

睡眠与冠心病的孟德尔随机化因果关联研究

高雪¹ 杨晓晨² 孟玲先¹ 孙鹤立¹ 王彤¹

¹山西医科大学公共卫生学院卫生统计教研室,太原 030001; ²中国人民大学统计学
院生物与医学统计学教研室,北京 100872

通信作者:王彤, Email: tongwang@sxmu.edu.cn

【摘要】目的 对睡眠与冠心病之间的因果关联进行研究与探讨。**方法** 研究包含5个睡眠相关性状[睡眠时间(连续变量)、长睡眠时间(二分类变量)、短睡眠时间(二分类变量)、早睡早起睡眠习惯、经常性失眠]的全基因组关联分析(GWAS)数据集与1个冠心病GWAS数据集。应用非独立工具变量异质性检测方法评估并排除无效工具变量,利用基于广义的汇总数据孟德尔随机化方法估计睡眠与冠心病之间的因果效应,应用Bonferroni法校正检验的显著性水平。**结果** 睡眠时间与冠心病具有显著的因果关联($OR=0.755$, $95\%CI:0.658\sim0.867$, $P=6.68E-05$)。短睡眠时间($OR=4.251$, $95\%CI:2.396\sim7.541$, $P=7.51E-07$)以及经常性失眠($OR=1.814$, $95\%CI:1.346\sim2.446$, $P=9.25E-05$)均会显著增加冠心病患病风险。而长睡眠时间以及早睡早起睡眠习惯与冠心病的因果关联不显著。**结论** 睡眠时间以及经常性失眠与冠心病存在因果关联。

【关键词】 冠心病;睡眠;孟德尔随机化

基金项目:国家自然科学基金(81872715);山西省研究生教育创新项目(2019SY268)

DOI:10.3760/cma.j.cn112338-20190624-00462

Causal relationship between sleep and coronary artery disease: a Mendelian randomization study

Gao Xue¹, Yang Xiaochen², Meng Lingxian¹, Sun Heli¹, Wang Tong¹

¹Department of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan 030001, China; ²Department of Medical and Biological Statistics, School of Statistics, Renmin University of China, Beijing 100872, China

Correspondence author: Wang Tong, Email: tongwang@sxmu.edu.cn

【Abstract】 Objective To understand the causal relationship between sleep and coronary artery disease (CAD). **Methods** This study included six genome-wide association parts; five for sleep related traits [sleep duration (continuous variable), long sleep duration (binary variable), short sleep duration (binary variable), early-to-bed/up habit (known as 'morningness') and frequently insomnia] and one for CAD. Heterogeneity in dependent instrument approach was used to assess and to remove the pleiotropic instruments. Generalized summary data-based Mendelian randomization was performed to estimate the causal relationships between sleep related traits and CAD. And a conservative Bonferroni was used for statistical tests. **Results** Sleep duration was putatively causal for CAD ($OR=0.755$, $95\%CI:0.658-0.867$, $P=6.68E-05$). Our results also indicated significant causal effects between both short sleep duration and frequently insomnia on CAD, with ORs as 4.251 ($95\%CI:2.396-7.541$, $P=7.51E-07$) and 1.814 ($95\%CI:1.346-2.446$, $P=9.25E-05$), respectively. There was no convincing evidence of causality between long sleep duration or morningness with CAD. **Conclusions** Our findings suggested that both sleep duration and frequently insomnia played causal roles on CAD, indicating that disease models should include sleep duration and insomnia as potential factors for CAD to develop effective interventions.

【Key words】 Coronary artery disease; Sleep; Mendelian randomization

Fund programs: National Natural Science Foundation of China (81872715); The Graduate Education Innovation Project of Shanxi Province (2019SY268)

DOI:10.3760/cma.j.cn112338-20190624-00462

冠心病是威胁人类健康的主要疾病之一,如何降低冠心病对社会、家庭和个人造成的巨大负担是

一项极大的挑战^[1]。有研究表明睡眠状况可能与冠心病的发生发展有关。现有流行病学研究对睡眠与

冠心病之间的关联进行了探索,然而得到的结论却不尽一致^[2-5]。近年,孟德尔随机化(Mendelian randomization, MR)在医学领域受到广泛应用。MR 是以遗传变异通常是单核苷酸多态性(SNP)作为工具变量估计暴露与疾病之间效应关联的模型^[6-7]。由于等位基因在配子形成过程中随机分离与组合,遗传变异实现了对于人群的随机分配,这就类比于传统的随机对照试验^[8]。因此,本研究应用基于汇总数据的两样本 MR 方法探讨睡眠时间、早睡早起睡眠习惯、失眠等睡眠性状与冠心病之间的因果关联。

资料与方法

1. 资料来源:睡眠数据自 2019 年最新发布的全基因组关联研究(Genome-Wide Association Study, GWAS),样本均自英国生物样本库(UK Biobank, UKB),汇总统计量源于睡眠疾病知识平台(Sleep Disorders Knowledge Portal, <http://sleepdisordergenetics.org/informational/data/>)。睡眠时间的 GWAS 包括 446 118 个样本,以研究者自我报告的每天习惯性睡眠时间定义睡眠时间。另外,在该项研究中还将睡眠时间定义为二分类变量,从而分别进行短睡眠时间(<7 h)以及长睡眠时间(≥ 9 h)相对于正常睡眠时间(7~8 h)的 GWAS 研究。分析中校正了年龄、性别、遗传主成分以及遗传相关性等因素^[9]。

早睡早起睡眠习惯的 GWAS 共纳入 449 734 个样本,参与者被要求回答“你认为你是哪种睡眠类型”这一问题,回答“早睡早起型”和“早睡早起多于晚睡晚起”者被定义为病例组,回答“晚睡晚起型”和“晚睡晚起多于早睡早起”者被定义为对照组,回答其余选项者列为缺失处理。分析中校正了年龄、性别、研究中心等因素^[10]。

经常性失眠的 GWAS 包含了 237 627 个样本,根据如下问题进行诊断:“您是否在夜里难以入睡,或者是否在半夜会醒”。回答经常出现的参与者被定义为病例组,而回答很少或从不出现的参与者被定义为对照组^[11]。

基于两样本的汇总数据 MR 模型需满足暴露与结局属于独立的样本,因此对冠心病数据集的选取,应排除与上述睡眠性状的队列相重叠的 GWAS(即排除应用 UKB 作为主要队列的分析)。基于此标准,我们选择了样本量较大($n=184\ 305$)的 1 项 GWAS-Meta 分析,数据源于冠状动脉疾病全基因组复制与 Meta 分析及冠状动脉疾病联合工作组(CARDIoGRAMplusC4D Consortium, [\[cardiogramplusc4d.org\]\(http://www.cardiogramplusc4d.org\)\)。病例按照冠心病\(如心肌梗死、急性冠状动脉综合征、慢性稳定型心绞痛、冠脉狭窄>50%等\)诊断标准判定^{\[12\]}。](http://www.</p>
</div>
<div data-bbox=)

上述各 GWAS 简要信息参见表 1。伦理审查以及知情同意已在各原始研究中获得。

表 1 睡眠与冠心病相关的全基因组关联研究

暴露	病例组	对照组	PubMed ID
睡眠时间 ^[9]	446 118	-	30846698
短睡眠时间 ^[9]	106 192	305 742	30846698
长睡眠时间 ^[9]	34 184	305742	30846698
早睡早起 ^[10]	252 287	150 908	30696823
经常性失眠 ^[11]	129 270	108 357	30804566
冠心病 ^[12]	60 801	123 504	26343387

2. 统计学分析:应用基于广义的汇总数据 MR (Generalized Summary data-based MR, GSMR) 方法进行睡眠与冠心病的因果关联分析^[13]。GSMR 方法首先计算每个工具变量的比例估计值,再应用广义最小二乘方法将各个比例估计值整合。GSMR 相比于传统 MR 模型具有更高的效能,因为它同时考虑了工具变量与结局及与暴露的效应估计值的误差,而其他 MR 模型只考虑工具变量与结局的效应估计值的误差,而假定工具变量与暴露的效应估计无误差^[14];另外,该方法校正了 SNP 之间的连锁不平衡(Linkage disequilibrium, LD),而其他方法则通常假定工具变量之间是独立的^[15-16]。

MR 的第一个假设是工具变量与暴露具有稳健的关联关系^[17]。因此,选择与 5 个睡眠变量具有全基因组统计显著性水平($P<5\times 10^{-8}$)的 SNP 作为初步的工具变量候选集。接下应用聚合算法筛选近似独立的 SNP($LD\ r^2<0.05$)以排除工具变量之间的强共线性对模型的影响。工具变量之间仍存在的弱相关性可以被 GSMR 方法所校正^[13]。现有的汇总水平的数据无法得到 SNP 之间的 LD 信息,因此应用 1 000 基因组工程第三阶段发布的个体水平的基因型数据作为参考数据集,估算 SNP 之间的相关性^[18]。

MR 的第二个假设是工具变量与影响“暴露—结局”关联的混杂因素独立,第三个假设是工具变量只能通过暴露对结局产生影响,而不能直接作用于结局。这两个假设的违背在本质上都是建立了除“工具变量—暴露—结局”之外的其他效应通路,通常被合并称为工具变量无多效性假设^[7]。在遗传学中,多效性是指遗传变异通过除感兴趣的因素之外的其他通路影响结局^[19]。多效性遗传变异常会导致 MR 模型中的检验统计量膨胀,从而使得估计结果

产生偏差。基于此,应用非独立工具变量异质性检测方法(Heterogeneity in Dependent Instrument-outlier test, HEIDI-outlier test)排除多效性工具变量,并利用其全局检验(HEIDI-global test)评估剩余工具变量的有效性^[20],以此规避由于工具变量与混杂因素相关或者工具变量直接作用于结局而导致因果效应估计出现偏倚。

为了得到更为稳健的估计结果,应用Bonferroni法校正检验的显著性水平(0.050/5,校正5种睡眠性状与冠心病的因果关联)。GCTA1.91.7软件(<http://cnsngomics.com/software/gcta>)用于读取参考数据集并估算SNP之间的相关性、进行聚合算法筛选近似独立的工具变量、剔除无效工具变量以及进行因果关联估计^[21]。

结 果

在睡眠时间与冠心病的因果关联分析中,聚合算法确定86个SNP作为工具变量候选集。HEIDI-outlier test识别rs13088093具有多效性予以排除,对剩余85个SNP进行HEIDI-global test,未发现证据表明剩余工具变量存在多效性($P=0.082$)。应用该85个SNP作为最终的工具变量进行GSMR分析,并得到显著的效应估计($OR=0.755$, $95\%CI:0.658 \sim 0.867$, $P=6.68E-05$) (表2)。

表2 睡眠相关性状与冠心病的因果关联分析

暴露	应用于MR模型的 工具变量个数	HEIDI全局 检验的 显著性	OR值 (95%CI)	P值
睡眠时间	85	0.082	0.755(0.658 ~ 0.867)	6.68E-05
短睡眠时间	29	0.370	4.251(2.396 ~ 7.541)	7.51E-07
长睡眠时间	10	0.091	0.208(0.048 ~ 0.897)	0.035
早睡早起	160	0.064	1.008(0.961 ~ 1.057)	0.743
经常性失眠	51	0.056	1.814(1.346 ~ 2.446)	9.25E-05

在短睡眠时间(<7 h)与冠心病的因果关联分析中,聚合算法确定29个与暴露关联显著的SNP。HEIDI-outlier test未检测出存在多效性的SNP, HEIDI-global test也未提示多效性($P=0.370$)。利用该29个SNP作为短睡眠时间的工具变量分析其对冠心病的因果效应,结果显示二者存在显著的因果关联($OR=4.251$, $95\%CI:2.396 \sim 7.541$, $P=7.51E-07$) (表2)。

在长睡眠时间(≥ 9 h)与冠心病的因果关联分析中,聚合算法确定10个与暴露关联显著的SNP。HEIDI-outlier test未检测出存在多效性的SNP, HEIDI-global test也没有证据表明这些SNP中存在

多效性($P=0.091$)。利用这10个SNP作为工具变量进行分析,得到显著性 $P=0.035$,该 P 值小于名义上的显著性水平($P=0.050$)但大于校正后的显著性水平($P=0.010$),在此情况下长睡眠时间与冠心病只存在名义上显著的因果关联,然而经Bonferroni法校正后关联没有统计学意义($OR=0.208$, $95\%CI:0.048 \sim 0.897$, $P=0.035$) (表2)。在早睡早起睡眠习惯与冠心病的因果关联分析中,聚合算法确定163个与暴露具有强关联的SNP。其中,HEIDI-outlier test识别并剔除3个存在多效性的SNP:rs1421085、rs2398144、rs67425923,保留下的SNP经过HEIDI-global test测定无多效性($P=0.064$)。GSMR模型得到二者的效应参数估计并不显著($OR=1.008$, $95\%CI:0.961 \sim 1.057$, $P=0.743$) (表2)。在经常性失眠与冠心病的因果关联分析中,聚合算法确定52个与失眠具有显著关联的SNP, HEIDI-outlier test移除多效性工具变量rs2297787,剩余51个SNP经HEIDI-global test证明不存在多效性($P=0.056$)。GSMR得到的效应估计具有统计学意义($OR=1.814$, $95\%CI:1.346 \sim 2.446$, $P=9.25E-05$) (表2)。

讨 论

本研究中采用MR模型探讨了睡眠与冠心病之间的因果关联。结果显示,睡眠时间与冠心病具有显著的因果关联,过短的睡眠时间可能会增加冠心病的发病风险,而长睡眠时间与冠心病并没有显著的因果关联。另外,研究结果显示经常性失眠会增加冠心病的患病风险,而早睡早起习惯则对冠心病没有显著影响。这些结论为睡眠与冠心病的因果关联研究提供了统计学上的支持。

一些研究论证了睡眠与冠心病之间的作用机制及原理,为本研究提供了理论上的支持。例如:有研究证明睡眠问题可引起交感神经激活,而交感神经的激活可加重动脉粥样硬化,引起斑块破裂和血栓形成,影响冠心病的发展和预后^[22];睡眠问题会影响生物节律,引起过度应激状态,导致心律失常,从而增加心血管疾病的发生率^[23];睡眠问题可引起下丘脑-垂体-肾上腺轴和肾素血管紧张素系统的激活,导致相应激素水平升高,而交感神经的激活与这些激素的协同作用进一步增加了心血管事件的风险^[24];睡眠问题可影响褪黑素的分泌,而褪黑素分泌的减少与冠心病的发生发展有关^[25]。

本研究是采用MR模型探讨睡眠与冠心病之间因果关联的研究。该研究具有以下优点:第一,MR

模型的应用控制了混杂因素及逆向因果对于估计的影响,从而基于观察性研究得到了可靠的因果效应估计^[26]。另外,汇总数据 MR 模型以基于大样本的 GWAS 作为数据集,相比于基于个体数据的小样本模型,大样本的纳入很大程度上提高了检验效能^[27];第二, GSMR 方法相比于其他 MR 模型具有更高的检验效能,因为它同时控制了工具变量与暴露及结局的效应误差,并且校正了工具变量之间的 LD 对于结果的偏倚影响^[13];第三,尽管在此研究中应用的冠心病 GWAS 数据集源于病例对照研究,应用 GSMR 模型获得的关于睡眠对冠心病的因果关联效应仍然可以在一般人群中进行解释,而不止在患病群体中解释^[13]。

本研究存在局限性。首先,基于汇总数据的 MR 模型假设暴露与结局之间呈线性关联,但实际上二者之间的关联可能是非线性的。基于此,本研究将睡眠时间同时作为连续变量与分类变量进行量化,并分别研究了短睡眠时间与长睡眠时间对冠心病的影响。在实际研究中,应在得到的效应估计的基础上,结合以往研究的结论给出合理的解释;另外,由于无法从现有的 GWAS 数据集中获得按性别或年龄分层的数据,因此该研究无法验证不同性别或年龄层中睡眠与冠心病的关联是否是一致的。基于此,当分层 GWAS 汇总数据可得时,可针对不同性别或年龄层群体的睡眠状况与冠心病的因果关联进行进一步研究。

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

- [1] Dai XM, Wiernek S, Evans JP, et al. Genetics of coronary artery disease and myocardial infarction [J]. *World J Cardiol*, 2016, 8 (1): 1-23. DOI: 10.4330/wjcv.v8.i1.1.
- [2] Sofi F, Cesari F, Casini A, et al. Insomnia and risk of cardiovascular disease: a Meta-analysis [J]. *Eur J Prev Cardiol*, 2014, 21(1): 57-64. DOI: 10.1177/2047487312460020.
- [3] Cappuccio FP, Cooper D, D'Elia L, et al. Sleep duration predicts cardiovascular outcomes; a systematic review and Meta-analysis of prospective studies [J]. *Eur Heart J*, 2011, 32 (12): 1484-1492. DOI: 10.1093/eurheartj/ehr007.
- [4] Laugsand LE, Vatten LJ, Platou C, et al. Insomnia and the risk of acute myocardial infarction: a population study [J]. *Circulation*, 2011, 124 (19): 2073-2081. DOI: 10.1161/circulationaha.111.025858.
- [5] Yang XR, Chen H, Li SY, et al. Association of sleep duration with the morbidity and mortality of coronary artery disease: a Meta-analysis of prospective studies [J]. *Heart Lung Circ*, 2015, 24(12): 1180-1190. DOI: 10.1016/j.hlc.2015.08.005.
- [6] Ioannidis JP, Haidich AB, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies [J]. *JAMA*, 2001, 286(7): 821-830. DOI: 10.1001/jama.286.7.821.
- [7] 高雪,王慧,王彤.孟德尔随机化中多效性偏倚校正方法简介 [J]. *中华流行病学杂志*, 2019, 40(3): 360-365. DOI: 10.3760/cma.j.issn.0254-6450.2019.03.020.
- [8] Gao X, Wang H, Wang T. Review on correction methods related to the pleiotropic effect in Mendelian randomization [J]. *Chin J Epidemiol*, 2019, 40 (3): 360-365. DOI: 10.3760/cma.j.issn.0254-6450.2019.03.020.
- [8] Katan MB. Commentary: Mendelian randomization, 18 years on [J]. *Int J Epidemiol*, 2004, 33 (1): 10-11. DOI: 10.1093/ije/dyh023.
- [9] Dashti HS, Jones SE, Wood AR, et al. Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates [J]. *Nat Commun*, 2019, 10(1): 1100. DOI: 10.1038/s41467-019-08917-4.
- [10] Jones SE, Lane JM, Wood AR, et al. Genome-wide association analyses of chronotype in 697 828 individuals provides insights into circadian rhythms [J]. *Nat Commun*, 2019, 10 (1): 343. DOI: 10.1038/s41467-018-08259-7.
- [11] Lane JM, Jones SE, Dashti HS, et al. Biological and clinical insights from genetics of insomnia symptoms [J]. *Nat Genet*, 2019, 51(3): 387-393. DOI: 10.1038/s41588-019-0361-7.
- [12] Nikpay M, Goel A, Won HH, et al. A comprehensive 1 000 Genomes-based genome-wide association Meta-analysis of coronary artery disease [J]. *Nat Genet*, 2015, 47 (10): 1121-1130. DOI: 10.1038/ng.3396.
- [13] Zhu ZH, Zheng ZL, Zhang FT, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data [J]. *Nat Commun*, 2018, 9(1): 224. DOI: 10.1038/s41467-017-02317-2.
- [14] Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods [J]. *Stat Med*, 2016, 35(11): 1880-1906. DOI: 10.1002/sim.6835.
- [15] Aissani B. Confounding by linkage disequilibrium [J]. *J Hum Genet*, 2014, 59(2): 110-115. DOI: 10.1038/jhg.2013.130.
- [16] Gao X, Meng LX, Ma KL, et al. The bidirectional causal relationships of insomnia with five major psychiatric disorders: A Mendelian randomization study [J]. *Eur Psychiatry*, 2019, 60: 79-85. DOI: 10.1016/j.eurpsy.2019.05.004.
- [17] Pierce BL, Ahsan H, van der Weele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants [J]. *Int J Epidemiol*, 2011, 40 (3): 740-752. DOI: 10.1093/ije/dyq151.
- [18] The 1 000 Genomes Project Consortium. A global reference for human genetic variation [J]. *Nature*, 2015, 526 (7571): 68-74. DOI: 10.1038/nature15393.
- [19] Hackinger S, Zeggini E. Statistical methods to detect pleiotropy in human complex traits [J]. *Open Biol*, 2017, 7 (11): 170125. DOI: 10.1098/rsob.170125.
- [20] Zhu ZH, Zhang FT, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets [J]. *Nat Genet*, 2016, 48(5): 481-487. DOI: 10.1038/ng.3538.
- [21] Yang J, Lee SH, Goddard ME, et al. GCTA: a tool for genome-wide complex trait analysis [J]. *Am J Hum Genet*, 2011, 88 (1): 76-82. DOI: 10.1016/j.ajhg.2010.11.011.
- [22] Mullington JM, Haack M, Toth M, et al. Cardiovascular, inflammatory, and metabolic consequences of sleep deprivation [J]. *Prog Cardiovasc Dis*, 2009, 51 (4): 294-302. DOI: 10.1016/j.pcad.2008.10.003.
- [23] Stamler JS, Goldman ME, Gomes J, et al. The effect of stress and fatigue on cardiac rhythm in medical interns [J]. *J Electrocardiol*, 1992, 25(4): 333-338. DOI: 10.1016/0022-0736(92)90039-3.
- [24] Sun XY, Dai XY, Yang TS, et al. Effects of mental resilience on neuroendocrine hormones level changes induced by sleep deprivation in servicemen [J]. *Endocrine*, 2014, 47(3): 884-888. DOI: 10.1007/s12020-014-0228-8.
- [25] Sahn E, Olmez E, Acet A. Effects of physiological and pharmacological concentrations of melatonin on ischemia-reperfusion arrhythmias in rats: can the incidence of sudden cardiac death be reduced? [J]. *J Pineal Res*, 2002, 32 (3): 194-198. DOI: 10.1034/j.1600-079x.2002.10853.x.
- [26] Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization [J]. *JAMA*, 2017, 318(19): 1925-1926. DOI: 10.1001/jama.2017.17219.
- [27] Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data [J]. *Genet Epidemiol*, 2013, 37 (7): 658-665. DOI: 10.1002/gepi.21758.

(收稿日期: 2019-06-24)

(本文编辑: 万玉立)