# **RESEARCH ARTICLE**





# Identifying the potential causal role of insomnia symptoms on 11,409 health-related outcomes: a phenome-wide Mendelian randomisation analysis in UK Biobank

Mark J. Gibson<sup>1,2\*</sup>, Deborah A. Lawlor<sup>1,3†</sup> and Louise A. C. Millard<sup>1,3†</sup>

# Abstract

**Background** Insomnia symptoms are widespread in the population and might have effects on many chronic conditions and their risk factors but previous research has focused on select hypothesised associations/effects rather than taking a systematic hypothesis-free approach across many health outcomes.

Methods We performed a Mendelian randomisation (MR) phenome-wide association study (PheWAS) in 336,975 unrelated white-British UK Biobank participants. Self-reported insomnia symptoms were instrumented by a genetic risk score (GRS) created from 129 single-nucleotide polymorphisms (SNPs). A total of 11,409 outcomes from UK Biobank were extracted and processed by an automated pipeline (PHESANT) for the MR-PheWAS. Potential causal effects (those passing a Bonferroni-corrected significance threshold) were followed up with two-sample MR in MR-Base, where possible.

**Results** Four hundred thirty-seven potential causal effects of insomnia symptoms were observed for a diverse range of outcomes, including anxiety, depression, pain, body composition, respiratory, musculoskeletal and cardiovascular traits. We were able to undertake two-sample MR for 71 of these 437 and found evidence of causal effects (with directionally concordant effect estimates across main and sensitivity analyses) for 30 of these. These included novel findings (by which we mean not extensively explored in conventional observational studies and not previously explored using MR based on a systematic search) of an adverse effect on risk of spondylosis (OR [95%CI] = 1.55 [1.33, 1.81]) and bronchitis (OR [95%CI] = 1.12 [1.03, 1.22]), among others.

**Conclusions** Insomnia symptoms potentially cause a wide range of adverse health-related outcomes and behaviours. This has implications for developing interventions to prevent and treat a number of diseases in order to reduce multimorbidity and associated polypharmacy.

Keywords Insomnia, Mendelian randomisation, MR-PheWAS, UK Biobank

<sup>†</sup>Deborah A. Lawlor and Louise A. C. Millard contributed equally to this work.

\*Correspondence:

Mark J. Gibson

mark.gibson@bristol.ac.uk

<sup>2</sup> School of Psychological Science, University of Bristol, Bristol, UK

<sup>3</sup> Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

# 

# Background

While there is still much debate over the exact purpose of sleep, it is clear that sleep is vital for healthy functioning and likely to be multifaceted. Experiments on rats have suggested that sleep is linked to antioxidative enzyme levels in the brain which regulate the levels of reactive oxygen species (by-products of the metabolization of

© The Author(s) 2023. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativeco mmons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

<sup>&</sup>lt;sup>1</sup> MRC Integrative Epidemiology Unit (IEU), University of Bristol, Bristol, UK

oxygen which damage cells) [1]. It has also been proposed that sleep is vital for the consolidation of information, learning, and memory [2, 3]. Insomnia is defined as regular dissatisfaction with the quality or quantity of sleep for a prolonged period and includes difficulty initiating or maintaining sleep [4]. Evidence suggests that 6–7% of the European population have a diagnosis of insomnia, while 33-37% self-report having insomnia symptoms [5-7]. It is the second most prevalent mental health disorder (after anxiety disorder) and is more common in women and the elderly [6, 7]. Multimorbidity, defined as patients living with two or more chronic health conditions, is associated with polypharmacy, poor quality of life and premature mortality [8, 9]. It is increasingly recognised as a threat to global health and identifying potential causes of multimorbidity is a research priority [10].

Given the high prevalence of insomnia symptoms, and their potentially causal associations with many diseases (including increased risk of depression [11, 12], substance use [13, 14], autism spectrum disorder and bipolar disorder [15], dementia [16], high body mass index and diabetes [17, 18], hypertension [19], cardiovascular disease [20-22], pain [23] and inflammation [24]), insomnia symptoms could lead to multimorbidity. However, studies to date have largely been observational and may not reflect causal effects, and/or have focused on hypothesised selected outcomes, predominantly mental, neurocognitive and cardiometabolic outcomes, rather than systematically, using a hypothesis free approach, searching for potential causal effects across a wide range of health and disease outcomes. If insomnia symptoms are a cause of multimorbidity then insomnia treatments, such as cognitive behavioural therapy for Insomnia [25] recommended by UK National Institute for Health and Care Excellence [26], might be an effective means of reducing other diseases and multimorbidity, in those with insomnia.

Mendelian randomisation (MR) is a method used for testing causal relationships that generally uses genetic variants that are robustly associated with the exposure of interest as instrumental variables (IV) [27]. MR is typically less prone to confounding of the exposure-outcome association and reverse causation than conventional observational epidemiology; as genetic variation is determined at conception, it cannot be altered by disease status [28]. However, it has other potential sources of bias, in particular those due to weak instruments, confounding of the instrument-outcome association and horizontal pleiotropy [29] (the core assumptions of MR have been previously reported in detail [30]). A MR-phenome-wide association study (MR-PheWAS) is a hypothesis-free approach that tests for causal effects of a trait of interest [31] on many phenotypes [32]. To our knowledge,

only one previous study has undertaken an MR-PheWAS of insomnia symptoms [33]. In that study, the automated tool PhenoScanner [34] was used to explore causal effects of maternal insomnia symptoms on 17,503 outcomes. It identified 2844 potential causal effects (p-value < 0.05) including on adiposity, mental health, musculoskeletal, respiratory/allergic and reproductive phenotypes. However, that MR-PheWAS was part of an illustrative example in a methodological paper focused on addressing one of the MR assumptions, and none of the potential causal effects were explored further with replication or sensitivity analyses. The aim of this study is to explore the causal effects of insomnia symptoms on a wide-range of disease and health-related traits. We followed the STROBE-MR reporting guidelines when writing this paper [35] and this study was not pre-registered.

# Methods

# Study population

We used data from UK Biobank, a large prospective cohort study (dataset ID 43017 of UK Biobank application 16729, phenotypic data extracted on 24/02/2021). UK Biobank recruited 503,325 adults aged from 37 to 73 years. They were recruited between 2006 and 2007 and attended one of the 22 test centres across the UK. Of the 503,325 participants, genetic data (see Additional file 1: Text S1) was successfully obtained for 487,406 participants [36]. Participants were then excluded from this sample if they did not meet the genetic quality control [37], they were not of white-British ancestry, they were not part of the maximal subset of individuals not related to any other individual to the third degree or higher or they had since withdrawn their consent (as of 09/08/2021). The remaining 336,975 participants were included in the MR-PheWAS (See Additional file 1: Fig. S1 for a flow diagram).

# Genetic risk score

We generated a weighted genetic risk score (GRS) using 129 independent single-nucleotide polymorphisms (SNPs) previously identified [18] to associate with selfreported insomnia symptoms (answering yes to any of eight questions about insomnia diagnosis, symptoms or treatment versus answering no to all these question plus three more questions about diagnosis and treatment of collections of diseases which include insomnia—see Additional file 1: Text S2) at GWAS significance (with  $p < 5 \times 10^{-8}$ ) in 23andMe, Inc. (Additional file 2: Table S1). These data were requested from 23andMe as they were not provided in the original GWAS paper. SNPs were weighted by their per-allele association with insomnia symptoms in the original GWAS. We used a linkage disequilibrium (LD) threshold of  $R^2 > 0.001$  to clump the GWAS significant SNPs into independent SNPs. LD was calculated in the 1000 Genomes European data [38], and the TwoSampleMR (MR-base) R package v0.5.6 [39] was used to clump GWAS significant SNPs into independent SNPs. One SNP (rs28458909) was not available in UK Biobank and thus was replaced by a proxy (rs28780988) that was in close LD ( $R^2$ =1). All palindromic SNPs had an effect allele frequency falling below 0.49 or above 0.51 in UK Biobank and 23andMe and therefore could be harmonised.

As the SNPs used to construct the GRS are not replicated, there is a higher chance that spurious SNPs could have been falsely detected. We created two sensitivity analysis GRS which used SNPs which were replicated in a meta-analysis of 23andMe and UK Biobank. These analyses are only sensitivity analyses as they are at risk of overfitting due to UK Biobank being used to identify SNPs (see Additional file 1: Text S3 and Additional file 2: Table S2).

# Outcomes

A total of 11,409 outcome variables were derived and analysed using PHESANT [40]. Outcomes included those obtained from responses to baseline and follow-up questionnaires, baseline assessments such as weight, height, blood pressure and bone density measurements, followup assessments such as accelerometer measurements and a range of different scans (including brain and cardiac scans), biomarker measures from blood or urine samples and outcomes from linkage to primary and secondary care, and the national cancer and death registers. In order to summarise our overall findings from the MR-PheWAS, outcomes were assigned to categories and subcategories based on their UK Biobank category (e.g. Online followup>Mental health>Anxiety). Measurements that were not health-related outcomes were assigned to the Auxiliary Variables category. These included outcomes such as hospital administration records and procedural metrics. Individual sleep variables from the mental health and physical health categories were then reassigned to a sleep category and medication variables in the physical health category that were for mental disorders were reassigned to the mental health category. We then manually assigned outcomes in these two categories to subcategories.

# **MR-PheWAS** analysis

The PHESANT package (v1.0) was used for the MR-PheWAS. We adjusted for age at assessment, sex and the top 10 genetic principal components to control for populations stratification [41]. A complete case analysis was undertaken by PHESANT meaning participant numbers differ between outcomes and we chose to exclude

outcomes with less than 100 cases. PHESANT derives outcomes from the UK Biobank data and defines whether they are continuous, binary, ordered categorical or unordered categorical and tests the association with a trait of interest, in our case the insomnia symptoms GRS, using linear (using inverse normal rank transformed data to ensure a normal distribution), logistic, ordered logistic, and multinomial logistic regression, respectively. The results are presented as difference in mean standard deviation (SD) of inverse rank normal transformed continuous outcomes and odds ratio (OR) for categorical outcomes, per 1 SD increase in the weighted GRS. We defined *potential causal effects* as any insomnia symptoms GRS-outcome association that passed the Bonferroni-corrected significance threshold of  $4.38 \times 10^{-6}$ (0.05/11,409) in the MR-PheWAS. The less conservative false discovery rate correction was also calculated and reported but was not used to identify potential causal effects for follow-up.

# Follow-up two-sample MR

We undertook follow-up analyses using two-sample MR for all outcomes for which the association with the GRS was identified as a potential causal effect of insomnia symptoms and an appropriate GWAS could be found. The purpose of this was to confirm the reliability of the potential causal effects identified in the MR-PheWAS and to provide a causal estimate. The TwoSampleMR package (MR-base) v0.5.6 [39] was used to conduct the follow-up. It was decided a priori that outcomes included in the auxiliary variables or sleep categories would not be followed up. We conducted an automated search for relevant GWAS using pre-specified search terms for each outcome and a predetermined workflow to select the most appropriate GWAS for each outcome. First, we conducted an automated search for relevant GWAS using pre-specified search terms for each outcome. The search automatically excluded GWAS that included solely UK Biobank data, included non-European populations or stratified by sex, based on the meta-data included in the MR-Base database. Of the remaining GWAS, we excluded those that did not match a follow-up outcome on manual inspection, those for which the origins of the data used could not be determined and those that used UK Biobank or 23andMe data. If the only GWAS available for a particular outcome included UK Biobank or 23andMe data (but did not only include UK Biobank or 23andMe data), we undertook follow-up in those GWAS and report the extent of overlap between the two samples. Of the remaining GWAS, we then chose the most suitable for a given trait. This was either the most suitable

match in terms of the trait used in that GWAS or where multiple GWAS had suitable traits, we chose the one with the larger sample size. All GWAS from FinnGen were then updated to the newest version when the fifth release was added to the MR-Base database.

The two-sample MR analysis used the same 129 SNPs and SNP-insomnia symptoms associations used by the MR-PheWAS GRS [18], and the SNP-outcome associations were extracted from the GWAS for each outcome. We used the TwoSampleMR (MR-base) package for the two-sample MR analyses, which has a built-in function for harmonising SNPs between the SNP-exposure and SNP-outcome summary results (in this study so that results reflect the effects of having symptoms on outcomes for each SNP). By default, SNPs are excluded if harmonisation is not possible (e.g. if a suitable proxy cannot be found for missing SNPs or if SNPs were palindromic with allele frequencies near to 0.5). We used the inverse-variance weighted (IVW) method for our main two-sample MR analyses [42] and weighted median regression MR [43] and MR-Egger [44] as sensitivity analyses to explore potential bias due to unbalanced horizontal pleiotropy. We did not correct for multiple testing as these analyses only followed up results which had past the very conservative Bonferroni-corrected threshold used in the MR-PheWAS. All code can be found at https://github.com/MRCIEU/PHESANT-MR-PheWAS-Insomnia v1.1.

# Systematic search of previous literature

At the suggestion of a peer reviewer, we undertook a systematic search to identify published MR studies of the effect of insomnia on health outcomes. This was used to explore the extent to which the MR-PheWAS identified novel findings that have not been previously studied with MR. We searched Embase and Web of Science on 8/12/2022 for articles containing "Insomnia" AND ("Mendelian randomisation" OR "Mendelian randomization") in any field. We excluded articles which were not fully peer-reviewed original research articles or were not investigating the causal effect of insomnia on an outcome through MR. We then extracted information on the relevant analyses from each article and whether they found evidence of a causal effect.

# Results

The study population had a mean age of 57 years, 54% were female and 32% were educated to degree level (Table 1). Self-reported insomnia symptoms were common, with 48% reporting these sometimes and 28% usually.

 Table 1
 Baseline characteristics for the white-British UK Biobank

 sample of 336,975 individuals included in the MR-PheWAS

	Mean (SD) or <i>N</i> (%) <sup>a</sup>
Age at assessment centre (years)	57 (8)
Townsend area deprivation score	— 1.58 (2.93)
Sex	336,975 (100%)
Male	155,702 (46%)
Female	181,269 (54%)
Insomnia	336,744 (99.9%))
Usually	95,380 (28%)
Sometimes	160,877 (48%)
Never/Rarely	80,483 (24%)
Education	333,846 (99%)
College or university degree	106,741 (32%)
A levels/AS levels or equivalent	38,439 (11%)
O levels/GCSEs or equivalent	74,089 (22%)
CSEs or equivalent	18,114 (5%)
NVQ/HND/HNC or equivalent	22,097 (7%)
Other professional qualifications (e.g. nursing or teaching)	17,284 (5%)
None	57,078 (17%)

A level advanced level, AS level advanced subsidiary level, CSE certificate of secondary education, GCSE General Certificate of Secondary Education, HND Higher National Diploma, HNC Higher National Certificate, NVQ National Vocational Qualification, SD standard deviation

<sup>a</sup> Mean (SD) for continuous variables and number and percentage for categorical variables

# **MR-PheWAS**

The insomnia symptoms GRS was associated with an increased risk of insomnia symptoms in UK Biobank: OR of self-report of usually versus never/rarely/sometimes having trouble falling or staying asleep = 1.08 [95% Confidence Interval (CI): 1.07, 1.09] per one standard deviation higher GRS ( $p=3.59 \times 10^{-84}$ , McFadden's pseudo  $R^2=0.01$ ). See Additional file 1: Fig. S2 for the association of each SNP with insomnia symptoms.

Of the 11,409 associations included in the MR-PheWAS, 437 were identified as potential causal effects (Additional file 2: Table S3). These included anxiety, stress, depression, mania, addiction, pain, body composition, immune, respiratory, endocrine, dental, musculoskeletal, cardiovascular and reproductive traits, as well as socioeconomic and behavioural traits. Figure 1 shows the proportion of potential causal effects of insomnia symptoms by broad categories of outcomes. For associations between insomnia symptoms and mental health-related outcomes, 96 of 301 (32%) were identified as potential causal effects. There were higher proportions of these in 10 out of 17 of the mental health subcategories (Fig. 2), including depression (38%), anxiety (48%), general (33%), well-being (87%), suicide and self-harm (24%) and mania (19%). Of the physical health category, 197 out of 6451



Fig. 1 Proportion of potential causal effects of insomnia on outcomes within different categories. *n* is the total number of outcomes in the category. Additional file 2: Table S3 gives the category for each outcome. Results shown in this figure are also provided in Additional file 2: Table S4

(3%) associations with the insomnia symptoms GRS were identified as potential causal effects. Higher proportions of potential causal effects (Fig. 3) were seen for the pain (30%) and body composition (19%) subcategories. For the family and childhood category, 17 out of 96 (18%) associations were identified as potential causal effects. This category included some outcomes that could not be plausibly affected by adult insomnia and might reflect shared family (inherited) predisposition to insomnia and its potential causal effects on fertility and health-related outcomes across family members. For the lifestyle/ behaviours category, 44 out of 854 outcomes (5%) were identified as potential causal effects, while for the sociodemographic category 38 out of 1053 (4%) were. There were 2 of 2160 (0.1%) outcomes identified as potential causal effects from the brain imaging category. Alternatively, the brain/cognition category had no potential causal effects. Full details of the numbers in each category/subcategory and the numbers and percentages of outcomes in those categories that are potentially influenced by insomnia symptoms are provided in Additional file 2: Tables S4 and S5. For the results of the sensitivity analyses, see Additional file 1: Text S4, Figs. S3-S4 and Additional file 2: Table S3.

# Follow-up two-sample MR

Of the 437 potential causal effects identified in the MR-PheWAS, we identified 71 with a relevant GWAS in MR-Base [45-132], and hence eligible for follow-up (see Additional file 1: Fig. S5 and Additional file 2: Tables S6-S8). Of these, 45 outcomes showed clear evidence of an effect of being a self-reported insomnia symptoms case versus not in the IVW MR analyses, having 95% CIs which excluded the null (Figs. 4a, b and 5 and Additional file 2: Tables S9-S10). Three of these estimates (HDL cholesterol, triglycerides and absolute leukocyte count) contradicted the direction of the MR-PheWAS estimate. Of the 42 remaining, 30 (7 continuous and 23 binary) of these had effect estimates in the same direction across all main and sensitivity two-sample MR analyses although with CIs often including the null. These 30 outcomes include a range of categories: substance use and mental



**Fig. 2** Proportion of potential causal effects of insomnia on outcomes within different mental health subcategories. *n* is the total number of outcomes in the category. Additional file 2: Table S3 gives the subcategory for each outcome. Results shown in this figure are also provided in Additional file 2: Table S5

health-related outcomes such as acute alcohol intoxication, mental and behavioural disorders due to tobacco, neuroticism, anxiety disorder and post-traumatic stress disorder; body composition outcomes such as obesity, body fat percentage, body mass index, hip circumference and waist circumference; musculoskeletal outcomes such as low back pain, gonarthrosis, unspecified arthrosis, unspecified joint disorders, shoulder lesions, unspecified soft tissue disorders, spondylosis and dorsalgia; digestive health-related outcomes such as irritable bowel syndrome, diverticular disease of intestine, unspecified gastritis (including duodenitis), gastro-oesophageal reflux disease, diaphragmatic hernia and oesophagitis; allergy or respiratory outcomes such as allergic disease (asthma, hay fever or eczema), asthma and bronchitis; and outcomes which were not related to others in the set such as unspecified headache syndromes, C-reactive protein level and HbA1c. Cochran's Q showed evidence of between SNP heterogeneity (p < 0.05) in both the IVW and MR-Egger analyses for 16 of these 30 outcomes: Anxiety, asthma, obesity, body mass index, body fat percentage, hip circumference, waist circumference, C-reactive protein level, unspecified arthrosis, unspecified joint disorders, unspecified soft tissue disorders, shoulder lesions, low back pain, gonarthrosis, dorsalgia and allergic disease. Only anxiety disorders showed evidence of unbalanced horizontal pleiotropy in the MRegger intercept, implying that heterogeneity in most SNP estimates is due to either balanced pleiotropy or different causal biological mechanisms of the SNP on insomnia symptoms.

# Systematic search of previous literature

After deduplication, abstract review and full-text review, 81 articles exploring the effect of insomnia on a health outcome via MR were identified in the systematic search (see Additional file 1: Fig. S6). Article information and a summary of the findings for each article included can be seen in Additional file 2: Table S11 (while information for articles excluded at full-text screening with the



**Fig. 3** Proportion of potential causal effects of insomnia on outcomes within different physical health subcategories. *n* is the total number of outcomes in the category. Additional file 2: Table S3 gives the subcategory for each outcome. Results shown in this figure are also provided in Additional file 2: Table S5

reason for exclusion can be seen in Additional file 2: Table S12). These articles showed evidence that insomnia may have causal effects on anxiety, neuroticism, posttraumatic stress disorder, subjective well-being, depressive symptoms, major mood disorder, a range of cardiovascular outcomes (including coronary heart disease, angina pectoris and hypertension), type 2 diabetes mellitus, cholesterol levels, body mass, osteoarthritis, rheumatoid arthritis, pain, migraine, gastro-oesophageal reflux disease, irritable bowel syndrome, miscarriage, allergic disease, asthma, smoking and alcohol use, among others. Of the 30 directionally consistent findings across the MR-PheWAS, two-sample follow-up MR and two-sample sensitivity analyses (for which the 95% CI excluded the null in the MR-PheWAS and the IVW two-sample follow-up), only spondylosis, unspecified joint disorders, shoulder lesions, unspecified soft-tissue disorders, gastritis (including duodenitis), oesophagitis, diverticular disease of intestine, diaphragmatic hernia, bronchitis, unspecified headache syndromes and C-reactive protein levels were not supported by previous MR literature (i.e. no clear evidence of a concordant evidence in the previous literature). While the systematic search identified no papers investigating the effects of insomnia on acute alcohol intoxication, mood and behavioural disorders due to tobacco, certain body composition outcomes and gonarthrosis (arthrosis of the knee) specifically, there was evidence for closely related and overlapping outcomes in the previous literature.

# Discussion

In this study, we conducted an MR-PheWAS of insomnia symptoms using 11,409 outcome variables. Of these GRS-outcome associations, 437 met our criteria for being potential causal effects, of which 71 were possible to follow-up using two-sample MR. Follow-up analyses showed consistent evidence of an adverse causal effect of insomnia symptoms on 30 outcomes including those related to anxiety disorders, respiratory disorders, musculoskeletal disorders, disorders of the digestive system

a			1 15 (1 01 1 30)	
Outcome	Body Composition/Metabolic	Disorders of lipoprotein metabolism and other lipidaemias	0.90 (0.59, 1.39) 1.09 (0.91, 1.30)	
		Obesity-	1.22 (1.02, 1.47) 1.35 (0.73, 2.53) 1.16 (0.91, 1.48)	
		Type 2 diabetes adjusted for BMI*-	0.98 (0.69, 1.40) 1.44 (0.36, 5.73) 1.07 (0.83, 1.38)	
		Angina pectoris-	1.08 (0.94, 1.25) 0.77 (0.47, 1.24) 0.97 (0.80, 1.17)	
		Hypertension, essential-	1.07 (0.96, 1.20) 0.98 (0.68, 1.43) 1.03 (0.90, 1.18)	
		Acute gastritis-	0.97 (0.69, 1.37) 1.01 (0.32, 3.23) 0.97 (0.58, 1.62)	
		Benign neoplasm: Colon-	1.13 (0.96, 1.32) 1.00 (0.58, 1.71) 1.24 (0.99, 1.55)	
		Cholelithiasis-	1.18 (1.04, 1.33) 0.96 (0.63, 1.48) 1.10 (0.93, 1.30)	
		Chronic gastritis -	1.16 (0.98, 1.39) 0.87 (0.48, 1.59) 1.05 (0.81, 1.35)	
		Diaphragmatic hernia-	1.35 (1.14, 1.61) 1.24 (0.68, 2.23) 1.33 (1.03, 1.72)	
	Digestive	Diseases of liver-	1.07 (0.82, 1.38) 0.78 (0.33, 1.87) 0.88 (0.60, 1.28)	
		Diseases of oesophagus, stomach and duodenum	1.27 (1.17, 1.39) 0.95 (0.72, 1.27) 1.24 (1.10, 1.40)	
		Diverticular disease of intestine -	1.17 (1.04, 1.32) 1.01 (0.67, 1.52) 1.16 (0.97, 1.38)	
		Gastritis incl. Duodenitis-	1.29 (1.06, 1.58) 1.25 (0.63, 2.48) 1.38 (1.02, 1.86)	
		Gastro-oesophageal reflux disease-	1.40 (1.25, 1.58) 1.01 (0.68, 1.50) 1.33 (1.12, 1.58)	
		Irritable bowel syndrome-	1.30 (1.08, 1.57) 1.47 (0.79, 2.76) 1.44 (1.09, 1.89)	
		Malignant neoplasm of colon-	1.07 (0.80, 1.43) 0.63 (0.23, 1.69) 1.10 (0.72, 1.69)	
		Oesophagitis-	1.99 (1.27, 3.12) 1.05 (0.23, 4.80) 2.07 (1.02, 4.20)	
	Headache	Headache syndromes-	1.23 (1.06, 1.42) 1.32 (0.81, 2.15) 1.30 (1.04, 1.61)	
	neauache	Migraine -	1.30 (1.09, 1.55) 0.93 (0.51, 1.69) 1.25 (0.99, 1.59)	
	Immune	Allergic disease asthma, hay fever or eczema*	1.20 (1.12, 1.30) 1.50 (1.14, 1.98) 1.10 (1.01, 1.19)	
		Allergic rhinitis	1.11 (0.91, 1.35) 0.64 (0.33, 1.22) 1.08 (0.83, 1.40)	
		-	1 0	1 2 3 4 5
				Odds Ratio (95% CI)

- Inverse variance weighted - MR Egger - Weighted median

Fig. 4 a, b Two-sample MR results of the effect (odds ratio), comparing genetically predicted self-reported insomnia cases versus non-cases for binary outcomes. \*GWAS has overlap with UK Biobank or 23andMe



Fig. 4 continued

and body composition measurements. A number of these had not previously been investigated using MR. These included respiratory disorders, soft-tissue disorders and digestive disorders. Together with the potential causal effects that we were not able to follow-up, these findings support a role for insomnia symptoms in multimorbidity. The findings also suggest that effective insomnia treatments, such as the cognitive behavioural therapy-insomnia [25], which has been shown to be an effective treatment for depression when comorbid with insomnia [133], could



Fig. 5 Two-sample MR results of the effect (mean difference), comparing genetically predicted self-reported insomnia cases versus non-cases, for continuous outcomes. \*GWAS has overlap with UK Biobank or 23andMe

be used to treat a range of other adverse health-related outcomes; however, this requires further investigation.

We found evidence (which was directionally consistent across the MR-PheWAS, the two-sample follow-up and

the two-sample sensitivity analyses, and for which the 95% CIs excluded the null in the former two) for a number of outcomes which have not been explored in MR research. These outcomes were spondylosis, unspecified

joint disorders, shoulder lesions, unspecified soft-tissue disorders, gastritis (including duodenitis), oesophagitis, diverticular disease of intestine, diaphragmatic hernia, bronchitis, unspecified headache disorders and C-reactive protein levels. The bidirectional relationship between insomnia and headache has been extensively researched in previous non-MR literature [134]. Furthermore, a positive association between insomnia and C-reactive protein levels has previously been shown in standard observational research [135]. C-reactive protein is a marker of inflammation which is itself a response of the immune system, providing evidence that insomnia may affect the immune system. The relationship between insomnia and the other outcomes has not been extensively researched in conventional epidemiology studies and these are, therefore, novel findings. However, diaphragmatic hernia, is a birth defect and so it is implausible this could be caused by insomnia, indicating the results are subject to violations of the core assumptions.

# Strengths and limitations

A key strength of our hypothesis-free MR-PheWAS is that it allows for many potential novel causal effects of insomnia symptoms to be identified. Furthermore, we used two-sample MR to follow up as many of the potential causal effects as possible and included sensitivity analyses to explore potential bias due to horizontal pleiotropy.

Limitations include variations in power due to the differing numbers of samples and cases across UK Biobank phenotypes meaning our MR-PheWAS analyses may have been underpowered for some outcomes. For the two-sample MR analyses, sample sizes ranged between 1000 and 360,838 for the outcome GWASs. With larger sample sizes, more precise estimates may have been obtained. Also, 366 (84%) potential causal effects could not be followed up because we were unable to identify suitable summary GWAS data in MR-Base. It is possible that for some outcomes, suitable GWASs may exist but may not have been added to MR-Base or may have become available after the search was conducted. As GWASs are conducted for a wider range of outcomes and GWASs increase in size, future research should explore avenues not currently explored in our follow-up and update the current analyses to increase power. We did update all FinnGen GWASs to the most recent versions which were released after the search for GWASs and screening was completed, but did not search for new GWASs specifically. In the two-sample MR follow-up, there was overlap between a number of the outcome GWASs and the exposure GWAS. This has the potential to bias the results away from the null; however, previous research has suggested sample overlap often does not have a large effect [136].

It is possible that some of the potential causal effects of insomnia that we have identified are driven by the health outcome in question causally influencing insomnia [33]. As GWASs get larger, they are more likely to identify genome-wide significant associations for phenotypes that are downstream of other healthrelated factors. For example, previous MR studies have shown that depression affects insomnia [12, 18], and a large GWAS of insomnia might identify statistically robust SNPs associated with insomnia, some of which are identified because of the relationship of depression with insomnia. Given the number of outcomes explored in this study, investigating reverse causality is left to future work. It is also possible that the results are subject to horizontal pleiotropy. In our two-sample follow-up, we used sensitivity analyses to explore bias due to unbalanced horizontal pleiotropy. These methods do not look at specific hypothesised pleiotropic paths but rather help to see whether pleiotropic paths might have biased estimates.

The questionnaires that were used in the GWAS that provided our genetic instruments are widely used in observational studies. They reflect a person's subjective reporting of symptoms, which may not be consistent of a diagnosis of insomnia. That said clinical diagnostic codes misclassify an important number who would meet diagnostic criteria as not everyone with symptoms will seek clinical help and not all of those who do will be diagnosed in the same way [137]. Furthermore, there may be differences in the health effects of short- and long-term insomnia and the insomnia definition used in the GWAS does not acknowledge the length of time the symptoms have been experienced, only whether they are present or not. Also, the non-representativeness of UK Biobank may also bias the results. Finally, it is important to note our presentation of MR-PheWAS results as proportions of potential causal effects in different phenotypic categories, which, although a useful summary, may be misleading if the correlations within each category differs across categories.

# Conclusions

Our results suggest that insomnia symptoms may have broad effects on health. In particular, we identified novel effects (that replicated in follow-up analyses) on respiratory disorders, soft-tissue disorders and digestive disorders and confirmed previously identified effects on mental health, hyperglycaemia, pain and body composition outcomes. These findings support a role for insomnia symptoms in multimorbidity and the possibility that effective insomnia treatments should be integrated into the treatment of other diseases. Future research should follow up individual outcomes in greater depth, including novel methods being developed for time-varying exposures and non-linear associations, to confirm novel findings.

#### Abbreviations

CI	Confidence interval
GRS	Genetic risk score
GWAS	Genome-wide association study
IVW	Inverse-variance weighted
LD	Linkage disequilibrium
MR	Mendelian randomisation
OR	Odds ratio
PheWAS	Phenome-wide association study
SD	Standard deviation
SNP	Single-nucleotide polymorphism

# **Supplementary Information**

The online version contains supplementary material available at https://doi. org/10.1186/s12916-023-02832-8.

Additional file 1: Text S1. UK Biobank Sample. Text S2. Insomnia Phenotype. Text S3. Sensitivity Analyses Methods. Text S4. Sensitivity Analysis Results. Figure S1. Flow chart of participant inclusion. Figure S2. Odds ratio and 95% confidence interval for association between each SNP used in the main GRS and insomnia in UK Biobank (Field 1200, with an answer of "usually" coded as an insomnia case). Figure S3. Odds ratio and 95% confidence interval for association between each SNP used in the S1 and S2 GRS and insomnia in UK Biobank (Field 1200, with an answer of "usually" coded as an insomnia case). Figure S4. Venn diagram of the number of GRS-outcome associations which passed the Bonferroni-corrected signifcance threshold for each MR-PheWAS (the percentages are with respect to the total number of associations (542) identified across all MR-pheWAS). Figure S5. Flow chart of GWAS inclusion for follow-up. Figure S6. Prisma style flow chart for article screening in systematic search.

Additional file 2: Table S1. GWAS significant SNPs in 23&Me used to construct GRS for PheWAS and used to conduct two-sample MR followup. \*rs28780988 used as a proxy for rs28458909 which was identified as an independent GWAs significant SNP in the clumping of the 23andMe GWAS results but was not available in UK biobank. Table S2. SNPs which were GWAS significant in both UK Biobank/23andMe meta-analysis and 23&Me, used to construct GRS for sensitivity PheWAS analyses. Table S3. Results from the MR-PheWAS and sensitivity analysis 1 and 2 for each outcome (Ordered by the p-value from the main MR-PheWAS). For linear regressions the beta is the mean difference per one standard deviation increase in GRS and for all others the beta is the odds ratio per one standard deviation increase in GRS. Table S4. Quantified details of the total number in each category, number and percentage of outcomes reaching criteria for potential causal effect in each category. Table S5. Quantified details of the total number in each subcategory, number and percentage of outcomes reaching criteria for potential causal effect in each subcategory. Table S6. Follow-up information for associations with Insomnia that passed the Bonferroni-corrected significance threshold in the main MR-PheWAS. Table S7. List of GWAS from TwoSampleMR package v0.5.6 included in follow-up. The potential causal effects from the MR-PheWAS these relate to are in the Outcome column separated by semicolons. Table S8. List of GWAS from TwoSampleMR package v0.5.6 returned in search but not included in follow-up. The Reason column gives the reason for exclusion. Table S9. Results from two-sample MR follow-up for binary outcomes. Table S10. Results from two-sample MR follow-up for continuous outcomes. Table S11. Articles identified in systematic search and included after screening, with a summary of findings. Table S12. Articles identified in systematic search and excluded at full text screening.

#### Acknowledgements

This research was conducted using the UK Biobank resource under application number 16729. This research also used data supplied by 23andMe under a confidentiality agreement. We would like to thank the research participants and employees of 23andMe, Inc. for making this work possible.

### Authors' contributions

MJG, DAL and LACM were all involved in the conception and planning of the study. MJG conducted the analysis and wrote the first version of the manuscript. DAL and LACM supervised the project. All authors critically reviewed and revised the manuscript.

#### Funding

This work was supported by a Medical Research Council (MRC) PhD studentship to MJG (grant code: MC\_UU\_00011/7), Diabetes UK (17/0005700), the MRC (MR/V033867/1), and the British Heart Foundation (AA/18/1/34219). DAL is further supported by a British Heart Foundation Chair (CH/F/20/90003) and National Institute of Health Research Senior Investigator award (NF-0616–10102). LACM is supported by a University of Bristol Vice-Chancellor's fellowship. All three authors work in a unit that is funded by the University of Bristol and Medical Research Council (MC\_UU\_00011/1, MC\_UU\_00011/6 and MC\_UU\_00011/7).

The funders had no role in the study design, collection or analysis of data or interpretation of results. The views expressed in this paper are those of the authors and not necessarily any funder or acknowledged person/institution.

#### Availability of data and materials

All data is available on request from the UK Biobank or 23andMe. The full GWAS summary statistics for the 23andMe discovery data set will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit https://research.23andme.com/collaborate/#dataset-access/ for more information and to apply to access the data. All code used in the analyses is available at https://github.com/MRCIEU/PHESANT-MR-PheWAS-Insomnia.

# Declarations

# Ethics approval and consent to participate

The data collection in UK Biobank was approved by the NHS National Research Ethics Service (ref 11/NW/0382). All data-sets used in this data obtained fully informed consent from participants.

# **Consent for publication**

Not applicable.

#### **Competing interests**

All authors have completed the ICMJE uniform disclosure form at http://www. icmje.org/disclosure-of-interest/. DAL has received support from Roche Diagnostics and Medtronic Ltd for biomarker research unrelated to this paper. MJG and LACM declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Received: 2 November 2022 Accepted: 13 March 2023 Published online: 03 April 2023

#### References

- Ramanathan L, Gulyani S, Nienhuis R, Siegel JM. Sleep deprivation decreases superoxide dismutase activity in rat hippocampus and brainstem. NeuroReport. 2002;13(11):1387–90.
- Morin A, Carrier J, Dostie V, Doyon J. Differences in time- and sleepdependent learning and memory consolidation of motor sequence and visuo-motor adaptation skills. Sleep. 2005;28:A355-A.
- 3. Stickgold R. Sleep-dependent memory consolidation. Nature. 2005;437(7063):1272–8.

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. 5th ed. Arlington: American Psychiatric Association; 2013.
- Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jonsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharm. 2011;21(9):655–79.
- Morphy H, Dunn KM, Lewis M, Boardman HF, Croft PR. Epidemiology of insomnia: a longitudinal study in a UK population. Sleep. 2007;30(3):274–80.
- Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. Sleep Med Rev. 2002;6(2):97–111.
- Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet. 2012;380(9836):37–43.
- Jani BD, Hanlon P, Nicholl BI, McQueenie R, Gallacher KI, Lee D, et al. Relationship between multimorbidity, demographic factors and mortality: findings from the UK Biobank cohort. BMC Med. 2019;17:74.
- Whitty CJM, MacEwen C, Goddard A, Alderson D, Marshall M, Calderwood C, et al. Rising to the challenge of multimorbidity. BMJ. 2020;368:16964.
- Baglioni C, Battagliese G, Feige B, Spiegelhalder K, Nissen C, Voderholzer U, et al. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. J Affect Disorders. 2011;135(1–3):10–9.
- Cai L, Bao YR, Fu XQ, Cao HB, Baranova A, Zhang XR, et al. Causal links between major depressive disorder and insomnia: a Mendelian randomisation study. Gene. 2021;768:145271.
- Pasman JA, Smit DJA, Kingma L, Vink JM, Treur JL, Verweij KJH. Causal relationships between substance use and insomnia. Drug Alcohol Depend. 2020;216:108151.
- Gibson M, Munafo MR, Taylor AE, Treur JL. Evidence for genetic correlations and bidirectional, causal effects between smoking and sleep behaviors. Nicotine Tob Res. 2019;21(6):731–8.
- Gao X, Meng LX, Ma KL, Liang J, Wang H, Gao Q, et al. The bidirectional causal relationships of insomnia with five major psychiatric disorders: a Mendelian randomization study. Eur Psychiat. 2019;60:79–85.
- de Almondes KM, Costa MV, Malloy-Diniz LF, Diniz BS. Insomnia and risk of dementia in older adults: systematic review and meta-analysis. J Psychiatr Res. 2016;77:109–15.
- Cappuccio FP, D'Elia L, Strazzlillo P, Miller MA. Quantity and quality of sleep and incidence of type 2 diabetes - a systematic review and metaanalysis. Diabetes Care. 2010;33(2):414–20.
- Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR, et al. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. Nat Genet. 2019;51(3):394–403.
- Meng L, Zheng Y, Hui RT. The relationship of sleep duration and insomnia to risk of hypertension incidence: a meta-analysis of prospective cohort studies. Hypertens Res. 2013;36(11):985–95.
- Li M, Zhang XW, Hou WS, Tang ZY. Insomnia and risk of cardiovascular disease: a meta-analysis of cohort studies. Int J Cardiol. 2014;176(3):1044–7.
- Sofi F, Cesari F, Casini A, Macchi C, Abbate R, Gensini GF. Insomnia and risk of cardiovascular disease: a meta-analysis. Eur Heart J. 2012;33:947.
- 22 Larsson SC, Markus HS. Genetic liability to insomnia and cardiovascular disease risk. Circulation. 2019;140(9):796–8.
- Broberg M, Karjalainen J, FinnGen, Ollila HM. Mendelian randomization highlights insomnia as a risk factor for pain diagnoses. Sleep. 2021;44(7):zsab025.
- Bos MM, Goulding NJ, Lee MA, Hofman A, Bot M, Pool R, et al. Investigating the relationships between unfavourable habitual sleep and metabolomic traits: evidence from multi-cohort multivariable regression and Mendelian randomization analyses. BMC Med. 2021;19:69.
- 25. Trauer JM, Qian MY, Doyle JS, Rajaratnam SMW, Cunnington D. Cognitive behavioral therapy for chronic insomnia: a systematic review and meta-analysis. Ann Intern Med. 2015;163(3):191–204.
- 26. The National Institute for Health and Care Excellence: Insomnia. https:// cks.nice.org.uk/topics/insomnia/. Accessed 09 Mar 2023.

- 27. Davey Smith G, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32(1):1–22.
- Davey Smith G, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. Plos Med. 2007;4(12):1985–92.
- 29. Davey Smith G, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. Int J Epidemiol. 2004;33(1):30–42.
- Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018;362:k601.
- Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al. PheWAS: demonstrating the feasibility of a phenomewide scan to discover gene-disease associations. Bioinformatics. 2010;26(9):1205–10.
- Millard LAC, Davies NM, Timpson NJ, Tilling K, Flach PA, Smith GD. MR-PheWAS: hypothesis prioritization among potential causal effects of body mass index on many outcomes, using Mendelian randomization. Sci Rep. 2015;5:16645.
- Yang Q, Sanderson E, Tilling K, Borges MC, Lawlor DA. Exploring and mitigating potential bias when genetic instrumental variables are associated with multiple non-exposure traits in Mendelian randomization. Eur J Epidemiol. 2022;37:683–700.
- 34 Staley JR, Blackshaw J, Kamat MA, Ellis S, Young R, Butterworth AS. PhenoScanner: a database of human genotype-phenotype associations. Genet Epidemiol. 2016;40(7):664.
- Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. BMJ. 2019;375:n2233.
- Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203–9.
- Mitchell RE, Hemani G, Dudding T, Corbin L, Harrison S, Paternoster L. UK Biobank Genetic Data: MRC-IEU Quality Control, version 2, 18/01/2019.. 2019.
- Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68–74.
- Hemani G, Zhengn J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. eLife. 2018;7:e34408.
- Millard LAC, Davies NM, Gaunt TR, Smith GD, Tilling K. Software Application Profile: PHESANT: a tool for performing automated phenome scans in UK Biobank. Int J Epidemiol. 2018;47(1):29–35.
- 41. Mitchell RE, Elsworth B, Mitchell R, Raistrick C, Paternoster L, Hemani G, et al. MRC IEU UK Biobank GWAS pipeline version 2. 2019.
- 42. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658–65.
- Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol. 2016;40(4):304–14.
- 44. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44(2):512–25.
- 45. Teslovich T. HDL cholesterol. OpenGWAS. (2010). https://gwas.mrcieu. ac.uk/datasets/ebi-a-GCST000755/
- Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010;466(7307):707–13.
- Lu Y. Body fat percentage. OpenGWAS. (2016). https://gwas.mrcieu.ac. uk/datasets/ebi-a-GCST003435/.
- Lu YC, Day FR, Gustafsson S, Buchkovich ML, Na JB, Bataille V, et al. New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. Nat Commun. 2016;7:10495.
- 49. Okbay A. Subjective well-being. OpenGWAS. (2016). https://gwas. mrcieu.ac.uk/datasets/ebi-a-GCST003766/.

- Okbay A, Baselmans BML, De Neve JE, Turley P, Nivard MG, Fontana MA, et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses (vol 48, pg 624, 2016). Nat Genet. 2016;48(12):1591.
- 51. Astle W. High light scatter reticulocyte count. OpenGWAS. (2016). https://gwas.mrcieu.ac.uk/datasets/ebi-a-GCST004611/.
- Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The allelic landscape of human blood cell trait variation and links to common complex disease. Cell. 2016;167(5):1415–29.
- Astle W. High light scatter reticulocyte percentage of red cells. OpenG-WAS. (2016). https://gwas.mrcieu.ac.uk/datasets/ebi-a-GCST004612/
- Astle W. Reticulocyte fraction of red cells. OpenGWAS. (2016). https:// gwas.mrcieu.ac.uk/datasets/ebi-a-GCST004619/.
- Astle W. Reticulocyte count. OpenGWAS. (2016). https://gwas.mrcieu.ac. uk/datasets/ebi-a-GCST004622/
- Astle W. Immature fraction of reticulocytes. OpenGWAS. (2016). https:// gwas.mrcieu.ac.uk/datasets/ebi-a-GCST004628/
- 57. Ferreira M. Allergic disease (asthma, hay fever or eczema). OpenGWAS. (2017). https://gwas.mrcieu.ac.uk/datasets/ebi-a-GCST005038/
- Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. Nat Genet. 2017;49(12):1752–7.
- Prins B. Serum alkaline phosphatase levels. OpenGWAS. (2017). https:// gwas.mrcieu.ac.uk/datasets/ebi-a-GCST005061/
- 60. Prins BP, Kuchenbaecker KB, Bao YC, Smart M, Zabaneh D, Fatemifar G, et al. Genome-wide analysis of health-related biomarkers in the UK Household Longitudinal Study reveals novel associations. Sci Rep. 2017;7:11008.
- 61. Demenais F. Asthma. OpenGWAS. (2017). https://gwas.mrcieu.ac.uk/ datasets/ebi-a-GCST006862/
- Demenais F, Margaritte-Jeannin P, Barnes KC, Cookson WOC, Altmuller J, Ang W, et al. Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. Nat Genet. 2018;50(1):42–53.
- 63. Mahajan A. Type 2 diabetes (adjusted for BMI). OpenGWAS. (2018). https://gwas.mrcieu.ac.uk/datasets/ebi-a-GCST007516/.
- Mahajan A, Wessel J, Willems SM, Zhao W, Robertson NR, Chu AY, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. Nat Genet. 2018;50(4):559–71.
- 65. Orru V. Leukocyte Absolute Count. OpenGWAS. (2020). https://gwas. mrcieu.ac.uk/datasets/ebi-a-GCST90001600/.
- 66. Orru V, Steri M, Sidore C, Marongiu M, Serra V, Olla S, et al. Complex genetic signatures in immune cells underlie autoimmunity and inform therapy. Nat Genet. 2020;52(10):1036–45.
- 67. Kettunen. Apolipoprotein A-I. OpenGWAS. (2016). https://gwas.mrcieu. ac.uk/datasets/met-c-842/.
- Kettunen J, Demirkan A, Wurtz P, Draisma HHM, Haller T, Rawal R, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun. 2016;7:11122.
- Sun B. Apolipoprotein A-V. OpenGWAS. (2018). https://gwas.mrcieu.ac. uk/datasets/prot-a-125/.
- Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. Nature. 2018;558(7708):73–9.
- 71. Sun B. Alanine aminotransferase 1. OpenGWAS. (2018). https://gwas. mrcieu.ac.uk/datasets/prot-a-1264/.
- Suhre K. Cystatin C. OpenGWAS. (2019). https://gwas.mrcieu.ac.uk/datas ets/prot-c-2609\_59\_2/.
- Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. Nat Commun. 2017;8:14357.
- Moor d. Neuroticism. OpenGWAS. (2014). https://gwas.mrcieu.ac.uk/ datasets/ieu-a-118/.
- de Moor MHM, van den Berg SM, Verweij KJH, Krueger RF, Luciano M, Vasquez AA, et al. Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. JAMA Psychiat. 2015;72(7):642–50.
- Shungin D. Hip circumference. OpenGWAS. (2015). https://gwas. mrcieu.ac.uk/datasets/ieu-a-49/.
- Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Lockes AE, Magi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature. 2015;518(7538):187-U378.

- Shungin D. Waist circumference. OpenGWAS. (2015). https://gwas. mrcieu.ac.uk/datasets/ieu-a-61/.
- Locke A. Body mass index. OpenGWAS https://gwas.mrcieu.ac.uk/ datasets/ieu-a-835/ (2015).
- Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Felix R, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197-U401.
- Rietveld C. College completion. OpenGWAS. (2013). https://gwas. mrcieu.ac.uk/datasets/ieu-a-836/.
- Rietveld CA, Medland SE, Derringer J, Yang J, Esko T, Martin NW, et al. GWAS of 126,559 Individuals identifies genetic variants associated with educational attainment. Science. 2013;340(6139):1467–71.
- Soranzo N. HbA1C. OpenGWAS. (2010). https://gwas.mrcieu.ac.uk/ datasets/ieu-b-103/.
- Soranzo N, Sanna S, Wheeler E, Gieger C, Radke D, Dupuis J, et al. Common variants at 10 genomic loci influence hemoglobin A(1C) levels via glycemic and nonglycemic pathways. Diabetes. 2010;59(12):3229–39.
- Ligthart S. C-Reactive protein level. OpenGWAS. (2018). https://gwas. mrcieu.ac.uk/datasets/ieu-b-35/.
- Ligthart S, Vaez A, Vosa U, Stathopoulou MG, de Vries PS, Prins BP, et al. Genome analyses of >200,000 individuals identify 58 loci for chronic inflammation and highlight pathways that link inflammation and complex disorders. Am J Hum Genet. 2018;103(5):691–706.
- 87. Acute alcohol intoxication (ALCOHOLACUTE10). OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-ALCOHOLACUTE10/.
- Mental and behavioural disorders due to alcohol, excluding acute intoxication. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/ finn-b-ALCOHOLMENTAL/.
- Allergic rhinitis. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datas ets/finn-b-ALLERG\_RHINITIS/.
- 90. Depression medications. OpenGWAS. (2021). https://gwas.mrcieu.ac. uk/datasets/finn-b-ANTIDEPRESSANTS/.
- Bronchitis. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/ finn-b-BRONCHITIS/.
- Malignant neoplasm of colon. OpenGWAS. (2021). https://gwas. mrcieu.ac.uk/datasets/finn-b-C3\_COLON/.
- 93. Benign neoplasm: colon. OpenGWAS. (2021). https://gwas.mrcieu.ac. uk/datasets/finn-b-CD2\_BENIGN\_COLON/.
- Disorders of lipoprotein metabolism and other lipidaemias. OpenG-WAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-E4\_LIPOPROT/.
- Obesity. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finnb-E4\_OBESITY/.
- 96. Bipolar affective disorders. OpenGWAS. (2021). https://gwas.mrcieu. ac.uk/datasets/finn-b-F5\_BIPO/.
- Depression. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/ finn-b-F5\_DEPRESSIO/.
- Post-traumatic stress disorder. OpenGWAS. (2021). https://gwas. mrcieu.ac.uk/datasets/finn-b-F5\_PTSD/.
- Mental and behavioural disorders due to tobacco. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-F5\_TOBAC/.
- 100. Other headache syndromes. OpenGWAS. (2021). https://gwas.mrcieu. ac.uk/datasets/finn-b-G6\_HEADACHE/.
- Migraine. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/ finn-b-G6\_MIGRAINE/.
- Other extrapyramidal and movement disorders+ in other diseases. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-G6\_ XTRAPYROTH/.
- Tinnitus. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finnb-H8\_TINNITUS/.
- 104. Angina pectoris. OpenGWAS. (2021) https://gwas.mrcieu.ac.uk/datas ets/finn-b-I9\_ANGINA/.
- Hypertension, essential. OpenGWAS. (2021). https://gwas.mrcieu.ac. uk/datasets/finn-b-I9\_HYPTENSESS/.
- Unspecified acute lower respiratory infection. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-J10\_ACUTELOWERNAS/.
- Other chronic obstructive pulmonary disease (J10\_COPDNAS). OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-J10\_ COPDNAS/.
- Emphysema. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/ finn-b-J10\_EMPHYSEMA/.

- 109. Acute gastritis. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datas ets/finn-b-K11\_ACUTGASTR/.
- 110. Cholelithiasis. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-K11\_CHOLELITH/.
- 111. Chronic gastritis. OpenGWAS. (2021) https://gwas.mrcieu.ac.uk/datasets/finn-b-K11\_CHRONGASTR/.
- 112. Diaphragmatic hernia. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/ datasets/finn-b-K11\_DIAHER/.
- 113. Other diseases of liver. OpenGWAS. (2021). https://gwas.mrcieu.ac. uk/datasets/finn-b-K11\_DISLIVOTH/.
- 114. Diverticular disease of intestine. OpenGWAS. (2021). https://gwas. mrcieu.ac.uk/datasets/finn-b-K11\_DIVERTIC/.
- 115. Irritable bowel syndrome. OpenGWAS. (2021). https://gwas.mrcieu.ac. uk/datasets/finn-b-K11\_IBS/.
- 116. Oesophagitis. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datas ets/finn-b-K11\_OESITIS/.
- 117. Diseases of oesophagus, stomach and duodenum. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-K11\_OESSTODUO/.
- 118. Other gastritis (incl. Duodenitis). OpenGWAS. (2021). https://gwas. mrcieu.ac.uk/datasets/finn-b-K11\_OTHGASTR/.
- 119. Gastro-oesophageal reflux disease. OpenGWAS. (2021). https://gwas. mrcieu.ac.uk/datasets/finn-b-K11\_REFLUX/.
- 120. Gonarthrosis [arthrosis of knee](FG). OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-KNEE\_ARTHROSIS/.
- Anxiety disorders. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/ datasets/finn-b-KRA\_PSY\_ANXIETY/.
- 122. Other arthrosis. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datas ets/finn-b-M13\_ARTHROSIS\_OTH/.
- Dorsalgia. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/ finn-b-M13\_DORSALGIA/.
- 124. Other enthesopathies. OpenGWAS. (2021). https://gwas.mrcieu.ac. uk/datasets/finn-b-M13\_ENTESOPATHYOTH/.
- 125. Low back pain. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datas ets/finn-b-M13\_LOWBACKPAIN/.
- Other joint disorders. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/ datasets/finn-b-M13\_OTHERJOINT/.
- 127. Shoulder lesions. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datas ets/finn-b-M13\_SHOULDER/.
- 128. Other soft tissue disorders, not elsewhere classified. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/finn-b-M13\_SOFTTISSUEOTH/.
- 129. Spondylosis. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/ finn-b-M13\_SPONDYLOSIS/.
- 130. Pain in thoracic spine. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/ datasets/finn-b-M13\_THORACISPINEPAIN/.
- Other arthritis (FG). OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/ datasets/finn-b-RHEU\_ARTHRITIS\_OTH/.
- Tobacco use. OpenGWAS. (2021). https://gwas.mrcieu.ac.uk/datasets/ finn-b-Z21\_TOBAC\_USE/.
- Cunningham JEA, Shapiro CM. Cognitive behavioural therapy for insomnia (CBT-I) to treat depression: a systematic review. J Psychosom Res. 2018;106:1–12.
- Uhlig B, Engstrøm M, Ødegård S, Hagen K, Sand T. Headache and insomnia in population-based epidemiological studies. Cephalalgia. 2014;34(10):745–51.
- Slavish DC, Graham-Engeland JE, Engeland CG, Taylor DJ, Buxton OM. Insomnia symptoms are associated with elevated C-reactive protein in young adults. Psychol Health. 2018;33(11):1396–415.
- Minelli C, Del Greco FM, Van Der Plaat DA, Bowden J, Sheehan NA, Thompson J. The use of two-sample methods for Mendelian randomization analyses on single large datasets. Int J Epidemiol. 2021;50(5):1651–9.
- 137. Stores G. Clinical diagnosis and misdiagnosis of sleep disorders. J Neurol Neurosur Ps. 2007;78(12):1293–7.

# **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

#### At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

