



ASSOCIATION STUDIES ARTICLE

Serum calcium and risk of migraine: a Mendelian randomization study

Peter Yin¹, Verner Anttila^{2,3,4}, Katherine M. Siewert⁵, Aarno Palotie^{2,3,4,6,7,8}, George Davey Smith⁹ and Benjamin F. Voight^{10,11,12,*}

¹Department of Biology, College of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, USA, ²Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA, ³Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA, ⁴Analytic and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA, ⁵Genomics and Computational Biology Program, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA, ⁶Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA, ⁷Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland, ⁸Department of Neurology, Massachusetts General Hospital, Boston, MA, USA, ⁹Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK, ¹⁰Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA, ¹¹Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA and ¹²Institute for Translational Medicine and Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

*To whom correspondence should be addressed at: University of Pennsylvania - Perelman School of Medicine, 3400 Civic Center Boulevard, 10-126 Smilow Center for Translational Research, Philadelphia, PA 19104, USA. Tel: +215-746-8083; Fax: +215-573-9135; Email: bvoight@upenn.edu

Abstract

Migraine affects ~14% of the world's population, though not all predisposing causal risk factors are known. We used electronic health records, genetic co-heritability analysis, and a two-sample Mendelian Randomization (MR) design to determine if elevated serum calcium levels were associated with risk of migraine headache. Co-morbidity was evaluated using electronic health records obtained from the PennOmics database comprising >1 million patient entries. Genetic co-heritability and causality via MR was assessed using data from the International Headache Consortium (23,285 cases, 95,425 controls) and circulating serum calcium levels (39,400 subjects). We observed co-occurrence of migraine and hypercalcaemia ICD-9 diagnoses (OR = 1.58, $P = 4 \times 10^{-13}$), even after inclusion of additional risk factors for migraine (OR = 1.23, $P = 2 \times 10^{-3}$). Second, we observed co-heritability ($r_g = 0.191$, $P = 0.03$) between serum calcium and migraine headache, indicating that these traits have a genetic basis in common. Finally, we found that elevation of serum calcium levels by 1 mg/dl resulting from our genetic score was associated with an increase in risk of migraine (OR = 1.80, 95% CI: 1.31–2.46, $P = 2.5 \times 10^{-4}$), evidence supporting a causal hypothesis. We also present multiple MR sensitivity analyses in support of this central finding. Our results provide evidence that hypercalcaemia is comorbid with migraine headache diagnoses, and that genetically elevated serum calcium over lifetime appears to increase risk for migraine. Further studies will be required to understand the biological mechanism, pathways, and clinical implication for risk management.

Received: July 29, 2016. Revised: November 28, 2016. Accepted: December 5, 2016

© The Author 2016. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Introduction

Migraine is the most common neurological disorder in the world with over a billion afflicted (1), with elevated prevalence in women versus men of European ancestry (2). A highly debilitating disorder, it ranks among the costliest in terms of economic impact and lost productivity (3). Migraine is classified into two common forms: migraine with and without aura, an associated perceptual distortion that occurs before or close to the onset of migraine (4). Genome-wide association studies have demonstrated numerous genetic factors that underlie susceptibility to migraine generally (5), suggestive of a common aetiology between migraine with and without aura. An important next step is to translate these data into clinically relevant insights of causal mechanisms underlying the disorder.

Identification of causal factors for migraine is critical to success in translational medicine. Specifically, interventions based on causal factors have the best chance for success in ameliorating disease. While the Randomized Control Trial is the standard to assess causal relationships, such trials are expensive, time-consuming, and ultimately require knowledge of an intervention hypothesized to be clinically beneficial. Thus, research efforts generating compelling evidence for causal hypotheses offer the best chances for translational clinical outcomes to emerge.

Epidemiological studies can provide evidence of association between measurable biomarkers and risk to disease. However, these studies often require knowing what biomarker will be studied at the start of the study, and often cannot survey all possible traits. For example, a tumour within parathyroid glands can cause primary parathyroidism (6), resulting in excess secretion of parathyroid hormone and circulating levels of serum calcium, with severe headache as a common symptom. Given the sporadic nature of this disease, it is challenging to evaluate the relationship between elevated parathyroid hormone or serum calcium levels with a disposition to headache, with only anecdotal reports scattered in the literature (7–9). While not perfect, biomedical informatics mining of electronic health records could provide one complementary source of data to identify potential and plausible candidate traits like these for further study (10).

Once correlated risk factors have been identified, a suite of computational methods can provide direct evidence for and against causal hypotheses using human genomics data (11–13). One approach, termed Mendelian Randomization, uses genetic variants associated with a biomarker of interest in an instrumental variable analysis to estimate a causal effect of the biomarker on a disease endpoint. The approach has some analogy to the classic Randomized Control Trial, where genotype operates as the randomized intervention (14,15). Because alleles are sorted randomly at birth and the assumption that trait endpoints do not change germline variation, the approach avoids issues due to reverse causality. In addition, careful selection of genetic variation for testing can address issues of confounding. While these methods require several key assumptions to hold, application of the approach has been used extensively to generate evidence for (16) and against (17) specific causal hypotheses, importantly where clinical trial data is not readily available. A second, recently described approach estimates genetic correlation – joint heritability – between a pair of traits (18). Application of the approach recapitulates expected causal associations such as low-density lipoprotein cholesterol and heart disease, as well as obesity and T2D, though one limitation of the approach is that this does not provide a direction of effect between the trait and endpoint. Nonetheless, both approaches are

complementary towards evaluating causal hypotheses using genomics data.

Serum calcium level is a continuous, quantitative biomarker that is maintained within a homeostatic range. The prevalence of extreme levels of circulating serum calcium (diagnosed as hypercalcaemia) has a range of underlying causes in the general population and is difficult to estimate given its frequency, though has been observed in 0.5–1% of patients who visit emergency departments (39). Genetic variation associated with population-level variability in serum calcium levels have also been reported (19), allowing the opportunity to apply the above approaches—mining electronic health records, estimation of genetic heritability, and direct causal inference testing—to evaluate the hypothesis that serum calcium is a causal risk factor for migraine headache. Here, we use biomedical electronic health records to show a correlation between elevated calcium levels and migraine, and subsequently large-scale genetic data sets to test the hypothesis that genetically elevated serum calcium levels increase risk for migraine headache.

Results

Hypercalcaemia and migraine headache are diagnoses that occur together frequently

Large-scale epidemiological association studies have not systematically examined a potential correlation between serum calcium levels and migraine headache. Thus, we aimed to test the hypothesis that Migraine headache diagnoses are associated with elevated serum calcium levels, diagnosed as hypercalcaemia. To do this, we obtained ICD-9 diagnosis codes in over 1 million de-identified health records curated as part of the PennOmics resource (Methods), quantifying the odds that a Migraine Headache diagnoses (MHD) co-occurs in the patient record with another ICD-9 diagnosis code (Methods). We observed co-occurrence between MHD ICD-9 codes and ICD-9 codes for Hypercalcaemia (OR = 1.58, $P = 4.75 \times 10^{-13}$, [Supplementary Material, Table S2](#)), including adjustments for age, sex, and ancestry. These data are consistent with the hypothesis that migraine and elevated serum calcium levels (diagnosed as hypercalcaemia) occur together frequently in the electronic registry of our patient cohort.

A migraine headache may have additional contributing factors or co-diagnoses that might occur, potentially confounding our initial analysis. Consequently, we looked for additional diagnoses that frequently occurred along with migraine headache in our patient records. We found that MHD was co-diagnosed with hypothyroid, hypertension, and hyperlipidemia (OR = 1.44, 1.48, and 1.80, each $P < 10^{-50}$, [Table 1](#)), as well as hyperparathyroid (OR = 0.75, $P = 1.25 \times 10^{-3}$, [Table 1](#)). After including all of these additional factors into the previous model, we again observed co-occurrence for MHD with a diagnosis of hypercalcaemia (OR = 1.23, $P = 1.75 \times 10^{-3}$, [Table 1](#)). These data are consistent with the hypothesis that migraine headache and elevated serum calcium levels are frequently diagnosed together, independently of other important diagnoses that also occur with migraine.

Evidence of a common genetic basis between serum calcium levels and migraine headache

If serum calcium has a genetic link to migraine headache, genetic risk factors should be shared in common across the entire genome, and therefore, both traits would share genetic

Table 1. Phenotypic co-occurrence between migraine headache, demographics, hypercalcaemia, and other risk factors. Adjusted odds ratio (OR) and estimates from logistic regression, including Age, Sex, hypertension, hyperlipidemia, hypothyroidism, hyperparathyroidism, and hypercalcaemia

Variables	Overall n=1038457	with MHD n=29607	Without MHD n=1008850	Adjusted OR of MHD (95% CI)	P-value
Age (overall) ^a				0.85 (0.84–0.86)	<0.001
Mean (SD)	48.7 (21.7) % ^b	45.9 (15.5) % ^c	48.7 (21.9) % ^c		
Sex					
Male	439862 (42.4%)	5272 (1.2%)	434590 (98.8%)	1.00	
Female	598595 (57.6%)	24335 (4.1%)	574260 (95.9%)	3.69 (3.58, 3.80)	<0.001
Race					
Black	314045 (30.2%)	7850 (2.5%)	306195 (97.5%)	1.00	
White	724412 (69.8%)	21757 (3.0%)	702655 (97%)	1.38 (1.34, 1.42)	<0.001
Diagnosis, Yes					
Hyperlipidemia	203393 (19.6%)	7812 (3.8%)	195581 (96.2%)	1.80 (1.74, 1.87)	<0.001
Hypertension	234285 (22.6%)	8042 (3.4%)	226243 (96.5%)	1.48 (1.43, 1.54)	<0.001
Hypothyroid	57846 (5.6%)	2976 (5.1%)	54870 (94.8%)	1.44 (1.38, 1.50)	<0.001
Hyperparathyroid	5190 (0.5%)	148 (2.9%)	5042 (97.1%)	0.75 (0.64–0.90)	1.25×10^{-3}
Hypercalcaemia	6306 (0.6%)	270 (4.3%)	6036 (95.7%)	1.23 (1.08–1.40)	1.75×10^{-3}

^aReference is for computation of OR per 10-year increase in age.

^bProportions are within the columns.

^cproportions are across rows.

heritability. To test this hypothesis, we obtained genome-wide association data for both serum calcium levels and migraine headache (5,19), and estimated genetic correlation between both traits (all subjects, Methods). In this analysis, we observed a moderate genetic correlation between migraine and calcium levels ($r_g = 0.19$, 95% CI: 0.02–0.36, $P = 0.03$), rejecting the hypothesis that these traits are independent of one another, and evidence that serum calcium levels and migraine headache may have a genetic basis in common.

Genetically elevated serum calcium levels are associated with migraine headache

We next turned to directly test the hypothesis that elevated serum calcium levels cause increased susceptibility to migraine headache, using a two-sample Mendelian Randomization study design. We constructed a risk score using all eight genetic factors associated with serum calcium levels (Fig. 1, Supplementary Material, Table S3, Methods). This genetic instrument explained 1.25% of the variability in serum calcium levels, and was sufficient strength to minimize effects from weak instrument bias for our analysis of migraine headache (F-statistic = 118.9). Based on 23,285 migraineurs and 95,425 controls (5), we found that elevation of serum calcium levels by a hypothetical 1 mg/dl resulting from our genetic score was associated with an increase in risk of migraine (OR = 1.80, 95% CI: 1.31–2.46, $P = 2.5 \times 10^{-4}$, Table 2). Furthermore, when stratified by aura status, we found that elevation of serum calcium levels by a hypothetical 1 mg/dl resulting from our genetic score was associated with an increase in risk of migraine with (OR = 2.66, 95% CI: 1.41–5.02, $P = 2.6 \times 10^{-3}$, Table 2) or without aura (OR = 2.57, 95% CI: 1.49–4.43, $P = 6.6 \times 10^{-4}$, Table 2).

Next, we identified four genetic risk factors exclusively associated with serum calcium levels and their effects (Supplementary Material, Table S3, Methods). This genetic instrument explains 0.87% of the variability in serum calcium levels and has sufficient strength to minimize effects from weak instrument bias for our analysis of migraine headache

(F-statistic = 82.5). We estimated the causal effect of genetic elevation of serum calcium levels with migraine headache, using a genotype risk score method (Methods). We found that elevation of serum calcium levels by a hypothetical 1 mg/dl resulting from this exclusive genetic score was associated with an increase in risk of migraine (OR = 1.81, 95% CI: 1.24–2.63, $P = 2 \times 10^{-3}$, Table 2). Furthermore, when stratified by aura status, we found that elevation of serum calcium levels by a hypothetical 1 mg/dl resulting from our genetic score was associated with an increase in risk of migraine with (OR = 2.72, 95% CI: 1.27–5.81, $P = 9.8 \times 10^{-3}$), or without aura (OR = 2.84, 95% CI: 1.50–5.40, $P = 1.4 \times 10^{-3}$, Table 2). As the confidence intervals for these aura-stratified analyses included the causal effect estimate based on the entirety of the data, we conclude that the estimates stratified by aura status are not statistically different from the unstratified analysis.

Sensitivity analysis for the serum calcium genetic instruments using alternative causal inference methods

We next evaluated the robustness of our primary observation by applying a suite of sensitivity analyses. First, an unweighted risk score based on all serum-calcium variants was associated with an increased risk of migraine in all samples ($P = 2.9 \times 10^{-3}$, Table 2) or stratified by aura ($P = 0.011$ with aura, $P = 0.037$ without aura, Table 2), suggesting that our result is robust to the choice of weights used in our risk score. We next obtained a weighted-median estimate for the causal effect using all variants, an approach that generates a causal estimate analogous to the risk score method used above. This approach is unbiased asymptotically but only requires 50% or more of the weight for the score to derive from valid instruments, a less stringent requirement compared to the GRS score method (40). The weighted-median estimated for 1 mg/dl genetic elevation of serum calcium on migraine risk agreed with estimates from the analysis above which included all samples (OR = 1.92, 95% CI: 1.30–2.84, $P = 1.6 \times 10^{-3}$, Table 2), or analysis that stratified either with aura (OR = 2.73, 95% CI: 1.29–5.97, $P = 9.6 \times 10^{-3}$, Table 2) or

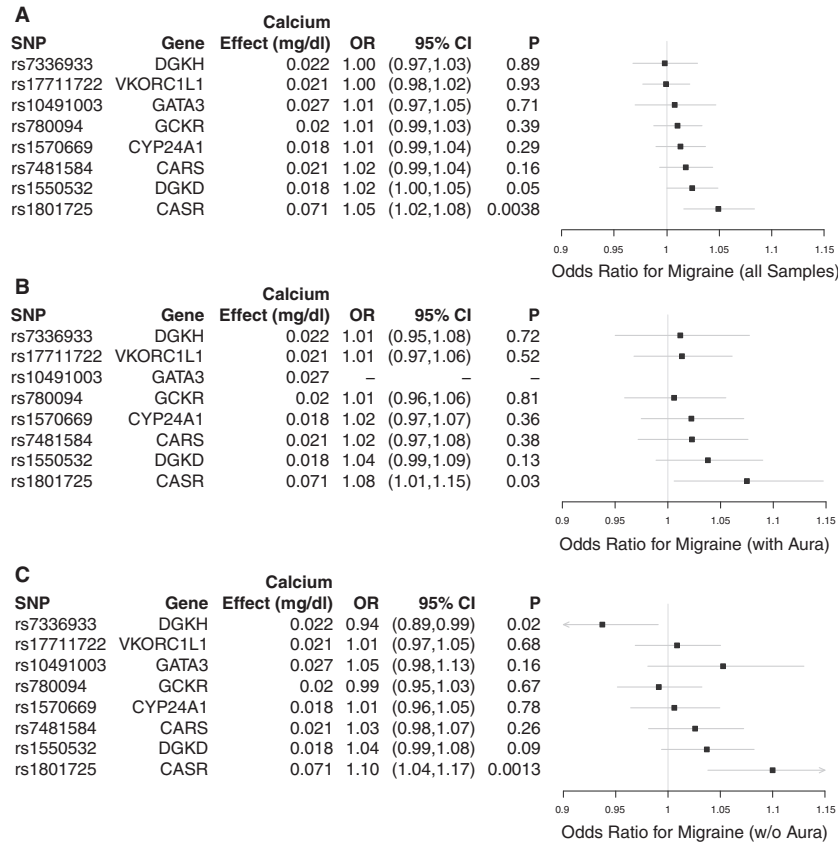


Figure 1. Summary SNP association data for each serum calcium variant for migraine. The effect size of the effect on serum calcium is given in units of mg/dl. (A) All samples. (B) Subset of migraine patients with aura. (C) Subset of migraine patients without aura. OR, Odds Ratio; CI, Confidence Interval.

Table 2. Summary statistics for genetic instruments used for causal inference analysis for serum calcium for migraine traits

	All Samples		With Aura		Without Aura	
	Odds Ratio ^a (95% CI)	P	Odds Ratio (95% CI)	P	Odds Ratio (95% CI)	P
Analyses using All Variants (n=8)						
Cumulative GRS ^b	1.80 (1.31–2.46)	2.5×10^{-4}	2.66 (1.41–5.02)	2.6×10^{-3}	2.57 (1.49–4.43)	6.6×10^{-4}
Unweighted GRS ^b	1.66 (1.19–2.31)	2.9×10^{-3}	2.43 (1.22–4.82)	0.011	1.87 (1.04–3.38)	0.037
Weighted-median GRS ^c	1.92 (1.30–2.84)	1.6×10^{-3}	2.73 (1.29–5.79)	9.6×10^{-3}	3.60 (1.81–6.85)	$<2.0 \times 10^{-4}$
Egger Regression (Causal Effect) ^c	1.97 (1.05–3.69)	0.037	2.70 (0.78–8.75)	0.11	5.42 (1.73–16.19)	2.2×10^{-3}
Egger Regression (Bias Term) ^d	–0.003 (–0.025–0.019)	0.74	–0.0007 (–0.048–0.046)	0.97	–0.026 (–0.065–0.013)	0.16
SIMEX MR-Egger (Causal Effect) ^d	2.01 (1.05–3.80)	0.037	2.85 (1.32–6.17)	0.021	5.80 (0.91–37.0)	0.066
SIMEX I ² (weighted)	0.933		0.943		0.938	
Analyses using Serum Calcium Exclusive variants (n=4)						
Cumulative GRS ^b	1.81 (1.24–2.63)	2.0×10^{-3}	2.72 (1.27–5.81)	9.8×10^{-3}	2.84 (1.50–5.40)	1.4×10^{-3}
Unweighted GRS ^b	1.61 (1.06–2.45)	0.025	2.45 (1.06–5.64)	0.036	1.86 (0.88–3.93)	0.10

^aOdds ratios are given in units of fold increase in migraine risk per unit increase in serum calcium (1 mg/dl).

^bP-value and 95% CI from the normal distribution.

^cP-value and 95% CI calculated by bootstrapping (n = 10,000).

^dP-value and 95% CI calculated from the Student's t distribution.

without aura (OR = 3.60, 95% CI: 1.81–6.85, $P < 2 \times 10^{-4}$, Table 2). To evaluate the potential for systematic bias in our data, we applied Egger regression to estimate bias and a causal effect in the context of bias (25). However, we did not observe evidence of bias ($P > 0.15$ for all instruments, Table 2), and found a modest association for the score that included all migraine-affected subjects ($P = 0.037$, Table 2). To further evaluate the potential of bias and violation of the no measurement error (NOME)

assumption, we applied the recently reported SIMEX approach (41). We found that each of our scores had a high I² (Table 2) indicative of a robust genetic instrument, and our estimated causal effects were similar to that obtained by Egger Regression, with modest association for scores using all migraine-affected or those with aura ($P < 0.05$, Table 2). Finally, we excluded our strongest variant from our constructed instrument (rs1801725, CASR). We observed modest nominal association using all

migraineurs (Supplementary Material, Table S4), consistent with an attenuated, residual contribution to migraine across all serum calcium loci. Taken collectively, these analyses suggest that inference is robust to the weights we selected for our genetic instruments, and our inference that genetically elevated serum calcium levels increase susceptibility to migraine headache does not seem to have evidence of clear violations in several specific assumptions made in our Mendelian Randomization analysis.

Discussion

Causal inference is one of the most challenging and important problems in biology and medicine. To achieve impactful translational outcomes efficiently, identification of causal risk factors is crucial. Computational approaches that rapidly facilitate causal hypothesis testing using the approaches we describe here are designed to begin the process of building a compelling, evidence-based case for causality, complementing or perhaps even motivating future experimental activities in models systems and in humans. We present three lines of support for an association between serum calcium levels and susceptibility to migraine headache: (i) comorbidity analysis using electronic health records indicates that diagnoses for migraine and serum calcium (using hypercalcaemia as a proxy) occur in patient records together more than expected, (ii) genetic heritability analysis indicates that serum calcium and susceptibility to migraine headache may have a genetic basis in common, and (iii) a genetic risk score analysis in the Mendelian Randomization framework indicates that a hypothetical 1 mg/dl genetic increase in serum calcium levels is associated with a 1.8-fold increase in risk to migraine headache. Taken collectively, these data support an epidemiological and genetic correlation, and a potentially causal connection between levels of circulating calcium and susceptibility to migraine.

As an alternative to epidemiological data, which in our case was not available, we extracted diagnosis ICD-9 codes from >1M electronic health records as a proxy. This design does have some limitations that should be acknowledged. First, the diagnosis codes of migraine that we obtained may not reflect a true clinical diagnosis of migraine headache obtained from a trained neurologist. Because the data are de-identified, we are not able to go back and 'verify' the diagnoses that are listed, or check specific biomarker measurements in patients (i.e. levels of serum calcium) in a comprehensive way. It may be important in the future to consider systematic, computational diagnosis of migraine headache using known affected subjects within the health record data to train a model to make accurate predictions of which subjects have clinically defined migraine headache. Second, because the collection of diagnoses begins at 2008, we are not able to determine the order of diagnoses. Thus, our analysis here may be susceptible to reverse causality, as is the case with other types of correlational studies, as well as confounding from factors that we did not include in our analysis.

Despite these limitations, there is some indication that our association does match previous epidemiological reports. Features of metabolic syndrome (hypertension, hyperlipidemia), as well as sex and hypothyroid are all positively associated with migraine (27–30), and our findings match these previous observations. As hyperparathyroidism is a frequent cause of hypercalcaemia, it is somewhat curious why the adjusted co-occurrence is in the opposite direction. At face, this implies that in the context of a patient presenting metabolic syndrome traits and features of headache, hyperparathyroid is

less often diagnosed. Further work would be required to rule out ascertainment against the diagnosis.

We also performed a genetic correlation analysis, to provide additional evidence supporting a relationship between these two traits. We note that this type of correlation evidence is different from, but complementary to, the Health record comorbidity analysis. Genetic correlation not only connects directly to the quantitative trait measured (serum calcium levels), but also directly estimates correlation between traits due to genetics, which should be correlated if a causal hypothesis is true. However, this analysis does not immediately give a direction for the relationship between traits, and still could be subject to confounding due to directional bias from unknown, pleiotropic associations.

Using the framework of Mendelian Randomization, large-scale genetic studies provide one of many pieces of evidence in support of consistent and robust causal relationships between measurable biomarkers and risk of clinical endpoints. However, the Mendelian Randomization analyses we performed here also have limitations that should be acknowledged. First, our analysis requires (i) a potent genetic instrument and effect on serum calcium, (ii) that the serum calcium genetic variants are not associated with one or more unmeasured confounding variables which also increase risk of migraine, and (iii) that the effect of change in migraine through our selected genetic instruments are mediated entirely through serum calcium levels. We assessed several of these assumptions in turn. While our genetic instrument explained ~1% of the variability in serum calcium levels, this was a sufficiently strong to minimize weak residual bias, assessed by *F*-statistic. While we cannot completely rule out additional confounding, we did examine the association of our SNP panel against a number of cardiometabolic traits, identifying a subset that was free of metabolic associations. We also generated causal effect estimates using both unweighted and weighted-median risk scores, and our results were consistent with the full genotype risk score approaches and suggest our inference is robust to potentially invalid instrument contributing to our score. Furthermore, we utilized Egger's regression and the SIMEX approach to evaluate systemic bias from our genetic instruments, and did not reject the null hypothesis of no bias. However, our negative results here do not completely rule out all forms of potential biases. While the MR analysis of the substrata estimated a higher causal effect than the entire population (Table 2), this could be reasonably explained by the fact that the 'all sample' migraine meta-analysis data includes studies that were not present in either of the aura stratified analyses.

Calcium channel blockers (CCBs) are often used as a migraine prophylactic, thus possibly implicating Ca^{2+} into the aetiology of migraine. However, the mechanism of action of these drugs is thought to target the vascular aetiology of migraine, e.g. lower cerebral vasoconstriction by reducing Ca^{2+} cellular influx (31,32). A recent meta-analysis of clinical trials data from a range of CCBs found no difference in reduction of migraine headaches relative to placebo (33), though clearly more data on specific drugs are required. Furthermore, CCBs may not ultimately impact levels of serum calcium, except in extreme cases, as changes in serum calcium levels are probably compensated by changes in levels of parathyroid hormone to maintain homeostasis. One possibility is that certain CCBs may act to raise parathyroid hormone by lowering serum calcium though hypercalciuria (34), which would be a possible, additional mechanism of action for some of these drugs that is consistent with our findings. Overall, our results suggest that direct targeting of

Ca²⁺ levels – in addition to the known vasodilatory effect of CCBs – may offer additional therapeutic benefits.

One of the variants we studied was rs1801725, a coding missense mutation (Ala986Ser) in the calcium sensing receptor gene, *CASR*. Indeed, the observed association at this genetic variant contributes heavily in our Mendelian Randomization result, in addition to the aggregated contribution of weaker serum calcium association in our risk score analysis. *CASR* is expressed in the parathyroid gland and kidney where it detects changes in circulating calcium levels and relays this information back to intracellular signalling pathways to regulate calcium homeostasis via secretion or absorption (35). Rare mutations in *CASR* can cause Mendelian forms of either hypercalcaemia (if loss of function) or hypocalcaemia (activating mutations). Here, the alanine residue at this site is a conserved amino acid (back to the model fish, Stickleback); carriage of the serine allele is associated with elevated serum calcium levels, consistent with a hypomorphic function *CASR* with carriage of serine at this position. Given the evidence from Mendelian disease studies, evolutionary constraint at this site, and that this presumed hypomorphic allele is coding, a strong hypothesis is that this variant is indeed causal, though functional assays are necessary to demonstrate the effect of this variant formally.

Because the serine allele is also associated with increased susceptibility to migraine headache ($P = 3.8 \times 10^{-3}$), one hypothesis is that modest pharmacological *CASR* agonism would both lower serum calcium and protect against migraine headache. Cinacalcet, an FDA approved agonist of *CASR* marketed as Sensipar, is approved in treatment for the limited indications of secondary hyperparathyroidism or hypercalcaemia in patients with parathyroid carcinoma. While the drug has a narrow applicability owing to the potential risk of hypocalcaemia, along with additional adverse side effects (nausea, muscle or chest pain, and osteoporosis), it still may be worthwhile to explore the repurposing potential of this compound for treatment of migraine headache in specific instances.

Our comorbidity analysis, while providing some correlation between hypercalcaemia and migraine, suggest the potential clinical impact. While patients with hypercalcaemia are potentially rare diagnoses overall (0.6% in our cohort), 1% of migraine subjects were also diagnosed with hypercalcaemia. If we assumed that elevated calcium levels were causal for migraine susceptibility, the implication is that this small population of migraineurs (~1%) could be managed by returning their serum levels to homeostatic levels. There could also remain a larger population of subjects who have undiagnosed hypercalcaemia, or even levels of serum calcium that are higher than the population average but who are otherwise sub-clinical, who could be tested (and potentially treated) in a similar manner. An important consideration here is if the amount of change in serum calcium we note is actually obtainable to achieve a clinical benefit. We demonstrate that genetic elevation of serum calcium by 1 mg/dl increase was associated with increased odds of migraine headache by 1.8-fold. One study has shown that >85% of patients taking Sensipar had their serum calcium levels decreased by >1 mg/dl (36); a second study demonstrated a median difference of ~1 mg/dl for treated versus placebo controls (37). These clinical trials indicate that the level of change in serum calcium associated with migraine risk is, in fact, clinically obtainable.

In summary, we provide evidence supporting the correlation in diagnoses of serum calcium levels (quantified as hypercalcaemia) and migraine headache, independent of additional, potentially confounding risk factors. In addition, we provide

evidence for shared narrow-sense genetic heritability, using large-scale data from human populations studied separately for serum calcium levels and migraine headache. Finally, we found that genetic variants exclusively associated with circulating calcium levels are associated with migraine headache, providing evidence of a causal relationship between elevated serum calcium levels over a lifetime and migraine headache.

Materials and Methods

Description of the PennOmics resource

PennOmics is a data warehouse that integrates research and clinical data from several separate data storage resources from area hospitals affiliated with the University of Pennsylvania Health System (UPHS). This includes records stored at the Penn Data Store (PDS), the Cancer Center tumour registry, the Velos Clinical Trial Management System, and genomics data from contributing clinical and research labs throughout UPHS. Records and clinical data contained in PennOmics are continually updated (>3 million potential patient records are contained there as of this writing) and are completely de-identified, meaning all direct patient identifiers have been removed, including adjustment of patient age information to further obfuscate the identification of individual subjects. Details for the prevalence of each ICD-9 diagnosis codes and demographic data used in this study are provided in Table 1.

Comorbidity analysis for ICD-9 codes using the PennOmics resource

ICD-9 diagnoses and demographic patient information from the PennOmics resource, comprising 1,098,023 subjects with records collected after 2008, was queried on November 6th, 2015. Initially, we obtained all diagnoses for subjects where sex and age (in years) was present, in non-hispanic and White ($n = 724,412$) and Black ($n = 314,045$) reported ethnicities. Hispanic white samples were excluded here, owing to the small number of reported MHD diagnoses given the total number of samples present in this extracted collection. For each individual, binary variables for each phenotype group were constructed, which included (i) migraine headache, (ii) hypercalcaemia, (iii) hypertension, (iv) hyperlipidaemia, (v) hypothyroid, and (vi) hyperparathyroid, based on available ICD-9 codes in the resources (Supplementary Material, Table S1). A subject was labelled '1' for the given variable if any of these assigned ICD-9 codes were present in the record, otherwise '0'. We used logistic regression (implemented in R (3.13), using `glm()`, binomial link function) to report the odds of co-occurrence of the migraine headache ICD-9 status diagnosis code with ICD-9 code variables for hypercalcaemia, including hypertension, hypothyroid, hyperparathyroid or hyperlipidaemia, age, and sex as additional covariates. These factors were selected given previous reports of correlation with migraine (27–30). Results are presented for the analysis that included ancestry as an additional covariate in the model, which were qualitatively similar when each ethnicity was analyzed individually (Supplementary Material, Table S2).

Heritability analysis for migraine headache and serum calcium levels

We used the bivariate heritability estimation method implemented in LDSC (v1.0.0) to estimate the genetic correlation

between migraine and serum calcium levels. The details of the method are described elsewhere (18). Briefly, assuming a polygenic model, the genetic correlation between two traits can be estimated using the summary association statistics (i.e. p-values converted to Z-scores via inverse-normal transformation) for both traits and for all tested genetic variants included in a genome-wide association panel, and the 'LD score', which measures the amount of genetic variation tagged (i.e. in linkage disequilibrium) by the SNP (38). Under some assumptions, the slope of the regression of the product of the association Z-scores on the LD score can be used to estimate the genetic correlation between the two traits. Precise details of the method, its efficiency, and accuracy are described elsewhere (38). For this analysis, we utilized common variants that were present in the Hapmap3 reference (20), which were also genotyped across most of the sample: specifically, the number of analyzed samples not less than two-thirds of the 90th percentile of the total in each meta-analysis. We obtained summary GWAS data for serum calcium levels from a recent report (19), which involve an inverse-variance fixed-effects Hapmap reference panel-based imputation meta-analysis of 17 population-based cohorts of European descent, all measured for circulating serum calcium levels. For migraine headache, we utilized association data from a recent report, based on an inverse-variance fixed-effects Hapmap reference panel-based imputation meta-analysis including 29 cohorts of European ancestry (5). In brief, contributing cohorts derived mostly from northern European (Finland, Germany, Netherlands, UK, Iceland, Sweden, Norway, Estonia, etc.), with cases/controls were matched within each study. Stratification was further controlled by inclusion of principal components for genetic similarity as covariates in the association analysis by each study. From these data, all subjects diagnosed with migraine headache included 23,285 individuals with migraine (cases defined from both population and clinically defined cohorts) and 95,425 population-matched controls.

Construction of genetic instruments for Mendelian randomization analysis

We used our previous reported tool, MeRP, to perform instrument construction and testing (21). Briefly, the National Human Genome Research Institute (NHGRI) maintains a compilation of data from GWAS publications (22), and was used as our starting point. All ($n=8$ total) serum calcium single nucleotide polymorphic (SNP) associations that were genome-wide significant ($P < 5 \times 10^{-8}$) were obtained. These SNPs then went through filtering steps for LD and confounding trait associations. First, associations with potential confounding traits were assayed from within the NHGRI GWAS catalog and a summary file containing association p-values for >2 million SNPs and 15 cardio/metabolic traits (these included blood pressure, cholesterol, anthropometric and obesity, glucose/insulin, along with type 2 diabetes and coronary heart disease) obtained from the public domain, and violating SNPs were removed. The remaining SNPs were grouped together based on pairwise LD ($r^2 > 0.05$) and the SNPs with the lowest trait associated p-value value from each LD group were selected to comprise the serum calcium genetic variants for MR analysis. This resulted in four ($n=4$: rs7336933, rs1801725, rs10491003, rs1570669) SNPs exclusively associated with serum calcium levels that comprised our final, serum calcium instrument for MR testing. The list of exclusive and non-exclusive serum calcium ($n=4$: rs780094, rs17711722, rs7481584, and rs1550532) associated SNPs along with annotations for

nearest genes are provided (Supplementary Material, Table S3). Phenotypic associations for non-exclusive variants includes the following: rs780094 has multiple established trait associations ($P < 5 \times 10^{-8}$) with triglycerides, total cholesterol, Height, fasting glucose; rs17711722 was nominally associated ($P < 0.05$) with triglycerides, body mass index (BMI), 2-hour glucose, fasting insulin, and HOMA-IR; rs7481584 was nominally associated with high-density lipoprotein cholesterol, BMI, and type 2 diabetes; rs1550532 was nominally associated with fasting insulin, HOMA-IR, and coronary heart disease.

Causal inference analysis between serum calcium levels and migraine headache

We utilized a statistical method for estimating causal effects based on multi-SNP genetic instruments using summary data alone, using a two-sample Mendelian Randomization (MR) design based on summary statistics. In the two-stage design, association data from both the intermediate trait (in this case, serum calcium levels) and the endpoint (migraine headache) are obtained from two separate genome-wide studies. This study design is a valid approach to infer causal effects (23). Briefly, if genetic variants (SNPs) are unlinked, and further if the effects of each SNP on levels of serum calcium levels are log-additive, a causal effect ($\hat{\alpha}$) and the error between a biomarker and outcome can be estimated by (24):

$$\hat{\alpha} = \frac{\sum_j w_j \hat{\beta}_j \sigma_j^{-2}}{\sum_j w_j^2 \sigma_j^{-2}} \quad (1)$$

$$SE(\hat{\alpha}) = \sqrt{\frac{1}{\sum_j w_j^2 \sigma_j^{-2}}} \quad (2)$$

where for all j SNPs, β_j represents the estimated natural log odds effect of the j -th SNP on the endpoint of interest, σ_j represents the standard error on the log odds effect of the j -th SNP on the endpoint, and w_j represents a weight for the SNP on the outcome. Each SNP was weighted using the reported estimated effect of the SNP on serum calcium levels (in units of mg/dl). The significance of the estimate was assessed via the ratio of the above quantities (i.e. $\hat{\alpha}/SE(\hat{\alpha})$), which is chi-squared distributed with 1 degree of freedom. Effects of each variant used in our genetic risk score on migraine headache were obtained from the genome-wide association data set from the International Migraine Headache Genetics Consortium (5), the same used in the heritability analysis described above, and we also included additional stratified analyses with (5,118 cases, 74,239 controls) and without (7,107 cases and 69,427 controls) aura. Our two-stage design assumed that samples are non-overlapping between the endpoint (Migraine Headache Genetics consortium) and the intermediate trait (the Serum Calcium). Only one study (TwinsUK) was common to both, and only contributed to replication for serum calcium associated variants, thus, we used estimates from the primary scan which did not include this study.

We conducted additional sensitivity analysis to examine potential bias in our instruments and robustness of our analyses. This includes (i) Egger regression using all eight associated serum calcium variants to evaluate systematic bias (25), and (ii)

the recently reported SIMEX method designed evaluate the ‘no measurement error’ assumption (41). We also computed a causal estimate using an unweighted genotype scores (*i.e.* weights set to one for all variants). As for the weighted scores, the migraine log-odds was polarized to the serum calcium-increasing allele. In addition, we performed a weighted-median method, which can also assess a causal effect (unbiased asymptotically as sample size increases) and only requires at least 50% of the weight for the score to derive from valid instruments (40). Weighted median estimation was carried out using the sample code previously made available (40). Standard error and 95% CI were estimated by bootstrapping for the median-weighted and Egger regression approaches. To measure the strength of the genetic instruments, we calculated F-statistics for each score as previously described (26), based on the total symmetric sample size of the asymmetric migraine study design ($N = 2 \times (n_{\text{aff}} \times m_{\text{controls}}) / (n_{\text{aff}} + m_{\text{controls}})$).

Supplementary Material

Supplementary Material is available at HMG online.

Acknowledgements

The authors would like to thank the International Headache Genetics Consortium for migraine association data sets, Caroline Fox, Murielle Bochud, and Qiong Yang for assistance in accessing the serum calcium association data, and the Penn Medicine Data Analytics Center for their assistance in assembling the information used in this study, and comments from the two anonymous reviewers.

Conflict of Interest statement. None declared.

Funding

A.P. was supported by the European Commission FP7 project no. 602633-2 EUROHEADPAIN, by the Academy of Finland grants no. 251704 and 286500, and by the Sigrid Juselius Foundation. B.F.V. was supported by grants from the American Heart Association (13SDG14330006) and the NIH (R01DK101478). G.D.S. works within the Medical Research Council Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council (MC_UU_12013/1). The funding bodies had no role in the design of the study, results, or decision to publish. Funding to pay the Open Access publication charges for this article was provided to B.F.V. by the NIH/NIDDK (DK101478).

References

- Global Burden of Disease Study 2013 Collaborators. (2015) Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*, **386**, 743–800.
- Stovner, L.J., Zwart, J.A., Hagen, K., Terwindt, G.M. and Pascual, J. (2006) Epidemiology of headache in Europe. *Eur. J. Neurol.*, **13**, 333–345.
- Gustavsson, A., Svensson, M., Jacobi, F., Allgulander, C., Alonso, J., Beghi, E., Dodel, R., Ekman, M., Faravelli, C., Fratiglioni, L., et al. (2011) Cost of disorders of the brain in Europe 2010. *Eur. Neuropsychopharmacol.*, **21**, 718–779.
- Headache Classification Subcommittee of the International Headache Society. (2004) The International Classification of Headache Disorders: 2nd edition. *Cephalalgia*, **24** Suppl 1, 9–160.
- Anttila, V., Winsvold, B.S., Gormley, P., Kurth, T., Bettella, F., McMahon, G., Kallela, M., Malik, R., de Vries, B., Terwindt, G., et al. (2013) Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat. Genet.*, **45**, 912–917.
- Kleeman, C.R., Massry, S.G. and Coburn, J.W. (1971) The clinical physiology of calcium homeostasis, parathyroid hormone, and calcitonin. II. *Calif. Med.*, **114**, 19–30.
- Geelhoed, G.W. (1982) Parathyroid adenolipoma: clinical and morphologic features. *Surgery*, **92**, 806–810.
- Basso, U., Maruzzo, M., Roma, A., Camozzi, V., Luisetto, G. and Lumachi, F. (2011) Malignant hypercalcemia. *Curr. Med. Chem.*, **18**, 3462–3467.
- Ozen, S., Celik, A., Alper, H., Simsek, D.G. and Darcan, S. (2010) Primary hyperparathyroidism in a young adult presenting with severe hypertension. *J. Pediatr. Endocrinol. Metab.*, **23**, 193–196.
- Blair, D.R., Lyttle, C.S., Mortensen, J.M., Bearden, C.F., Jensen, A.B., Khiabani, H., Melamed, R., Rabadan, R., Bernstam, E.V., Brunak, S., et al. (2013) A nondegenerate code of deleterious variants in Mendelian loci contributes to complex disease risk. *Cell*, **155**, 70–80.
- Davey Smith, G. and Ebrahim, S. (2003) ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.*, **32**, 1–22.
- Davey Smith, G. and Ebrahim, S. (2004) Mendelian randomization: prospects, potentials, and limitations. *Int. J. Epidemiol.*, **33**, 30–42.
- Sheehan, N.A., Didelez, V., Burton, P.R. and Tobin, M.D. (2008) Mendelian randomisation and causal inference in observational epidemiology. *PLoS Med.*, **5**, e177.
- Hingorani, A. and Humphries, S. (2005) Nature’s randomised trials. *Lancet*, **366**, 1906–1908.
- Davey Smith, G. and Ebrahim, S. (2005) What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ*, **330**, 1076–1079.
- Cohen, J.C., Boerwinkle, E., Mosley, T.H., Jr. and Hobbs, H.H. (2006) Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N. Engl. J. Med.*, **354**, 1264–1272.
- Voight, B.F., Peloso, G.M., Orho-Melander, M., Frikke-Schmidt, R., Barbalic, M., Jensen, M.K., Hindy, G., Hólm, H., Ding, E.L., Johnson, T., et al. (2012) Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet*, **380**, 572–580.
- Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R. and Loh, P.R. ReproGen Consortium; Psychiatric Genomics Consortium. (2015) Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, Duncan L, et al. An Atlas of Genetic Correlations across Human Diseases and Traits. *Nat. Genet.*, **47**, 1236–1241.
- O’Seaghdha, C.M., Wu, H., Yang, Q., Kapur, K., Guessous, I., Zuber, A.M., Köttgen, A., Stoudmann, C., Teumer, A., Kutalik, Z., et al. (2013) Meta-analysis of genome-wide association studies identifies six new Loci for serum calcium concentrations. *PLoS Genet.*, **9**, e1003796.
- The International Hapmap Consortium. (2010) Integrating common and rare genetic variation in diverse human populations. *Nature*, **467**, 52–58.

21. Yin, P. and Voight, B.F. (2015) MeRP: a high-throughput pipeline for Mendelian randomization analysis. *Bioinformatics*, **31**, 957–959.
22. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorf, L. and Parkinson, H. (2014) The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic. Acids. Res.*, **42**, D1001–D1006.
23. Evans, D.M. and Davey Smith, G. (2015) Mendelian Randomization: New Applications in the Coming Age of Hypothesis-Free Causality. *Annu. Rev. Genomics. Hum. Genet.*, **16**, 327–350.
24. Dastani, Z., Hivert, M.F., Timpson, N., Perry, J.R., Yuan, X., Scott, R.A., Henneman, P., Heid, I.M., Kizer, J.R., Lyytikäinen, L.P., et al. (2012) Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. *PLoS Genet.*, **8**, e1002607.
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.
26. Burgess, S., Dudbridge, F. and Thompson, S.G. (2016) Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Stat. Med.*, **35**, 1880–1906.
27. Breslau, N. and Rasmussen, B.K. (2001) The impact of migraine: Epidemiology, risk factors, and co-morbidities. *Neurology*, **56**, S4–12.
28. Scher, A.I., Terwindt, G.M., Picavet, H.S., Verschuren, W.M., Ferrari, M.D. and Launer, L.J. (2005) Cardiovascular risk factors and migraine: the GEM population-based study. *Neurology*, **64**, 614–620.
29. Winsvold, B.S., Sandven, I., Hagen, K., Linde, M., Midthjell, K. and Zwart, J.A. (2013) Migraine, headache and development of metabolic syndrome: an 11-year follow-up in the Nord-Trøndelag Health Study (HUNT). *Pain*, **154**, 1305–1311.
30. Lisotto, C., Mainardi, F., Maggioni, F. and Zanchin, G. (2013) The comorbidity between migraine and hypothyroidism. *J. Headache Pain*, **14**, P138.
31. Meyer, J.S. and Hardenberg, J. (1983) Clinical effectiveness of calcium entry blockers in prophylactic treatment of migraine and cluster headaches. *Headache*, **23**, 266–277.
32. Peroutka, S.J. (1983) The pharmacology of calcium channel antagonists: a novel class of anti-migraine agents? *Headache*, **23**, 278–283.
33. Jackson, J.L., Cogbill, E., Santana-Davila, R., Eldredge, C., Collier, W., Gradall, A., Sehgal, N. and Kuester, J. (2015) A Comparative Effectiveness Meta-Analysis of Drugs for the Prophylaxis of Migraine Headache. *PLoS One*, **10**, e0130733.
34. Zaheer, S., de Boer, I., Allison, M., Brown, J.M., Psaty, B.M., Robinson-Cohen, C., Ix, J.H., Kestenbaum, B., Siscovick, D. and Vaidya, A. (2016) Parathyroid Hormone and the Use of Diuretics and Calcium-Channel Blockers: The Multi-Ethnic Study of Atherosclerosis. *J. Bone. Miner. Res.*, **31**, 1137–1145.
35. Hebert, S.C., Brown, E.M. and Harris, H.W. (1997) Role of the Ca(2+)-sensing receptor in divalent mineral ion homeostasis. *J. Exp. Biol.*, **200**, 295–302.
36. Khan, A., Bilezikian, J., Bone, H., Gurevich, A., Lakatos, P., Misorowski, W., Rozhinskaya, L., Trotman, M.L. and Tóth, M. (2015) Cinacalcet normalizes serum calcium in a double-blind randomized, placebo-controlled study in patients with primary hyperparathyroidism with contraindications to surgery. *Eur. J. Endocrinol.*, **172**, 527–535.
37. EVOLVE Trial Investigators, Chertow, G.M., Block, G.A., Correa-Rotter, R., Drüeke, T.B., Floege, J., Goodman, W.G., Herzog, C.A., Kubo, Y., London, G.M., et al. (2012) Effect of cinacalcet on cardiovascular disease in patients undergoing dialysis. *N. Engl. J. Med.*, **367**, 2482–2494.
38. Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Patterson, N., Daly, M.J., Price, A.L. and Neale, B.M. (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.*, **47**, 291–295.
39. Lindner, G., Felber, R., Schwarz, C., Marti, G., Leichtle, A.B., Fiedler, G.M., Zimmermann, H., Arampatzis, S. and Exadaktylos, A.K. (2013) Hypercalcemia in the ED: prevalence, etiology, and outcome. *Am. J. Emerg. Med.*, **31**, 657–660.
40. Bowden, J., Davey Smith, G., Haycock, P.C. and Burgess, S. (2015) Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.*, **40**, 304–314.
41. Bowden, J., Del Greco, M.F., Cosetta, M., Davey Smith, G., Sheehan, N.A. and Thompson, J.R. (2016) Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I^2 statistic. *Int. J. Epidemiol.*, in press, doi: 10.1093/ije/dyw220.