex phenotypes and diseases, and provided tremendous opportunities for further analyses using summary association statistics. Recently, Pickrell et al. developed a robust method for causal inference using independent putative causal SNPs. However, this method may fail to infer the causal relationship between two phenotypes when only a limited number of independent putative causal SNPs identified. Here, we extended Pickrell's method to make it more applicable for the general situations.

Method—We extended the causal inference method by replacing the putative causal SNPs with the lead SNPs (the set of the most significant SNPs in each independent locus) and tested the performance of our extended method by using both simulation and empirical data.

Results—Simulations suggested that when the same number of genetic variants are used, our extended method had similar distribution of test statistic under the null model as well as comparable power under the causal model compared with the original method by Pickrell et al. But in practice, our extended method would generally be more powerful because the number of independent lead SNPs was often larger than the number of independent putative causal SNPs. And including more SNPs on the other hand would not cause more false positives. By applying our extended method to summary statistics from GWAS for blood metabolites and femoral neck bone mineral density (FN-BMD), we successfully identified 10 blood metabolites that may causally influence FN-BMD.

*Corresponding author Hong-Wen Deng, Ph. D., Center of Bioinformatics and Genomics, Department of Global Biostatistics and Data Science, School of Public Health and Tropical Medicine, Tulane University, 1440 Canal Street, Suite 1610, New Orleans, LA 70112, Phone: 504-988-1310, Fax: 504-988-1706, hdeng2@tulane.edu.

[†]These authors contributed equally to this work.

Conflict of interest

The authors have no financial interests to disclose.

Conclusion—We extended a causal inference method for inferring putative causal relationship between two phenotypes using summary statistics from GWAS, and identified a number of potential causal metabolites for FN-BMD, which may provide novel insights into the pathophysiological mechanisms underlying osteoporosis.

Keywords

Causal relationship; Osteoporosis; BMD; Metabolomics

Introduction

Inferring the causal direction between correlated phenotypes is a pervasive issue in biology that simple regression and association analysis cannot answer. Randomized controlled trial is a scientific method which randomly allocated the study participants to one or another treatment (Chalmers et al. 1981), and is widely taken as the gold standard for establishing causal inference (Cartwright 2009). In the era of genome-wide association studies (GWAS) and high-throughput genomic technologies, genetic data are routinely available on large well-phenotyped studies (Davey Smith and Hemani 2014). Mendelian randomization uses genetic variants as instrumental variables to estimate and test for the causal effect of an exposure variable on an outcome (Ference et al. 2012). Mendelian randomization is widely used to investigate the causal relationship between two phenotypes (Burgess et al. 2016; Dalbeth et al. 2015). However, Mendelian randomization requires the strong assumption that the variant does not influence the two phenotypes via independent mechanisms (Ziegler et al. 2015). Genetic correlation, which is the proportion of variance of two phenotypes shared, is useful to provide etiological insights and estimate the degree of pleiotropy or causal overlap (Bulik-Sullivan et al. 2015). Pickrell et al. (Pickrell et al. 2016) recently developed a robust method to infer evidence of causal relationship between two phenotypes by using independent putative causal SNPs inferred from GWAS summary data. This method was similar like a bi-directional Mendelian randomization which is useful to distinguish between an exposure causing an outcome and an outcome causing a phenotype (Vimaleswaran et al. 2013; Welsh et al. 2010). Pickrell's method infers the causal relationship between the two phenotypes by distinguishing ρ_{XY} (the correlation between effect sizes of putative causal SNPs of phenotype X on phenotype X and phenotype Y) and ρ_{YX} (the correlation between effect sizes of putative causal SNPs of phenotype Y on phenotype Y and phenotype X). However, this method may fail to give causal reference when there were only a limited few independent putative causal SNPs (Pickrell et al. 2016). In Mendelian randomization studies, researchers used the uncorrelated lead SNPs to investigate whether the exposure has a causal effect on the outcome (Mokry et al. 2015; Zhu et al. 2016). Inspired by this, we extended Pickrell's causal inference method by using the independent lead SNPs (the set of the most significant SNPs in each independent locus) to replace the set of putative causal SNPs. Our extended method will be shown here to have comparable or better performance with the Pickrell's method (Pickrell et al. 2016).

Early studies suggested that metabolites may act as intermediary playing important roles in the development of phenotype variation, and genetic variants generally explained a greater proportion of phenotype variance for metabolites than what is generally observed for

complex phenotypes (Kettunen et al. 2012). This larger contribution of genetic variants will lead to an increased power for detecting the causal relationship between metabolites and complex phenotypes by using genetic variants as instrumental variables (Brion et al. 2013; Freeman et al. 2013). Here, we applied our extended method to infer the putative causal relationships between metabolites in blood and femoral neck bone mineral density (FN-BMD). Low FN-BMD is a major risk factor for osteoporosis, an increasingly serious public health problem worldwide (Cummings and Melton 2002; Harvey et al. 2010). Numerous metabolites have been found to be associated with bone metabolism (Lv et al. 2016), however, the causal relationship for many of these metabolites and BMD was not clear and will be assessed here using our extended method.

Materials and methods

Causal inference

Pickrell et al. (Pickrell et al. 2016) developed a robust method to infer whether there was a causal relationship between two phenotypes. The principle is as follows (Pickrell et al. 2016): assuming phenotype X is one of a large number of factors that causally influence phenotype Y, then all the significant variants associated with phenotype X should also have an effect on phenotype Y, and importantly, the effect sizes of these variants on X and Y should be correlated. For example, if one variant has a large effect on phenotype X, then the effect size of this variant on phenotype Y is also expected to be relatively large. On the other hand, the reverse may not be true, that is, for variants that influence phenotype Y, the effect size of these variants on X and Y are not expected to be correlated, since these variants may influence Y through many other mechanisms other than through phenotype X.

As defined by Pickrell et al., we assumed that there were N_X independent genetic variants significantly associated with phenotype X, and N_Y independent genetic variants associated with phenotype Y identified from GWAS. β_{XX} was defined as the vector of effect sizes on phenotype X for the N_X variants ascertained through the association study of X. β_{XY} was the vector of the N_X variants' effect sizes on phenotype Y. Similarly, β_{YY} was defined as the vector of effect sizes on phenotype Y for the N_Y variants ascertained through the association study of Y, and β_{YX} was the vector of the N_Y variants' effect sizes on phenotype X. ρ_X was defined as Spearman rank correlation between β_{XX} and β_{XY} . Then the rank correlation between β_{YY} and β_{YX} could be expressed as ρ_Y . After Fisher's Z-transformation: $\hat{Z}_X = \text{arctanh}(\rho_X)$, the sampling distributions of ρ_X and ρ_Y can be approximated as:

$\hat{Z}_X \sim N(Z_X, \frac{1}{N_X - 3})$, $\hat{Z}_Y \sim N(Z_Y, \frac{1}{N_Y - 3})$, where $N(Z_X, \frac{1}{N_X - 3})$ was the density of a normal distribution with mean Z_X and variance $\frac{1}{N_X - 3}$. The approximate likelihood for joint

probability of the two coefficients can be expressed as (Pickrell et al. 2016):

$L(Z_X, Z_Y, \hat{Z}_X, \hat{Z}_Y, N_X, N_Y) = N(\hat{Z}_X; Z_X, \frac{1}{N_X - 3})N(\hat{Z}_Y; Z_Y, \frac{1}{N_Y - 3})$, where $N(\hat{Z}_X; Z_X, \frac{1}{N_X - 3})$ was the density of a normal distribution with mean Z_X and variance $\frac{1}{N_X - 3}$ evaluated at $\hat{Z}_X$

As defined by Pickrell et al., there were four models to compare (Pickrell et al. 2016):

1. Model 1: If phenotype X causally affected phenotype Y, then ρ_Y has a normal distribution with mean zero, we let Z_Y be 0 and estimated Z_X
2. Model 2: If phenotype Y causally affected phenotype X, then ρ_X has a normal distribution with mean zero, we let Z_X be 0 and estimated Z_Y
3. Model 3: If there was no causal relationship between X and Y, then ρ_X and ρ_Y have normal distributions with mean zero, we let both Z_X and Z_Y be zero
4. Model 4: If the correlation did not depend on how the variants were ascertained, then ρ_X and ρ_Y have normal distributions with the same mean, we let Z_X be equal to Z_Y .

The first two models are causal models, while the last two models could be interpreted as non-causal models. Then each model was fitted by maximizing the corresponding approximate likelihood. To compare the fitting effects, the Akaike information criterion (AIC) was calculated for each model and the number of parameters for each model was 1, 1, 0 and 1, respectively. The smallest AIC of causal models (AIC_{causal}) was chosen from the first two models, and the smallest AIC of non-causal models ($AIC_{\text{non-causal}}$) was chosen from the last two models. A relative likelihood (r) of the best non-causal model compared to the best causal model was then calculated as: $r = \exp\left(\frac{AIC_{\text{causal}} - AIC_{\text{non-causal}}}{2}\right)$. As such, $r = 0.01$ could be interpreted as the relationship between these two phenotypes were 100 folds more likely to be causal rather than non-causal (Pickrell et al. 2016). A less stringent significance threshold of $r = 0.05$ was also used in this study.

To assess the type I error, we introduced a p value calculated by χ^2_1 test (O'Connor and Price 2017) to give the significance of the difference between Z_X and Z_Y :

$$\chi^2 = \frac{1}{\frac{1}{N_X - 3} + \frac{1}{N_Y - 3}} (\hat{Z}_X - \hat{Z}_Y)^2. \text{ The p values were adjusted by using the Benjamini-}$$

Hochberg method (Benjamini 1995) to control the false discovery rate (FDR) at level 0.05 if multiple tests were conducted.

For inference of putative causal SNPs, Pickrell et al. (Pickrell et al. 2016) split the genome into 1703 approximately independent blocks based on linkage disequilibrium (LD) patterns in the European populations (Berisa and Pickrell 2016). The 1703 independent blocks were separated by a heuristic approach for choosing segment boundaries. Two sets SNPs in different blocks were independent if the pairwise r^2 between SNPs in different sets was close to zero (Berisa and Pickrell 2016). The set of independent putative causal SNPs associated with the phenotype was used to represent N_X or N_Y . The inferred putative causal SNPs were those with the largest posterior probability of being causal from each block where the regional posterior probability of containing one causal SNP associated with the phenotype was larger than 0.9 (Pickrell et al. 2016).

In our extended method, we also used the 1703 approximately independent blocks and extracted the significant lead SNPs from each block (i.e., the most significant SNPs ranked by their p values) to represent N_X and N_Y in the causal inference analysis.

Simulation study

In order to simulate the summary statistics with no LD, we only draw effect sizes of lead SNPs from 1703 independent blocks. The effect sizes were drawn from a multivariate normal distribution with mean of zero, the correlation between phenotype X and Y of ρ and variance of σ_X^2 and σ_Y^2 for phenotype X and phenotype Y, respectively. To test type I error in the null situations, we let ρ equal to zero, which indicates there is no correlation between the two phenotypes. We assumed that those SNPs with relative large effect sizes in each phenotype were significant SNPs which were further used to give a causal inference. We also varied the number of significant SNPs in each simulation to test whether the type I error rate was inflated when more significant SNPs were used.

To evaluate the power of our extended method, we assumed that phenotype X caused phenotype Y with ρ_X at different levels. The effect sizes were drawn from a mixture of multivariate normal distribution for two phenotypes, where a component has ρ at different level (0.1, 0.3, 0.5, 0.7 and 0.9) and a component has ρ equal to zero. The first component was used to generate the effect sizes of those significant lead SNPs of trait X on both phenotypes. The second component was used to generate the effect sizes of other lead SNPs on either of the two study phenotypes. And this would make sure that the correlation between effect sizes of significant SNPs of phenotype Y on both phenotypes was random and closed to zero. Before the causal inference, we calculated the rank correlation ρ_X (the rank correlation between effect sizes of N_X significant lead SNPs on phenotype X and effect sizes of N_X significant lead SNPs on phenotype Y) to make sure that phenotype X caused phenotype Y with a certain ρ_X . The causal inference was conducted only when ρ_X was located in an interval (e.g., 0.25-0.35 when ρ was equal to 0.3). We also simulated that there were different number (N_X and/or N_Y) of significant lead SNPs. In each simulation, we calculated a p value. After 10,000 simulations, the power can then be calculated as the fraction of simulation results with a FDR less than 0.05.

GWAS data

We tested the performance, as an empirical validation, of our extended method by using empirical data (GWAS data for low-density lipoprotein cholesterol and coronary artery disease), and further applied it to infer the putative causal relationships between metabolites in blood and FN-BMD by using GWAS data for blood metabolites and FN-BMD. All the analyses were conducted in European-ancestry individuals.

GWAS data for low-density lipoprotein cholesterol (LDL-c)—The Global Lipids Genetics Consortium examined genetic associations with lipid levels (LDL-c, high-density lipoprotein cholesterol, triglycerides and total cholesterol) in 188,577 European-ancestry individuals using genome-wide and custom genotyping arrays (Global Lipids Genetics 2013). The summary statistics data for associations with LDL-c were downloaded from <http://www.sph.umich.edu/csg/abecasis/public/lipids2013>.

GWAS data for coronary artery disease (CAD)—The CARDIoGRAM Consortium performed a meta-analysis of 14 GWAS of CAD comprising 22,233 cases and 64,762 controls of European descent (Schunkert et al. 2011). The summary statistics data for this

study (Schunkert et al. 2011) were downloaded from the CARDIoGRAM consortium (<http://www.cardiogramplusc4d.org/data-downloads/>).

GWAS data for blood metabololites—A GWAS meta-analysis of two cohorts (TwinsUK and KORA F4) containing 7,824 European adult individuals was conducted to identify significant loci associated with metabolite concentrations. A total of 529 metabolites in serum or plasma were detected by using liquid-phase chromatography and gas chromatography separation coupled with tandem mass spectrometry. After stringent quality controls, a subset of 486 metabolites was included for genetic analysis in both cohorts, including 309 known and 177 unknown metabolites for which chemical identities were not yet determined (Shin et al. 2014). TwinsUK cohort was genotyped with Illumina arrays, and KORA F4 cohort was genotyped with Affymetrix GeneChip array 6.0. Subsequently, two cohorts were imputed using reference haplotypes from the HapMap 2 project. The details of data processing were described in previous studies (Illig et al. 2010; Kolz et al. 2009). SNPs with minor allele frequency (MAF) > 1% were included in the study. The full results including associations of the 486 metabolites were downloaded from the metabolomics GWAS server (<http://metabolomics.helmholtz-muenchen.de/gwas>).

GWAS data for FN-BMD—The GENetic Factors for OSteoporosis (GEFOS) Consortium conducted a meta-analysis for FN-BMD with 53,236 individuals of European ancestry from the general population (Zheng et al. 2015). Associations for FN-BMD were derived from whole-genome sequencing, whole-exome sequencing, deep imputation of genotyped samples using a combined UK10K/1000 Genomes reference panel, and *de-novo* replication genotyping (Zheng et al. 2015). The summary statistics data were downloaded from the GEFOS website (<http://www.gefos.org/>).

For all the above data, we only included SNPs with minor allele frequency (MAF) > 1% in the causal inference analysis.

Results

To test the performance of our extended causal inference method, we simulated two phenotypes X and Y and assumed different numbers of significant lead SNPs for each phenotype (i.e., $N_X = N_Y = 10, 20, 50$ and 100). In each simulation, we performed χ^2 test to assess the significance of the difference between Z_X and Z_Y and calculated the relative likelihood (r) of the best non-causal model to the best causal model. After 10,000 times of simulations under the non-causal model, we obtained the distribution of p values and r under the null situations (Fig. 1 and S1). The type I errors were very well controlled by our extended method, with no inflation of p values detected when over 10 lead SNPs were used in the causal inference (Fig. 1B-D). When a small number (i.e., 10) of genetic variants were used in the causal inference, the rank correlation of these variants' effect sizes on two phenotypes can vary significantly in the null situations. However, this case was not expected if there is a sufficiently large number of genetic variants used in the causal inference. And even when $N_X = N_Y = 10$ (Fig. 1A), we only observed 4 simulations out of 10,000 that gave a causal conclusion with a FDR < 0.05. It suggested that the thresholds of FDR < 0.05 was useful to control this kind of inflation due to statistical and sampling fluctuation. The

distribution of r was also similar to the simulation results of Pickrell's (Pickrell et al. 2016), for which only a small fraction of simulations produced $r < 0.01$ (Fig. S1). In addition, when we used more variants in the analysis, the fraction became smaller (Fig. S1).

We also used simulations to assess the power of our extended method at different significance levels (Fig. 2). We assumed that phenotype X caused phenotype Y with ρ_X at different levels. In each time, the effect sizes were drawn from a mixture of multivariate normal distribution for two phenotypes, where a component has ρ at different level (0.1, 0.3, 0.5, 0.7 and 0.9) and a component has ρ equal to zero, and the causal inference was conducted. After 10,000 simulations, the power was calculated as the fraction of simulation result with a $FDR < 0.05$. As expected, the power for causal inference was enhanced with the increasing number of lead SNPs used as input and the increased true rank correlation (ρ_X) in their effect sizes (Fig. 2). We also gave the power when it was calculated as the fraction of simulation result with $r < 0.01$ and 0.05 (Fig. S2). The results from power simulation (Fig. 2 and S2) suggested that in many situations $r < 0.05$ was a more conservative threshold than $FDR < 0.05$.

As LDL-c has been widely accepted as a causal risk factor for CAD (Ference et al. 2012), we next tested our extended method by inferring the causal relationship between LDL-c (phenotype X) and CAD (phenotype Y) at different significance levels. As expected, the rank correlation ρ_X increased along with the increase of the significance levels (Fig. 3A). When the significance level was more stringent than 1×10^{-5} , the relative likelihood r was consistently smaller than 0.05 (i.e., $\log_{10} r < -1.30$) (Fig. 3B). Therefore, we selected 1×10^{-5} as the significance level to be used in our extended method.

Next, we attempted to identify potential causal relationships between blood metabolites and FN-BMD. The association data of 486 metabolites used for causal inference were from a GWAS meta-analysis study which was conducted to identify significant loci associated with metabolite concentrations (Shin et al. 2014). For all the metabolites tested in the metabolomics GWAS (Shin et al. 2014), the mean numbers of independent putative causal SNPs were less than 5. Therefore, our extended method was more applicable than Pickrell's method to infer causal relationships between blood metabolites and FN-BMD. By using all the independent lead SNPs with p values $< 1 \times 10^{-5}$ (3-209 lead SNPs for different metabolites with a mean value of 19), we identified four metabolites (nonadecanoate, 1-stearoylglycerophosphoinositol, 1-palmitoleoylglycerophosphocholine and leucylleucine) causally (relative likelihood $r < 0.01$) influenced FN-BMD (Table 1 and Fig. 4). With a less stringent threshold of $r < 0.05$, we identified additional 6 metabolites (aspartylphenylalanine, ascorbate and four metabolites whose chemical identity were not yet determined) which may have causal effects on FN-BMD (Table 1 and Fig. S3). On the other hand, we did not identify any metabolite that may be causally influenced by FN-BMD (data not shown).

Discussion

In this study, we extended Pickrell's causal inference method by using the set of independent lead SNPs to replace the set of independent putative causal SNPs and further applied it to infer the putative causal relationship between metabolites in blood and FN-BMD.

Simulations suggested that when the same number of genetic variants were used, our extended method had similar distribution of test statistic under the null model as well as comparable power under the causal model compared to Pickrell's method (Pickrell et al. 2016). This was not expected, since the putative causal SNPs used by Pickrell's method were also likely to be the lead SNPs. For example, all the putative causal SNPs identified in the LDL-c data were also the lead SNPs. However, with the stringent criteria (the block posterior probability of containing one causal SNP associated with the phenotype > 0.9) (Pickrell et al. 2016), putative causal SNPs may not be inferred from some significant or suggestive significant blocks. Therefore, the number of independent lead SNPs that can be identified in a GWAS is expected to be larger than the number of independent putative causal SNPs that can be inferred, which leads to higher statistical power of our extended method than Pickrell's method in most simulated situations. If there was no causal relationship between phenotype X and phenotype Y, including more non-significant SNPs gave a smaller ρ_X (Fig. 3), and a smaller ρ_X was more unlikely to give a causal conclusion (Fig. 2). It hence suggested that including more non-significant SNPs on the other hand would not cause more false positives. The higher statistical power of our extended method is also demonstrated in the causal inference analysis of metabolites and FN-BMD. Specifically, Pickrell's method failed to identify any significant causal metabolite for FN-BMD (Table 1), whereas our extended method successfully identified 10 metabolites that may causally influence FN-BMD.

Several studies have attempted to improve Mendelian randomization. For example, the generalized summary data based Mendelian randomization analysis (GSMR) allows using the correlated genetic variants (Zhu et al. 2017). The Mendelian Randomization Pleiotropy RESidual Sum and Outlier (MR-PRESSO) test can correct for distortion in the causal estimate in most cases by removal of offending instrumental variables that exhibit pleiotropy (Verbanck et al. 2017). The latent causal variable (LCV) model can quantify the degree of partial genetic causality (O'Connor and Price 2017). The causal inference test (CIT) would be more likely to give an unbiased result even when measurement error existed (Hemani et al. 2017). Compared with these methods, our extended method allowed using more genetic variants while maintaining well-controlled type I error rate. Previous studies (Brion et al. 2013; Freeman et al. 2013) have shown that when the genetic variants that are used to estimate the causal effects collectively explain a larger proportions of the variance in the causal phenotype (the exposure in Mendelian randomization studies), this will lead to more precise estimates of causal effects and thus increase the power for detecting the causal relationships. And if a small number of genetic variants have pleiotropic effects on the phenotypes, it would cause little change of ρ_{XY} (the rank correlation between the effect sizes of significant phenotype X SNPs on phenotype X and phenotype Y), so the results would be robust (data not shown).

Metabolomics is an emerging "omics" science that can identify low molecular weight metabolites from tissues, cells or biological fluids (Dettmer K 2007; Wilson et al. 2005). Metabolites in blood play important roles in numerous biological pathways and reflect the dynamic biological changes closest to the phenotype (Kettunen et al. 2012; Smilde Ak Fau - Westerhuis et al. 2010). Therefore, identifying the causal relationships between metabolites and phenotypes will be helpful in fully understanding the mechanism of diseases

development. Osteoporosis is a major skeletal disorder characterized by low BMD and an increased susceptibility to fractures (Cummings and Melton 2002). Hip was one of the most prevalent osteoporotic fracturesites, and FN-BMD has the highest predictive value for hip fracture (Cartwright 2009). In this study, we identified 10 blood metabolites that may causally influence FN-BMD. For example, genetic variants that influenced ascorbate, which was also known as Vitamin C, had correlated effect on FN-BMD (Table 1 and Fig. S3D). The positive correlation coefficient ($\rho_X = 0.66$) can be roughly interpreted that the increased concentration of ascorbate caused the increased FN-BMD. This was consistent with previous finding that high plasma levels or dietary intake of antioxidants is associated with increased BMD (Maggio et al. 2003; Sugiura et al. 2011). In addition, ascorbate is a cofactor in collagen formation and hydroxylation of lysine and proline (Peterkofsky et al. 1994), which is important for bone maintenance (Bonucci 2012). Another example was aspartylphenylalanine, which is a dipeptide composed of aspartate and phenylalanine. Aspartylphenylalanine can be major decomposition products of aspartame degraded from aspartame via intramolecular cyclization and liberation of methanol (Lin SY 2000). Aspartame, given orally, improves the bone cortical density and muscle mass in mouse (Manion et al. 2011). This was consistent with our finding (Fig. S3A) that aspartylphenylalanine probably helped the increase of FN-BMD ($\rho_X = 0.86$). However, little was known about the functions of the top four metabolites (Table 1) and their roles on bone metabolism warrants further functional analysis. On the other hand, our analysis did not detect several metabolites that were previously known to be causally associated with BMD, such as vitamin K (Palermo et al. 2017). This is mainly because that some of these metabolites (e.g., vitamin K) were not included in the metabolomics GWAS dataset (Shin et al. 2014).

Conclusions

Our extended method was applicable to give a causal estimation between phenotypes even when there were few independent putative causal SNPs identified. We identified a list of metabolites that may causally influence FN-BMD, which provided novel insights in the mechanism of osteoporosis development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank all the study subjects for volunteering to participate in the study. We thank Genetic Factors for Osteoporosis Consortium provided the BMD replication dataset to download for us. Also we thank Shin SY et al. to provide metabolomics GWAS results for download.

Funding

This study was supported in part by Natural Science Foundation of China (NSFC; 81570807, 30900810, 31271344 and 31071097), Hunan Provincial Construct Program of the Key Discipline in Ecology (0713) and the Cooperative Innovation Center of Engineering and New Products for Developmental Biology of Hunan Province (20134486). The investigators of this work were also partially supported by grants from the NIH (R01 AR069055, U19 AG055373, R01 MH104680, R01 AR059781 and P20 GM109036), and the Edward G. Schlieder Endowment as well as the Drs. W. C. Tsai and P. T. Kung Professorship in Biostatistics from Tulane University.

References

- Benjamini Y, Yosef H. 1995; the false discovery rate_ a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B.* 57(1):289–300.
- Berisa T, Pickrell JK. 2016; Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics.* 32:283–5. DOI: 10.1093/bioinformatics/btv546 [PubMed: 26395773]
- Bonucci E. 2012; Bone mineralization. *Front Biosci (Landmark Ed).* 17:100–28. [PubMed: 22201735]
- Brion MJ, Shakhbazov K, Visscher PM. 2013; Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol.* 42:1497–501. DOI: 10.1093/ije/dyt179 [PubMed: 24159078]
- Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. 2015; An atlas of genetic correlations across human diseases and traits. *Nat Genet.* 47:1236–41. DOI: 10.1038/ng.3406 [PubMed: 26414676]
- Burgess S, Dudbridge F, Thompson SG. 2016; Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Stat Med.* 35:1880–906. DOI: 10.1002/sim.6835 [PubMed: 26661904]
- Cartwright N. 2009; What are randomised controlled trials good for? *Philosophical Studies.* 147:59–70. DOI: 10.1007/s11098-009-9450-2
- Chalmers TC, Smith H Jr, Blackburn B, Silverman B, Schroeder B, Reitman D, et al. 1981; A method for assessing the quality of a randomized control trial. *Control Clin Trials.* 2:31–49. [PubMed: 7261638]
- Cummings SR, Melton LJ. 2002; Epidemiology and outcomes of osteoporotic fractures. *Lancet.* 359(9319):1761–7. [PubMed: 12049882]
- Dalbeth N, Topless R, Flynn T, Cadzow M, Bolland MJ, Merriman TR. 2015; Mendelian randomization analysis to examine for a causal effect of urate on bone mineral density. *J Bone Miner Res.* 30:985–91. DOI: 10.1002/jbmr.2434 [PubMed: 25502344]
- Davey Smith G, Hemani G. 2014; Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet.* 23:R89–98. DOI: 10.1093/hmg/ddu328 [PubMed: 25064373]
- Dettmer KAP, Hammock BD. 2007; Mass spectrometry-based metabolomics, *Mass Spectrom. Mass Spectrom Rev.* 26(1):51–78. [PubMed: 16921475]
- Ference BA, Yoo W, Alesh I, Mahajan N, Mirowska KK, Mewada A, Kahn J, et al. 2012; Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. *J Am Coll Cardiol.* 60(25):2631–9. [PubMed: 23083789]
- Freeman G, Cowling BJ, Schooling CM. 2013; Power and sample size calculations for Mendelian randomization studies using one genetic instrument. *Int J Epidemiol.* 42:1157–63. DOI: 10.1093/ije/dyt110 [PubMed: 23934314]
- Global Lipids Genetics C. 2013; Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 45:1274–1283. DOI: 10.1038/ng.2797 [PubMed: 24097068]
- Harvey N, Dennison E, Cooper C. 2010; Osteoporosis: impact on health and economics. *Nat Rev Rheumatol.* 6(2):99–105. [PubMed: 20125177]
- Hemani G, Tilling K, Davey Smith G. 2017; Orienting the causal relationship between imprecisely measured traits using GWAS summary data. 13:e1007081.doi: 10.1371/journal.pgen.1007081
- Illig T, Gieger C, Zhai G, Romisch-Margl W, Wang-Sattler R, Prehn C, et al. 2010; A genome-wide perspective of genetic variation in human metabolism. *Nat Genet.* 42:137–41. DOI: 10.1038/ng.507 [PubMed: 20037589]
- Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A, Tikkanen E. 2012; Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet.* 44:269–76. DOI: 10.1038/ng.1073 [PubMed: 22286219]
- Kolz M, Johnson T, Sanna S, Teumer A, Vitart V, Perola M, et al. 2009; Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. *PLoS Genet.* 5:e1000504.doi: 10.1371/journal.pgen.1000504 [PubMed: 19503597]

- Lin SY, Lin CY. 2000; Simultaneous formation and detection of the reaction product of solid-state aspartame sweetener by FT-IRDSC microscopic system. *Food Addit Contam.* 17(10):821–7. [PubMed: 11103265]
- Lv H, Jiang F, Guan D, Lu C, Guo B, Chan C, et al. 2016; Metabolomics and Its Application in the Development of Discovering Biomarkers for Osteoporosis Research. *Int J Mol Sci.* 17doi: 10.3390/ijms17122018
- Maggio D, Barabani M, Pierandrei M, Polidori MC, Catani M, Mecocci P, et al. 2003; Marked decrease in plasma antioxidants in aged osteoporotic women: results of a cross-sectional study. *J Clin Endocrinol Metab.* 88:1523–7. DOI: 10.1210/jc.2002-021496 [PubMed: 12679433]
- Manion CV, Hochgeschwender U, Edmundson AB, Hugli TE, Gabaglia CR. 2011; Dietary aspartyl-phenylalanine-1-methyl ester delays osteoarthritis and prevents associated bone loss in STR/ORT mice. *Rheumatology (Oxford).* 50:1244–9. DOI: 10.1093/rheumatology/ker089 [PubMed: 21372000]
- Mokry LE, Ross S, Ahmad OS, Forgetta V, Smith GD, Goltzman D, et al. 2015; Vitamin D and Risk of Multiple Sclerosis: A Mendelian Randomization Study. *PLoS Med.* 12:e1001866.doi: 10.1371/journal.pmed.1001866 [PubMed: 26305103]
- O'Connor LJ, Price AL. 2017; Distinguishing genetic correlation from causation across 52 diseases and complex traits. *bioRxiv.* doi: 10.1101/205435
- Palermo A, Tuccinardi D, D'Onofrio L, Watanabe M, Maggi D, Maurizi AR, et al. 2017; Vitamin K and osteoporosis: Myth or reality? *Metabolism.* 70:57–71. DOI: 10.1016/j.metabol.2017.01.032 [PubMed: 28403946]
- Peterkofsky B, Gosiewska A, Kipp DE, Shah V, Wilson S. 1994; Circulating insulin-like growth factor binding proteins (IGFBPs) 1 and 2 induced in vitamin C-deficient or fasted guinea pigs inhibit IGF-I action in cultured cells. *Growth Factors.* 10(4):229–41. [PubMed: 7528515]
- Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY, Hinds DA. 2016; Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet.* 48:709–17. DOI: 10.1038/ng.3570 [PubMed: 27182965]
- Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al. 2011; Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet.* 43:333–8. DOI: 10.1038/ng.784 [PubMed: 21378990]
- Shin SY, Fauman EB, Petersen AK, Krumsiek J. 2014; An atlas of genetic influences on human blood metabolites. *Nat Genet.* 46:543–50. DOI: 10.1038/ng.2982 [PubMed: 24816252]
- Smilde AK, Westerhuis JA, Hoefsloot HC, Bijlsma S, Rubingh CM, Vis DJ, Jellema RH, et al. 2010; Dynamic metabolomic data analysis: a tutorial review. *Metabolomics.* 6(1):3–17. [PubMed: 20339444]
- Sugiura M, Nakamura M, Ogawa K, Ikoma Y, Ando F, Shimokata H, Yano M, et al. 2011; Dietary patterns of antioxidant vitamin and carotenoid intake associated with bone mineral density: findings from post-menopausal Japanese female subjects. *Osteoporos Int.* 22(1):143–52. [PubMed: 20480147]
- Verbanck M, Chen C-Y, Neale B, Do R. 2017; Widespread pleiotropy confounds causal relationships between complex traits and diseases inferred from Mendelian randomization. *bioRxiv.* doi: 10.1101/157552
- Vimaleswaran KS, Berry DJ, Lu C, Tikkanen E, Pilz S, Hiraki LT, et al. 2013; Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts. *PLoS Med.* 10:e1001383.doi: 10.1371/journal.pmed.1001383 [PubMed: 23393431]
- Welsh P, Polisecki E, Robertson M, Jahn S, Buckley BM, de Craen AJ, et al. 2010; Unraveling the directional link between adiposity and inflammation: a bidirectional Mendelian randomization approach. *J Clin Endocrinol Metab.* 95:93–9. DOI: 10.1210/jc.2009-1064 [PubMed: 19906786]
- Wilson ID, Plumb R, Granger J, Major H, Williams R, Lenz EM. 2005; HPLC-MS-based methods for the study of metabonomics. *J Chromatogr B Analyt Technol Biomed Life Sci.* 817(1):67–76.
- Zheng HF, Forgetta V, Hsu YH, Estrada K, Rosello-Diez A, Leo PJ, et al. 2015; Whole-genome sequencing identifies EN1 as a determinant of bone density and fracture. *Nature.* 526:112–7. DOI: 10.1038/nature14878 [PubMed: 26367794]

- Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. 2016; Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet.* 48:481–7. DOI: 10.1038/ng.3538 [PubMed: 27019110]
- Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R, et al. 2017; Causal associations between risk factors and common diseases inferred from GWAS summary data. *bioRxiv*. doi: 10.1101/168674
- Ziegler A, Mwambi H, König IR. 2015; Mendelian Randomization versus Path Models: Making Causal Inferences in Genetic Epidemiology. *Hum Hered.* 79:194–204. DOI: 10.1159/000381338 [PubMed: 26201704]

Fig. 1. Distribution of p value in null situations. The effect sizes were drawn from a multivariate normal distribution with mean of zero, the correlation between phenotype X and Y of ρ and variance of σ_X^2 and σ_Y^2 for phenotype X and phenotype Y, respectively. To test type I error, we set no correlation between the two phenotypes ($\rho = 0$). In each figure, N_X represented the number of independent lead SNPs that influenced phenotype X, and N_Y represented the number of independent lead SNPs that influenced phenotype Y. Each distribution was obtained based on 10,000 simulations.

Fig. 2.

Power to detect a causal relationship. We simulated data under a model that phenotype X caused phenotype Y with rank correlation (ρ_X) at different levels. The effect sizes were drawn from a mixture of multivariate normal distribution, where a component has ρ at different level (0.1, 0.3, 0.5, 0.7 and 0.9) and a component has ρ equal to zero. In each simulation, we assumed the number of independent lead SNPs was 10, 20, 50 or 100, and calculated the fraction of the results with $FDR < 0.05$.

Fig. 3. Inferring the causal relationship between LDL-c (phenotype X) and CAD (phenotype Y) at different significance levels. In each significance level, lead SNPs, whose p values were less than the significant level, were extracted from two traits to estimate the rank correlation ρ_X (A) and relative likelihood (B).

Fig. 4.

Putative causal relationships ($r < 0.01$) between FN-BMD and 4 blood metabolites, including A) nonadecanoate (M01356), B) 1-stearoylglycerophosphoinositol (M19324), C) 1-palmitoleoylglycerophosphocholine (M33230) and D) leucylleucine (M36756). Each panel shows the genetic variants associated with phenotype X – the specific metabolite (left plot) and FN-BMD (right plot), and their effect sizes on phenotype X (x axis) and phenotype Y (y axis). Each line represented one standard error.

Table 1

10 putative causal metabolites for FN-BMD

ID*	Metabolite name	N _x	ρ _x	r	p value	FDR	r [#]
M01356	nonadecanoate	16	-0.88	6.39×10 ⁻⁶	2.24×10 ⁻⁶	0.001	NULL
M19324	1-stearoylglycerophosphoinositol	7	-0.89	3.89×10 ⁻³	4.46×10 ⁻³	0.36	NULL
M33230	1-palmitoleoylglycerophosphocholine	14	0.72	8.81×10 ⁻³	3.30×10 ⁻³	0.36	NULL
M36756	leucylleucine	9	-0.83	9.40×10 ⁻³	8.59×10 ⁻³	0.39	NULL
M22175	aspartylphenylalanine	7	0.86	0.011	1.23×10 ⁻²	0.42	NULL
M32755	NA	24	0.57	0.017	5.38×10 ⁻³	0.36	NULL
M34499	NA	8	-0.79	0.031	8.54×10 ⁻³	0.39	NULL
M01640	ascorbate	14	0.66	0.035	1.30×10 ⁻²	0.42	NULL
M19364	NA	20	-0.57	0.044	2.38×10 ⁻⁴	0.054	NULL
M36134	NA	14	0.73	0.045	1.57×10 ⁻²	0.46	NULL

* ID provided by Metabolon Inc.

N_x: Number of independent lead SNPs associated with the concentration of each metabolite.

ρ_x: rank correlation coefficient between the effect sizes of these lead SNPs on metabolite and the effect sizes of these lead SNPs on FN-BMD.

r: relative likelihood calculated by the extended method, for example $r = 0.01$ could be interpreted as the relationship between these two phenotypes were 100 folds more likely to be causal rather than non-causal.

[#]: relative likelihood calculated by Pickrell's method.

NA: chemical identity not yet determined.

p value: calculated from χ^2 test.

FDR: calculated using Benjamini-Hochberg method.

NULL: failed to give a r value due to the number of putative causal SNPs less than 2.