

油性坚果摄入量与原发肺癌的两样本孟德尔随机化研究

郭学云¹⁾, 李娜²⁾, 陶国仲²⁾, 黄正²⁾

(1) 天长市人民医院肿瘤科; 2) 感染科, 安徽 天长 239321)

[摘要] **目的** 以两样本孟德尔随机化方法探讨油性坚果摄入量与原发肺癌的因果关系。**方法** 从全基因组关联研究 openGWAS 公开数据库中筛选出与油性坚果摄入量相关的($P < 5 \times 10^{-7}$)单核苷酸多态性位点作为工具变量, 以逆方差加权法为该研究评估因果效应的主要分析方法, 并采用 MR-Egger 回归和加权中位数法用于补充说明, 以 OR 值及 95%CI 评价油性坚果摄入量与原发肺癌之间是否存在因果关系。采用 mr_pleiotropy_test 法及 MR-PRESSO 法检验水平多效性, leave-one-out 留一法进行敏感性检验。**结果** 从 ukb-b-16576 数据集共提取 100 个 SNP 作为工具变量, MR-Egger 回归、加权中位数和逆方差加权法计算得 OR 值和 95% 可信区间分别 [0.63, 95%CI(0.11 ~ 3.65), $P = 0.611$], [0.34, 95%CI(0.19 ~ 0.61), $P = 0.001$], [0.32, 95%CI(0.21 ~ 0.48), $P = 0.001$], 所有 b 值结果方向均小于 0。mr_pleiotropy_test 检验显示 MR-Egger 截距为 $-7.66E-03$, 接近于 0, $P = 0.441$, 同样 MR-PRESSO 包 GlobalTest 检验 $P = 0.162$, 所选的工具变量不存在水平多效性, 经过 leave-one-out 敏感性分析, 未发现任何 SNP 位点对结果产生显著影响, 表明结果具有较高的稳定性和可靠性。**结论** 两样本孟德尔随机化分析表明油性坚果摄入量与原发肺癌的发病风险呈负向因果关系, 但需要更大样本量和不同种族背景的人的数据来进一步验证。

[关键词] 油性坚果; 原发性肺癌; 孟德尔随机化; 因果推断

[中图分类号] R151.3 **[文献标志码]** A **[文章编号]** 2095-610X(2024)06-0044-07

A Two-sample Mendelian Randomization Study of Oily Nut Intake and Primary Lung Cancer

GUO Xueyun¹⁾, LI Na²⁾, TAO Guozhong²⁾, HUANG Zheng²⁾

(1) Dept. of Oncology; 2) Dept. of Infectious Diseases, Tianchang People's Hospital, Tianchang Anhui 239321, China)

[Abstract] **Objective** To investigate the potential causal relationship between oily nut consumption and primary lung cancer by two-sample Mendelian randomization method. **Methods** Single nucleotide polymorphisms ($P < 5 \times 10^{-7}$) associated with oily nut intake were screened from the openGWAS public database of genome-wide association study as instrumental variables. The inverse variance weighting method was the main analysis method to evaluate the causal effect of this study, and MR-Egger regression and weighted median method were used to supplement the explanation. OR value and 95%CI were used to evaluate whether there was a causal relationship between oily nut intake and primary lung cancer. The mr_pleiotropy_test and MR-PRESSO were used to test the level pleiotropy, and the leave-one-out method was used to test the sensitivity. **Results** A total of 100 SNPs were extracted from the ukb-b-16576 dataset as instrumental variables. The MR-Egger regression, weighted median, and inverse variance weighting methods yielded OR values and 95% confidence intervals [0.63, 95%CI (0.11-3.65), $P = 0.611$], [0.34, 95%CI (0.19-0.61), $P = 0.001$], [0.32, 95%CI (0.21-0.48), $P = 0.001$], respectively. The mr_pleiotropy_test showed that the MR-Egger intercept was close to 0 ($-7.66E-03$) with a p-

[收稿日期] 2024-03-10

[基金项目] 中央级公益性科研院所基本科研业务费专项资金资助项目(CAFYBB20ZI001-04); 滁州市卫生健康科研课题基金资助项目(CZWJ2022C004)

[作者简介] 郭学云(1989~), 女, 安徽天长人, 医学硕士, 主治医师, 主要从事肿瘤内科研究工作。

[通信作者] 黄正, E-mail: 1217748705@qq.com

value of 0.441, indicating no significant horizontal pleiotropy. Similarly, the MR-PRESSO package GlobalTest yielded a *p*-value of 0.162, further confirming the absence of horizontal pleiotropy among the selected instrumental variables. After conducting the leave-one-out sensitivity analysis, no significant influence on the results was observed from any specific SNP locus, indicating that the results are highly stable and reliable. **Conclusions** The two-sample Mendelian randomization analysis revealed a negative causal relationship between oily nut intake and the risk of primary lung cancer. However, further verification is needed with larger sample sizes and data from individuals of different ethnic backgrounds to strengthen the findings.

[**Key words**] Oily nuts; Primary lung cancer; Mendel randomization; Causal inference

原发性肺癌(primary pulmonary carcinoma)是一种常见且具有极大危害性的恶性肿瘤。根据全球癌症流行病学统计数据显示, 2020年全球新发220.7万肺癌病例, 新增179.6万死亡肺癌病例, 分别占全部恶性肿瘤新发和死亡病例的11.4%和18.0%^[1], 而在中国2022年新发肺癌病例约87.1万, 新增肺癌死亡病例约76.7万, 分别占有恶性肿瘤发病和死亡病例的18.1%和23.9%, 肺癌新增和死亡病例均位居我国恶性肿瘤的首位^[2]。研究测算了预计在2020~2050年间, 癌症将使世界损失25.2万亿美元, 其中肺癌的经济负担最重, 将造成全球3.9万亿美元的损失, 通过预防肺癌发生对减轻社会经济负担意义重大^[3]。坚果, 即果实成熟后为干燥状态的果子, 例如板栗、榛子、腰果、瓜子、松仁、杏仁、白果、开心果、夏威夷果等, 其对增强体质、预防疾病有极好的功效^[4]。1项基于人群的肺癌病因学病例对照研究以及美国国立卫生研究院有关饮食与健康研究的前瞻性队列研究均显示, 油性坚果(核桃、榛子、杏仁和花生)的消费量与肺癌发病率呈负相关, 吸烟人群从坚果消费中获益更为明显^[5], 同样荷兰1项对120852名老年人进行20.3 a的随访研究显示增加坚果摄入量可能有助于预防男性小细胞肺癌^[6]。因上述研究属于观察性队列研究, 由于观察法更易受制于暴露与结局之外的混调变量干扰, 故该方法难以深刻揭示暴露与结局之间的因果效应, 无法明确摄入坚果能否降低原发性肺癌发病率。因此, 本研究采用两样本MR法计算油性坚果摄入量与原发性肺癌之间的因果效应值大小, 评价两者是否存在因果关系, 为肺癌防治提供理论参考依据。

1 资料与方法

1.1 数据来源

通过IEU openGWAS数据库(<http://www.mrbase.org>)^[7]获取近年来样本量较大原发性肺癌及坚果

摄入量的GWAS数据, 2组数据均为欧洲人群, 其中坚果摄入量数据集ukb-b-16576样本数为421764例, 共9851867个SNP位点。原发性肺癌数据ieu-a-966的总样本量为27209, 包含11348例患者及15861例对照人群, 894589个SNP位点。

1.2 工具变量筛选

通过R语言TwoSampleMR包^[8]extract_instruments函数获取工具变量, 设置p1参数为参数为 5×10^{-7} , clumpTRUE去掉连锁不平衡LDA的工具变量, 设置KB和r2参数分别为1000和0.001, 去掉在1Mb范围内与最显著SNP的r2大于0.001的SNP, 然后与结局变量原发性肺癌的SNP汇总取交集, 构建SNPs-暴露因素-结局因素关系, 删除经phenoscanner函数检索出的与支气管哮喘^[9]、慢性阻塞性肺疾病^[10]相关位点以及与原发性肺癌直接相关的SNP位点。采用弱工具变量可能引起相应偏倚, 因此通过计算F值评估工具变量与暴露因素的关联强度计算公式为: $F = [R^2 \times (N-1-K)] / [K \times (1-R^2)]$, $R^2 = MAF \times (1-MAF) \times \beta / SD$, 式中N为暴露的样本量; K为纳入的SNP个数; R^2 为筛选SNP所解释的变异比例; β 为油性坚果摄入量相关SNP的效应值; MAF为最小等位基因频率^[11]。F值<10提示可能存在弱工具变量偏倚, 值越大表示工具变量是弱工具变量的可能性越小。

1.3 因果效应估计

MR-Egger回归、加权中位数及逆方差加权法检验油性坚果摄入量与原发性肺癌之间的因果关联。由于两样本随机化的暴露和结局分别来自于不同的样本, 可能存在人群不同等异质性, 因此需要通过Cochran's Q异质性检验说明, 若 $P < 0.05$ 说明存在异质性。

1.4 可靠性评价

水平多效性是指当暴露因素为0时, 工具变量对结局仍有效应, 即通过混杂因素起作用, 如果水平多效性检验显著, 说明工具变量不仅与暴

表 2 5 种 MR 分析方法 OR 估计值及 95%CI
 Tab. 2 OR estimates and 95% CI of five MR analysis methods

method	b	se	P	OR 95%CI	Q值	Q_pval
MR Egger	-0.47	0.89	0.611	0.63(0.11, 3.65)	99.32	0.15
Weighted median	-1.07	0.29	0.001	0.34(0.19, 0.61)	100.01	0.16
Inverse variance weighted	-1.14	0.21	0.001	0.32(0.21, 0.48)		
Simple mode	-0.99	0.76	0.192	0.37(0.08, 1.64)		
Weighted mode	-0.99	0.73	0.173	0.37(0.08, 1.51)		

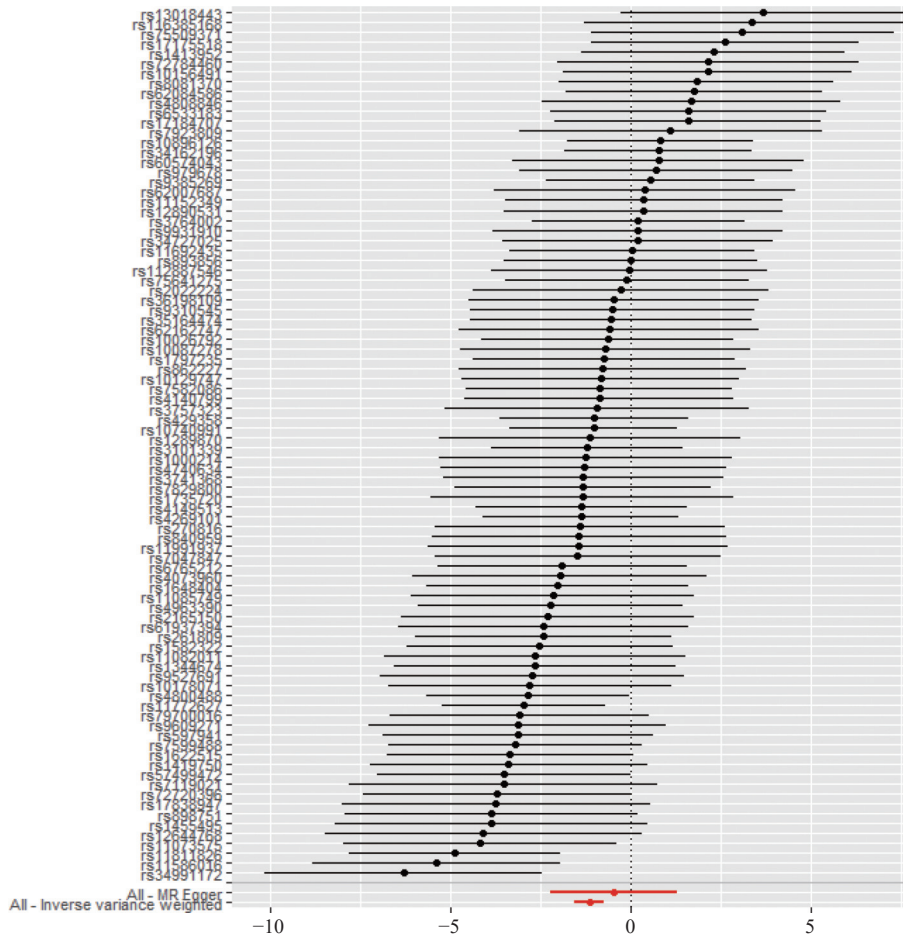


图 1 孟德尔随机化效应森林图
 Fig. 1 Mendelian randomization effect forest

03, 接近于 0, 标准误为 0.001, $P=0.441$, 见图 3。MR-PRESSO 包 global test 检验 $P=0.161$, 同样表明油性坚果摄入量对原发性肺癌的因果效应分析结果受基因水平多效性影响的可能性较小。利用留一法对 IVW 法的结果进行敏感性分析发现, 当逐个移除每个 SNP 位点时, 剩余的 90 个 SNP 与 IVW 法合并的 OR 效应值接近, 且黑点连线流畅, 没有发现显著的异常偏倚位点, 代表之前由 IVW 法得出的效应 OR 值相对稳健, 见图 4。综合而言, 这一敏感性分析表明 IVW 法的结果相对可靠。

3 讨论

3.1 主要创新点

本研究以大样本的 GWAS 汇总数据为基础, 应用两样本孟德尔随机化方法探究油性坚果(核桃、榛子、杏仁和花生)摄入量与原发性肺癌发病风险之间的因果关系, 结果表明油性坚果摄入量与原发性肺癌发病风险之间存在显著负向因果关系, 随着油性坚果摄入量每增加一个标准差原发性肺癌的发病风险下降 0.68 倍, 可以认为坚果摄入量

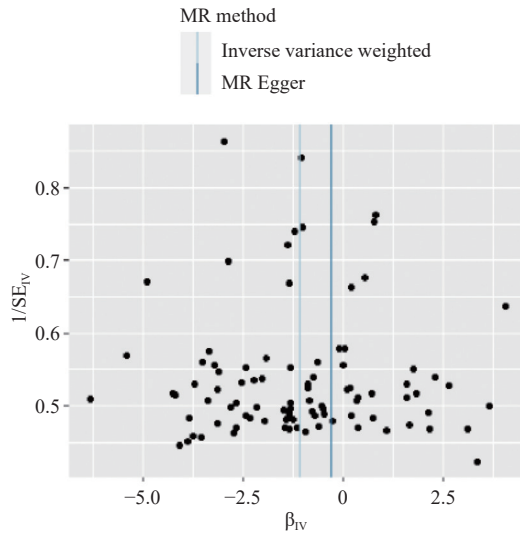


图 2 异质性检验漏斗图

Fig. 2 Heterogeneity test funnel plot

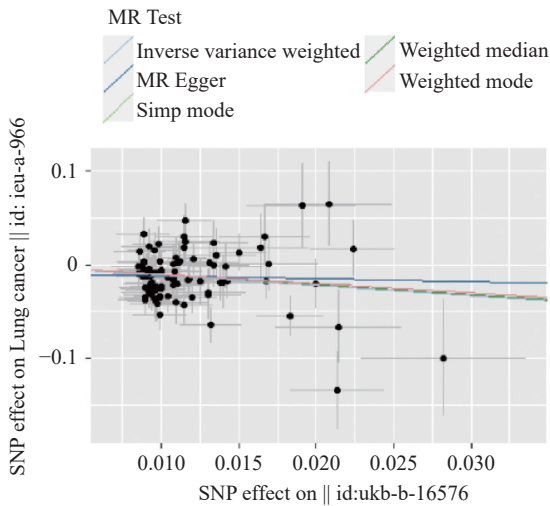


图 3 工具变量的孟德尔随机化效应散点图

Fig. 3 Scatter plot of Mendel's randomization effect

是原发性肺癌的保护因素。

3.2 坚果的营养成分及防癌机理

坚果是一种营养丰富的食物，主要含有碳水化合物、蛋白质和脂肪。此外，还富含多种维生素(如维生素B和维生素E)、微量元素(如磷、钙、锌和铁)以及多酚类化合物，特别值得一提的是油性坚果所含的脂肪酸以不饱和脂肪酸为主，如亚麻酸和亚油酸等，这些脂肪酸是人体必需的^[14]，坚果的丰富营养成分使其成为一种健康的食物选择。动物研究表明多不饱和脂肪酸(polyunsaturated fatty acids, PUFA)具有抗肿瘤和抗炎特性，1项对121970名研究参与者有关膳食PUFA摄入量与肺癌风险之间的前瞻性队列研究发现总PUFA以及n-6 PUFAs和n-3 PUFA之间的比率与肺癌风险呈负相关^[15]，特别是对于鳞癌和小细胞肺癌^[16]。

锌为DNA结合蛋白的重要组成部分，如参与DNA修复的铜/锌超氧化物歧化酶，锌在转录因子功能、抗氧化防御和DNA修复中起着重要作用，饮食中锌的缺乏会引起自由基应激反应，导致单链和双链DNA断裂，损伤DNA，从而增加癌症发展的风险，而核桃、葵花子、松子、腰果、杏仁以增加锌的人体含量，血锌通过减慢端粒损耗和调节癌症相关基因的表达可预防肺癌^[17]。多酚被认为具有防癌和抗癌的作用，其机理主要涉及防突变、抗氧化、免疫调节、抑制肿瘤细胞增殖、诱导癌细胞凋亡、抑制致癌基因的表达以及调控肿瘤细胞信号转导等方面。部分坚果富含多酚类物质，如儿茶素、黄酮类、花青素，儿茶素可通过抗氧化、抑制肿瘤新生血管生成、阻滞肿瘤细胞分裂周期、诱导癌细胞凋亡及调控microRNA等方面多途径、多靶点地发挥防癌抑癌作用^[18]。黄酮类通过靶向蛋白激酶C (protein kinase C, PKCs) 干扰受体酪氨酸激酶级联反应 (receptor tyrosine kinase, RTKs) 的激活，参与调控细胞分化、增殖、存活、迁移，抑制肿瘤发生^[19]。

3.3 孟德尔随机化及本研究的局限性

孟德尔随机化是基于全基因组关联分析数据，以单核苷酸多态性SNP作为天然的工具变量，用以推断因果关系的新颖流行病学方法，相较于队列研究等观察性研究，暴露在受精时便已确定，不易受到混杂因素及反向因果的干扰，因而能够有效减少偏倚。本研究筛选出与坚果摄入量紧密关联的SNP作为工具变量，在满足三大假设的前提下运用MR模拟伪随机化试验分析坚果摄入量与原发性肺癌的因果关系，排除模型内生性问题，得到的模型参数为一致性估计，结论更加可靠^[20]。但本研究存在相对不足之处，因GWAS数据库主要为欧洲血统，缺乏其他人种数据，结论外推性可能不佳，其次MR分析运用的2阶段最小二乘法本质上仍属于线性模型，无法拟合非线性关系^[21]，而真实世界中变量关系往往为非线性，另外因无个体属性数据，MR分析只能得到平均因果效应，无法进行亚组分析，更难以推断个体因果效应。

本研究为接下来的研究供了流行病学调查基础，油性坚果摄入量与原发性肺癌的因果关联有待于大样本随机对照试验及不同种族背景人群数据来进一步证实。此外，未来的研究还需继续探究坚果预防原发性肺癌的作用机制，需要强调的是，油性坚果最佳用量尚不明确，也是未来研究

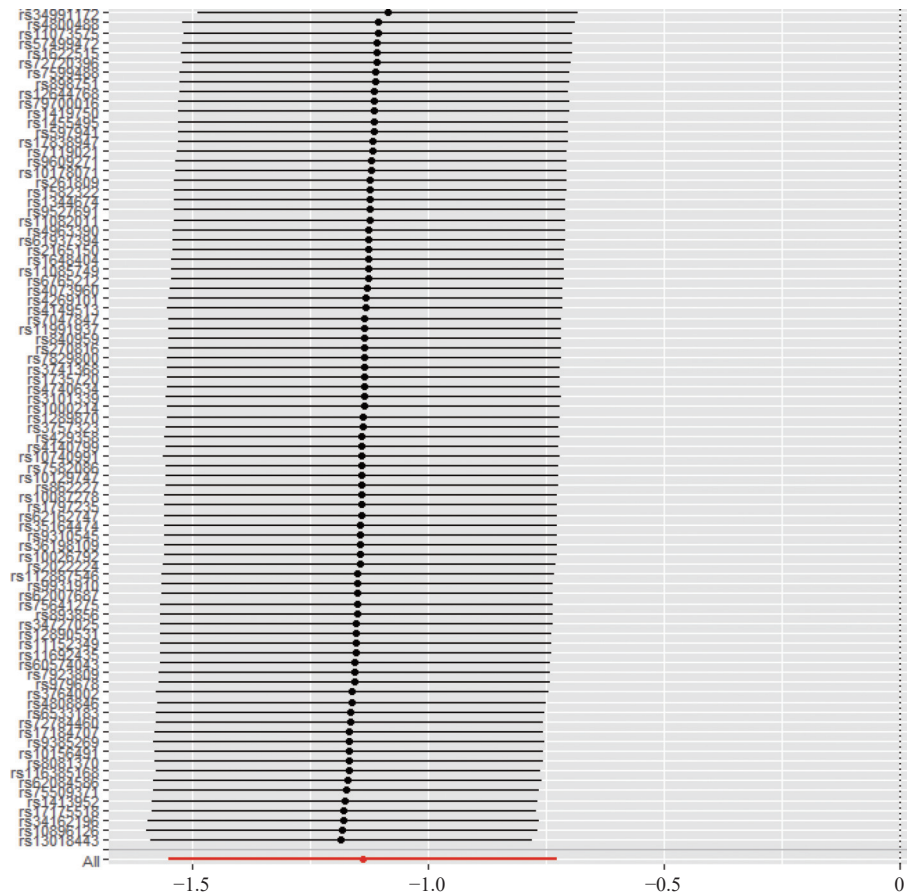


图4 留一法敏感性分析

Fig. 4 Sensitivity analysis of leave-one-out method

的重要着眼点。简而言之，本研究采用两样本孟德尔随机化方法对油性坚果摄入量与原发性肺癌之间的因果关系进行初步探究，结果表明增加油性坚果摄入量可能降低原发性肺癌的罹患率，值得宣传及推广。

[参考文献]

- [1] Sung H, Ferlay J, Siegel R L, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2021, 71(3): 209-249.
- [2] Xia C, Dong X, Li H, et al. Cancer statistics in China and United States, 2022: Profiles, trends, and determinants[J]. *Chin Med J (Engl)*, 2022, 135(5): 584-590.
- [3] Chen S, Cao Z, Prettner K, et al. Estimates and projections of the global economic cost of 29 cancers in 204 countries and territories from 2020 to 2050[J]. *JAMA Oncol*, 2023, 9(4): 465-472.
- [4] 开比努尔·再比布力, 姑丽切克然·艾斯克, 韩加. 坚果营养成分及保健功效的研究[J]. *粮食与食品工业*, 2022, 29(3): 32-36.
- [5] Chang C, Xinyan G, Yan H, et al. Association between nut consumption and cancer risk: A meta-analysis[J]. *Nutrition and cancer*, 2022, 75(1): 11-13.
- [6] Nieuwenhuis L, van den Brandt P A. Nut and peanut butter consumption and the risk of lung cancer and its subtypes: A prospective cohort study[J]. *Lung Cancer*, 2019, 128(1): 57-66.
- [7] Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenotype[J]. *Elife*, 2018, 7(1): e34408.
- [8] Rasooly D, Patel C J. Conducting a reproducible mendelian randomization analysis using the R analytic statistical environment.[J]. *Current Protocols in Human Genetics*, 2019, 101(1): 82.
- [9] Paraskevas A, Bellou E, Keramida E, et al. Bronchial asthma and lung cancer[J]. *Current Respiratory Medicine Reviews*, 2021, 17(2): 78-82.
- [10] Forder A, Zhuang R, Souza V, et al. Mechanisms contrib-

- uting to the comorbidity of COPD and lung cancer[J]. *Int J Mol Sci*, 2023, 24(3): 2859–2861.
- [11] Levin M G, Judy R, Gill D, et al. Genetics of height and risk of atrial fibrillation: A Mendelian randomization study[J]. *PLoS Med*, 2020, 17(10): e1003288.
- [12] Fu R, Kim S J. Inferring causality from observational studies: the role of instrumental variable analysis[J]. *Kidney Int*, 2021, 99(6): 1303–1308.
- [13] Wang X, Jiang Y, Zhang N R, et al. Sensitivity analysis and power for instrumental variable studies[J]. *Biometrics*, 2018, 74(4): 1150–1160.
- [14] 周启武, 唐兴莹, 李志宏, 等. 5种市售坚果类产品部分营养成分及重金属元素分析[J]. *安徽农业科学*, 2022, 50(19): 193–197.
- [15] Luu H N, Cai H, Murff H J, et al. A prospective study of dietary polyunsaturated fatty acids intake and lung cancer risk[J]. *Int J Cancer*, 2018, 143(9): 2225–2237.
- [16] Yang J J, Yu D, Takata Y, et al. Dietary Fat intake and lung cancer risk: A pooled analysis[J]. *J Clin Oncol*, 2017, 35(26): 3055–3064.
- [17] Bai Y, Wang G, Fu W, et al. Circulating essential metals and lung cancer: risk assessment and potential molecular effects[J]. *Environ Int*, 2019, 127(6): 685–693.
- [18] 李晓阳, 吴志平, 王梦馨, 等. 表没食子儿茶素没食子酸酯抗癌分子机制及其应用的研究进展[J]. *草药*, 2019, 50(13): 3217–3229.
- [19] Zanoaga O, Braicu C, Jurj A, et al. Progress in research on the role of flavonoids in lung cancer[J]. *Int J Mol Sci*, 2019, 20(17): 4291–4293.
- [20] Baer Benjamin R, Strawderman Robert L, Ertefaie Ashkan, et al. Discussion on Instrumental variable estimation of the causal hazard ratio[J]. *Biometrics*, 2023, 79(2): 554–558.
- [21] Fei W, Dylan S, Nandita M. A general approach to evaluating the bias of 2-stage instrumental variable estimators[J]. *Statistics in Medicine*, 2018, 37(12): 1997–2015.

征稿启事

为积极支持和推动昆明医科大学学科建设和发展,进一步提升《昆明医科大学学报》的办刊质量,不断扩大本刊的学术影响力,特对高质量稿源进行广泛征集。(1)国家自然科学基金资助课题的综述类论文可以在学报正刊发表;(2)对国家自然科学基金资助课题、云南省自然科学基金资助课题及昆明医科大学“十三五”省级、校级重点学科立项建设的研究论文,给予优先刊登。另外,投稿时需同时附:投稿单位证明、论文诚信承诺书以及基金项目证明。欢迎广大临床科研教学人员、硕士及博士研究生踊跃投稿。

网上投稿 <http://kmykdx.cnjournals.cn>, 电话: 0871—65936489, 65393133, 65922951, 65922874。

昆明医科大学学报编辑部

2024年1月1日