

网络出版时间: 2024-07-03 09:58:42 网络出版地址: <https://link.cnki.net/urlid/34.1065.R.20240702.1450.022>

# 两样本孟德尔随机化探讨欧洲 类风湿关节炎人群与缺铁性贫血的因果关系

王勇, 姚晓玲, 杨豫正, 凌益, 姚血明, 马武开

**摘要** 目的 通过两样本孟德尔随机化法探讨欧洲类风湿关节炎(RA)人群与缺铁性贫血(IDA)的因果关系。方法 RA与IDA的单核苷酸多态性(SNPs)均使用公开的全基因组关联研究(GWAS)数据进行汇总分析。使用逆方差加权法(IVW)作为主要分析方法评估RA对IDA的因果效应。以MR-Egger回归法、加权中位数法(WM)、加权模式和简单模式作为回归补充说明,进而评估敏感性分析结果的稳健性。用Heterogeneity函数计算P值检验异质性、用截距项Intercept检验水平多效性。结果 在全基因组水平FINNGEN数据库中选择出了经孟德尔随机分析去除连锁不平衡且满足 $P < 5.0 \times 10^{-8}$ 的强相关SNPs,整合暴露与结局数据后获得31个SNPs作为最终的有效工具变量。IVW法显示RA是IDA的风险因素(在RA的患病人群中,发生IDA的风险是无RA人群的1.064倍,OR=1.064,95%CI:1.028~1.103);WM和MR-Egger回归结果支持RA对IDA存在正向关联;截距项数值趋近于0,表明暴露与结局不存在水平多效性;Heterogeneity函数 $P < 0.05$ ,表明暴露与结局存在异质性,但经随机效应模型检验 $P < 0.05$ ,表明因果关系即使存在异质性,但总体趋于稳健。结论 RA是IDA的危险因素,RA与IDA存在正向关联。

**关键词** 类风湿关节炎; 缺铁性贫血; 两样本孟德尔随机化; 因果关联

中图分类号 R 593.22

文献标志码 A 文章编号 1000-1492(2024)07-1251-06  
doi: 10.19405/j.cnki.issn1000-1492.2024.07.022

类风湿关节炎(rheumatoid arthritis, RA)是一种病因不明的多因素自身免疫性疾病,主要累及关节及骨和软骨,疾病后期常出现肺纤维化、干燥综合

征、慢性贫血等并发症<sup>[1]</sup>。中国大陆RA发病率为0.42%,患病人群约500万,男女患病比约为1:4<sup>[2]</sup>。RA相关的贫血类型包括缺铁性贫血(iron deficiency anemia, IDA)、再生障碍性贫血、溶血性贫血、慢性病贫血等,是影响RA患者健康的常见特征性疾病<sup>[3-4]</sup>。RA与IDA关系的报道主要以横断面研究为主,较少分析其因果关系,且RA与IDA常以合并症的形式进行描述性分析,并未对这些关联进行因果关系的统计分析。

孟德尔随机化(mendelian randomization, MR)通过基因组水平上由单个核苷酸的变异所引起的DNA序列多态性(即单核苷酸多态性)作为工具变量(instrumental variables, IVs)来估计暴露与结局之间的因果效应,且SNP的等位基因遵循孟德尔遗传学定律,在出生时随机分配的SNP遗传变异受到潜在混杂因素干扰的概率接近于零,并且遗传变异的发生在疾病发生之前业已建立,为此则能够减少反向因果关联的可能性<sup>[5-6]</sup>。MR规避了随机对照试验伦理限制、费用昂贵、持续时间长等受限因素,逐渐成为业内热点分析方法。该研究运用两样本孟德尔随机化(two sample mendelian randomization, TSMR)研究RA与IDA之间的因果关联,为临床治疗RA提供新的参考。

## 1 材料与方法

**1.1 研究设计** 本研究以RA作为暴露因素,以IDA作为结局变量,使用TSMR进行因果关联分析,以探究RA和IDA的因果效应。MR分析满足三大假设:①关联性假设:IVs与暴露因素强相关;②独立性假设:IVs独立于观察到或未观察到的混杂因素;③排他性假设:IVs仅通过暴露因素影响结局变量。具体流程见图1。

**1.2 数据来源** 通过FINNGEN数据库获得RA和IDA的全基因组关联研究(genome-wide association study, GWAS)数据。其数据细节见表1。

**1.3 工具变量的选择** 筛选强相关P值并且规避连锁不平衡(linkage disequilibrium, LD)偏倚<sup>[7]</sup>,参

2023-01-29 接收

基金项目:国家自然科学基金(编号:82274678);贵州省中医风湿免疫病临床研究中心(编号:黔科合平台人才[2020]2202号);中医药、民族医药科学技术研究专项课题项目(编号:QZYY-2023-020);贵州中医药大学国家与省级科技创新人才团队培育项目(编号:贵中医TD合字[2022]004号)

作者单位:贵州中医药大学第二附属医院风湿免疫科,贵阳 550001

作者简介:王勇,男,硕士研究生;

马武开,男,主任医师,博士生导师,责任作者,E-mail:walker55@163.com

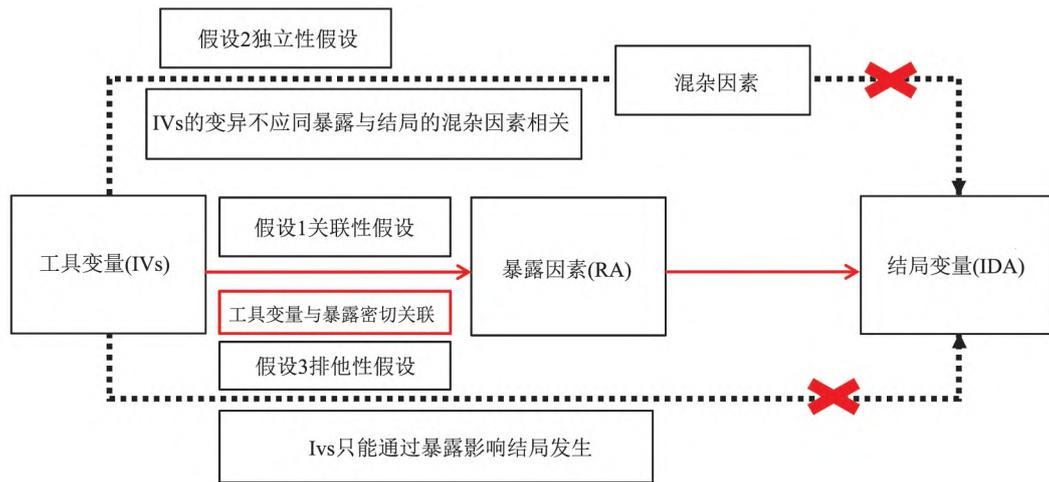


图1 TSMR 设计流程图

表1 RA 与 IDA 数据资料

表型	人群	病例组( n )	对照组( n )	单核苷酸多态性( 个 )	数据网址
RA	欧洲	9 139	368 029	20 170 234	<a href="https://storage.googleapis.com/finngen-public-data+9/summary_stats/finngen_R9_RHEUMA_SEROPOS_OTH.gz">https://storage.googleapis.com/finngen-public-data+9/summary_stats/finngen_R9_RHEUMA_SEROPOS_OTH.gz</a>
IDA	欧洲	10 208	360 528	20 170 106	<a href="https://storage.googleapis.com/finngen-public-data+9/summary_stats/finngen_R9_D3_ANAEMIA_IRONDEF_NAS.gz">https://storage.googleapis.com/finngen-public-data+9/summary_stats/finngen_R9_D3_ANAEMIA_IRONDEF_NAS.gz</a>

数设置为  $r^2 < 0.001$ 、遗传距离为 1 Mb，从中选择了 31 个 SNPs 作为 IVs 并记录相应的效应等位基因 (effect allele, EA) 和非效应等位基因 (other allele, OA)、效应值包括: P 值、*eaf* 值、*beta* 值、*se* 值。以逆方差加权法 (inverse variance weighted, IVW) 作为主要分析方法评估 RA 与 IDA 的因果关系; 以 MR-Egger 回归法<sup>[8]</sup>、加权中位数法 (weighted median, WM)、加权模式和简单模式作为回归补充说明; 敏感性分析运用 Heterogeneity 函数检验异质性, MR-Egger-intercept 检验是否存在水平多效性来验证结果的可靠性。选取  $F > 10$  的 SNPs 以避免弱工具变量偏倚以确保 IVs 与暴露之间的强相关性。

**1.4 统计学处理** 所有数据分析用 R 语言 (4.3.1 版) 软件 R 包 “TSMR”; 通过比值比 (odds ratio, OR)、95% 置信区间 (95% confidence interval, 95% CI) 等结果表示  $P < 0.05$  为差异有统计学意义。

## 2 结果

### 2.1 敏感性分析的结果

**2.1.1 水平多效性检验** 经过 “R 包: TSMR” 筛选确定出 31 个 SNPs 作为 IVs, 且单一的 SNP 对应的  $F$  值统计量分布范围为 30 ~ 1 461, 说明因果关联受到弱工具变量偏倚影响的可能性较小 (单一 SNP 的  $F$  值均  $> 10$ )。经 MR 筛选后的暴露与结局基本信

息见表 2。MR-Egger-intercept = 0.003 接近 0,  $P = 0.617$ , 表明 IVs 不存在水平多效性, 减少了影响 MR 分析结果的可能性。

**2.1.2 异质性检验** Heterogeneity 函数  $P = 0.045$  表示存在异质性, 但随机效应模型估计 MR 效应量显示  $P < 0.05$ , 间接表明暴露与结局即使存在异质性, 但总体结果趋于稳健, 见表 3。漏斗图因果效应分布基本呈轴对称, 发生偏倚的可能性小, 说明 IVs 之间差异小 (图 2A)。使用留一法揭示敏感性分析结果, 逐一剔除每个 SNP, 然后计算剩余 SNP 的效应值, 观察剔除 SNP 后的结果, 用来排查因果关系是否受单个工具变量影响, 最后与 MR 分析比较, 图示均在 “0” 的右侧呈正向分布, 说明 MR 分析所得结果稳健, 即便剔除任意一个 SNP 都难以对因果相关的推算产生较大影响 (图 2B)。森林图所有 IVW 估值均在 0 右侧, 表明结果稳健可靠 (图 2C)。散点图斜率大于 0, 表明 RA 是 IDA 的危险因素 (图 2D)。

**2.2 TSMR 分析结果** IVW 法分析结果显示: 遗传预测的 RA 与 IDA:  $OR = 1.064$ 、95%  $CI(1.028 \sim 1.103)$ 、 $P < 0.05$ , 表明因果效应显著, 二者具有因果关系。MR-Egger 回归法结果显示:  $OR = 1.054$ 、95%  $CI(0.999 \sim 1.111)$ 、 $P = 0.064$ , 表明因果效应不显著, 但可对敏感性分析起补充说明作用。WM、

简单模式、加权模式统计值分别为:  $OR = 1.059$ 、 $95\% CI(1.012 \sim 1.108)$ 、 $P = 0.014$ ;  $OR = 1.042$ 、 $95\% CI(0.935 \sim 1.162)$ 、 $P = 0.462$ ;  $OR = 1.049$ 、 $95\% CI(1.006 \sim 1.093)$ 、 $P = 0.033$ ; 三者  $OR$  值方向相同,推测 RA 与 IDA 之间具有正向联系,进一步表明 RA 与 IDA 具有因果关系,见表 4 及图 3。

表 2 RA 与 IDA 工具变量的基本信息

SNP	EA	OA	RA				IDA				F 值
			beta 值	se 值	P 值	eaf 值	beta 值	se 值	P 值	eaf 值	
rs10174238	A	G	-0.116	0.017	<0.001	0.764	-0.016	0.017	0.331	0.764	47
rs1024162	T	A	-0.134	0.016	<0.001	0.325	-0.007	0.015	0.658	0.325	70
rs10517086	A	G	0.110	0.016	<0.001	0.282	-0.028	0.016	0.078	0.282	47
rs10821948	A	C	0.094	0.015	<0.001	0.425	0.026	0.015	0.071	0.425	39
rs117292830	A	G	0.655	0.032	<0.001	0.038	0.058	0.037	0.121	0.038	419
rs117753409	A	C	0.148	0.026	<0.001	0.085	0.051	0.026	0.049	0.085	32
rs13180950	C	T	0.115	0.020	<0.001	0.155	0.000	0.020	0.981	0.155	33
rs1451723	T	C	0.084	0.015	<0.001	0.536	0.030	0.014	0.036	0.536	31
rs16903065	A	C	-0.143	0.024	<0.001	0.115	-0.005	0.022	0.819	0.116	36
rs2013002	C	T	-0.108	0.015	<0.001	0.618	0.010	0.015	0.505	0.619	52
rs2249354	A	G	0.376	0.068	<0.001	0.983	0.034	0.056	0.542	0.983	31
rs2476601	G	A	-0.414	0.019	<0.001	0.852	-0.070	0.020	<0.001	0.852	475
rs28579922	A	G	-0.685	0.026	<0.001	0.125	-0.009	0.022	0.665	0.125	694
rs3114891	G	A	0.087	0.015	<0.001	0.428	0.009	0.014	0.517	0.428	34
rs3118470	C	T	0.090	0.015	<0.001	0.441	-0.006	0.014	0.656	0.441	36
rs3128927	T	C	-0.324	0.019	<0.001	0.227	0.023	0.017	0.175	0.227	291
rs34536443	C	G	-0.404	0.050	<0.001	0.030	-0.019	0.042	0.655	0.030	65
rs3757387	C	T	0.105	0.015	<0.001	0.425	0.018	0.014	0.216	0.425	49
rs4129745	G	A	-0.100	0.017	<0.001	0.286	-0.039	0.016	0.015	0.286	35
rs6032664	T	A	0.096	0.017	<0.001	0.727	0.002	0.016	0.915	0.727	32
rs60600003	G	T	0.134	0.024	<0.001	0.102	0.032	0.024	0.172	0.102	31
rs6456160	C	T	-0.112	0.015	<0.001	0.558	0.006	0.014	0.696	0.558	56
rs66654254	A	G	-0.087	0.015	<0.001	0.435	-0.002	0.015	0.915	0.435	34
rs71565324	A	C	0.590	0.070	<0.001	0.008	0.134	0.078	0.087	0.009	71
rs72854535	C	T	0.219	0.021	<0.001	0.129	0.058	0.021	0.006	0.129	109
rs73510898	A	G	-0.156	0.028	<0.001	0.085	-0.002	0.026	0.936	0.085	31
rs74912105	G	A	0.363	0.056	<0.001	0.015	-0.013	0.059	0.826	0.015	42
rs7731626	A	G	-0.099	0.017	<0.001	0.279	0.013	0.016	0.424	0.279	34
rs78248443	T	C	-0.164	0.030	<0.001	0.075	-0.019	0.027	0.484	0.075	30
rs78782944	T	C	0.163	0.025	<0.001	0.095	0.030	0.025	0.221	0.095	43
rs9268543	T	A	0.688	0.018	<0.001	0.152	0.045	0.020	0.025	0.152	1461

表 3 随机效应模型及截距

暴露因素	结局变量	分析方法	SNPs	beta 值	se 值	P 值	截距项
RA	IDA	逆方差加权法	31	0.062	0.018	<0.001	
RA	IDA	MR-Egger	31		0.006	0.617	0.003
RA	IDA	异质性	31			0.045	

表 4 TSMR 分析结果

分析方法	SNPs	beta 值	se 值	P 值	OR 值	95% CI
MR-Egger	31	0.052	0.027	0.064	1.054	0.999 ~ 1.111
WM	31	0.057	0.023	0.014	1.059	1.012 ~ 1.108
IVW	31	0.062	0.018	<0.001	1.064	1.028 ~ 1.103
简单模式	31	0.041	0.056	0.462	1.042	0.935 ~ 1.162
加权模式	31	0.048	0.021	0.033	1.049	1.006 ~ 1.093

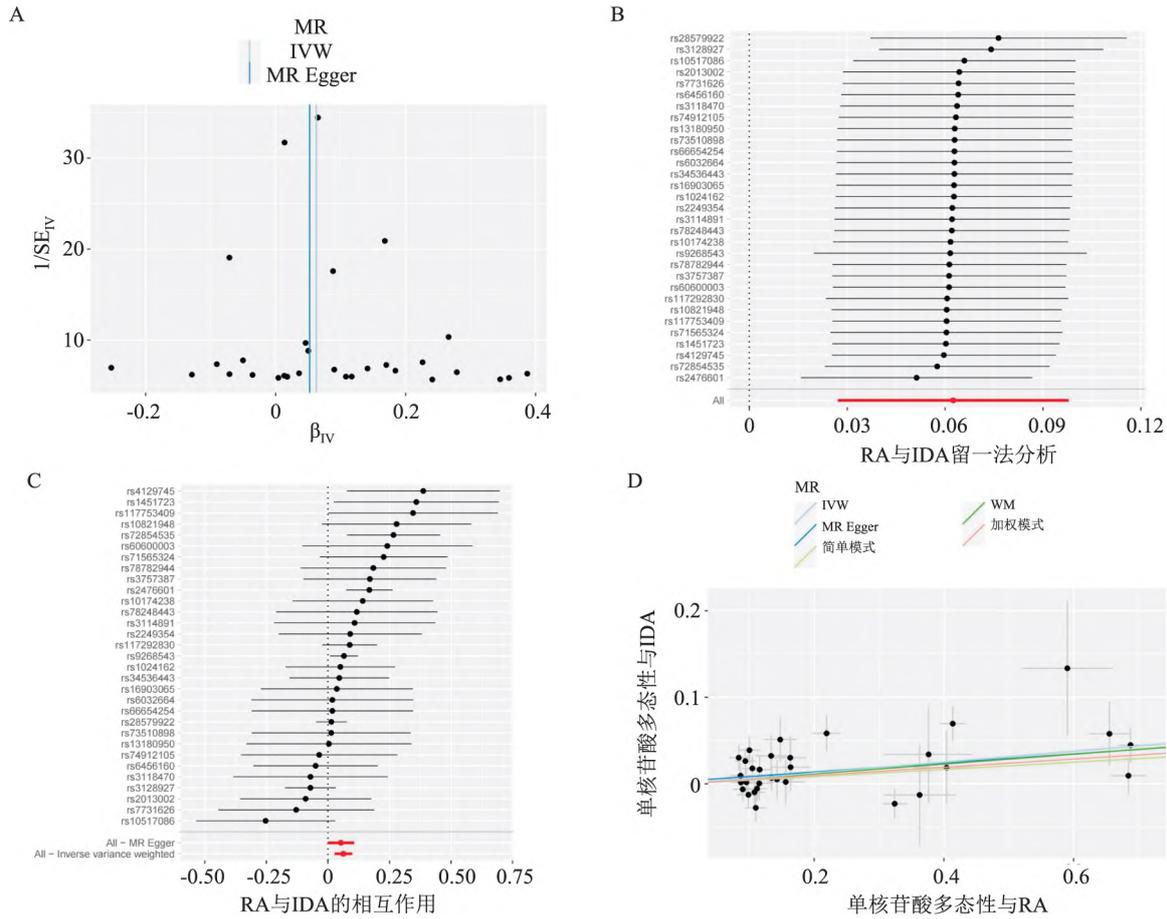


图2 RA 与 IDA 的 TSMR 分析结果示意图

A: 漏斗图; B: 留一法; C: 森林图; D: 散点图

分析方法	SNPs	OR(95%CI)	P值
MR-Egger	31	1.054(0.999 ~ 1.111)	0.064
WM	31	1.059(1.012 ~ 1.108)	0.014
IVW	31	1.064(1.028 ~ 1.103)	<0.001
简单模式	31	1.042(0.935 ~ 1.162)	0.462
加权模式	31	1.049(1.006 ~ 1.093)	0.333

图3 RA 与 IDA 风险评估图

### 3 讨论

MR 设计是一种基于孟德尔遗传的第一定律和第二定律且不受环境与行为因素干扰的自然随机分配方法,研究样本量大,统计效能高,利用在疾病发生发展前就建立好的遗传变异作为工具变量来评估因果效应,即以遗传变异代替临床实验的随机分配,在这个过程中,这种设计方法能够有效减少混杂因

素的影响<sup>[9]</sup>。

贫血是 RA 患者最常见的关节外表现之一,常伴随轻至中度贫血,而 RA 作为一种炎症性慢性疾病,对铁的合成不足、丢失过多或者破坏增加尚未有深入的认识。研究<sup>[10]</sup>表明可溶性转铁蛋白在 IDA 中发挥一定的作用,它是衡量红细胞生成活性的指标,在细胞铁含量降低时升高,不受炎症介质影响。IDA 以慢性失血造成的可能性大,较少由单纯铁摄

入不足所致,它既可以是一个独立性的疾病,也可以作为其他疾病的一个症状出现。RA 发展到颞颌关节、腕掌关节、掌指关节、肩关节等关节破坏阶段时,相应的会降低生活质量,小关节破坏导致的精细动作运行受阻所致拾筷困难、咀嚼功能的下降相继影响到饮食营养的摄入而成为 IDA 的安全隐患,尽管这些都是宏观外源性引起的 IDA,但在 RA 中识别其他或共存的常见贫血原因至关重要<sup>[11]</sup>。本研究表明 RA 与 IDA 存在因果关系,且经临床观察 RA 发生发展中贫血的发生也不在少数,亦以 IDA 居多,且贫血的程度影响着 RA 预后,应当引起重视。本研究发现 31 个与 RA 以及 IDA 具有强相关的 SNPs 为后续从微观内源性基因表达上探索 RA 与 IDA 提供方向,或许能够为药物靶点基础实验提供有力的理论依据也尚未可知;研究<sup>[12-13]</sup>还发现,不同年龄段的 RA 患者贫血程度也有所差异,所以依据年龄段的不同,采取不同的干预措施能否减少贫血的发病率甚至纠正贫血需要进一步的探究。另一方面,非甾体类抗炎药对治疗 RA 起到抗炎止痛效应的同时也相应的伴随着胃肠道不适等不良反应,在缓解 RA 疼痛的同时也伴随着因药物引起胃黏膜损伤的风险继而诱发慢性失血;炎性细胞因子不仅对治疗 RA 起反作用还会诱导炎性贫血,故而在治疗 RA 的过程中需兼顾抑制炎症因子的释放降低 IDA 的发生率,其中肿瘤坏死因子- $\alpha$  抑制剂作为治疗 RA 的有效药物之一,其使用对治疗 RA 以及延缓 IDA 的发生需进一步探索<sup>[14]</sup>。无论 RA 在缓解期或活动期,用药干预的同时仍需要预防 IDA 的发生,避免疾病的恶性发展。

本研究存在相应的局限性:① 本研究所使用的暴露和结局的人群样本仅为欧洲人群,全球 GWAS 数据仍在完善中,故本结论尚不能代入整体人群,其他人群尚待整合荟萃分析后进一步研究;② 本研究数据来源尚未进行年龄阶段划分、性别划分等一般资料的细则划分,尚需进一步拓宽信息数据以增强结果的适应性;③ 本研究仅通过 TSMR 分析 RA 和 IDA 的因果关系,但贫血分类型种类繁多,且缺乏 GWAS 数据来源,对其他的贫血种类有待进一步的研究。

综上所述,此次研究结果从基因层面揭示了 RA 与 IDA 之间存在正向因果关联,提示未来在临床治疗 RA 时应考虑 IDA 发生的可能性,关注血常

规及铁蛋白等相应检验指标。偏食摄入不足、青少年铁元素需求增加、胃肠道疾病吸收不佳、钩虫病丢失过多等原因也会引发 IDA,RA 与 IDA 的因果关系仅从理论分析数据获得,尚需实验验证,本次分析结果仅作参考。

## 参考文献

- [1] Radu A F, Bungau S G. Management of rheumatoid arthritis: An overview [J]. *Cells* 2021, 10(11): 2857.
- [2] 中华医学会风湿病学分会. 2018 中国类风湿关节炎诊疗指南 [J]. *中华内科杂志* 2018, 57(4): 242-51.
- [3] Ali E T, Jabbar A S, Mohammed A N. A comparative study of interleukin 6, inflammatory markers, ferritin, and hematological profile in rheumatoid arthritis patients with anemia of chronic disease and iron deficiency anemia [J]. *Anemia* 2019; 2019: 1-7.
- [4] Kumar A, Sharma E, Marley A, et al. Iron deficiency anaemia: Pathophysiology, assessment, practical management [J]. *BMJ Open Gastroenterol* 2022, 9(1): e759.
- [5] Allman P H, Aban I B, Tiwari H K, et al. An introduction to Mendelian randomization with applications in neurology [J]. *Mult Scler Relat Disord* 2018, 24: 72-8.
- [6] 王玉琢, 沈洪兵. 孟德尔随机化研究应用于因果推断的影响因素及其结果解读面临的挑战 [J]. *中华流行病学杂志* 2020, 41(8): 1231-6.
- [7] Slatkin M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future [J]. *Nature Nat Rev Genet*, 2008, 9(6): 477-85.
- [8] Burgess S, Thompson S G. Interpreting findings from mendelian randomization using the MR-Egger method [J]. *Eur J Epidemiol*, 2017, 32(5): 377-89.
- [9] Gala H, Tomlinson I. The use of mendelian randomisation to identify causal cancer risk factors: Promise and limitations [J]. *J Pathol*, 2020, 250(5): 541-54.
- [10] Günther F, Straub R H, Hartung W, et al. Usefulness of soluble transferrin receptor in the diagnosis of iron deficiency anemia in rheumatoid arthritis patients in clinical practice [J]. *Int J Rheumatol* 2022, 2022: 1-9.
- [11] Demir S, Kucuk A, Altas M, et al. Restless leg syndrome and sleep disorders in patients with rheumatoid arthritis and its relation with anemia parameters [J]. *Acta Medica (Hradec Kralove)* 2021, 64(3): 137-44.
- [12] Chai K L, Cole-Sinclair M. Review of available evidence supporting different transfusion thresholds in different patient groups with anemia [J]. *Ann N Y Acad Sci* 2019, 1450(1): 221-38.
- [13] Sharif N, Das B, Alam A. Prevalence of anemia among reproductive women in different social group in india: Cross-sectional study using nationally representative data [J]. *PLoS One* 2023, 18(2): e0281015.

[14] Corrado A ,Bello V D ,dOnofrio F ,et al. Anti-TNF- $\alpha$  effects on anemia in rheumatoid and psoriatic arthritis[J]. *Int J Immunopathol*

*Pharmacol* 2017 ,30(3) :302-7.

## To explore the causal relationship between rheumatoid arthritis and iron deficiency anemia in European population by two-sample Mendelian randomization

Wang Yong ,Yao Xiaoling ,Yang Yuzheng ,Ling Yi ,Yao Xueming ,Ma Wukai  
( Dept of Rheumatology and Immunology ,The Second Affiliated Hospital of Guizhou University of Traditional Chinese Medicine ,Guiyang 550001)

**Abstract** *Objective* To explore the causal relationship between rheumatoid arthritis ( RA ) and iron deficiency anemia ( IDA ) in European population by two-sample Mendelian randomization analysis. *Methods* The single nucleotide polymorphisms ( SNPs ) of RA and IDA were analyzed using public genome-wide association studies ( GWAS ) . The inverse variance weighting method ( IVW ) was used as the main analysis method to evaluate the causal effect of RA on IDA. MR-Egger method ,weighted median method ( WM ) ,weighted model method and simple model method were used as regression supplements to evaluate the robustness of sensitivity analysis results. The heterogeneity function was used to calculate the *P*-value to test the heterogeneity and the intercept term intercept was used to test the level pleiotropy. *Results* In the FINNGEN database at the genome-wide level ,strong-related SNPs that removed linkage disequilibrium and met the  $P < 5.0 \times 10^{-8}$  by Mendelian randomization analysis were selected. After integrating exposure and outcome data 31 SNPs were obtained as the final effective instrumental variables. IVW showed that RA was a risk factor for IDA ( the risk of IDA in RA patients was 1.064 times higher than that in non-RA patients ,*OR* = 1.064 ,95% *CI*: 1.028 - 1.103 ) . The weighted median method and MR-Egger method results supported the positive correlation between RA and IDA. The intercept value was close to 0 ,indicating that there was no horizontal pleiotropy between exposure and outcome. The heterogeneity function's  $P < 0.05$  indicated that there was heterogeneity between exposure and outcome ,but the random effect model test showed  $P < 0.05$  ,indicating that even if there was heterogeneity in causality ,the overall trend was stable. *Conclusion* RA is a risk factor for IDA ,and there is a positive correlation between RA and IDA.

**Key words** rheumatoid arthritis; iron deficiency anemia; two-sample mendelian randomization; causal association