

基于双样本孟德尔随机化方法分析 2 型糖尿病与肾癌的因果关系

吴思宇¹ 陈小楠¹

[摘要] **目的:**采用双样本孟德尔随机化探究 2 型糖尿病和肾癌的因果关系。**方法:**从 GWAS Catalog 数据库中筛选与 2 型糖尿病相关的单核苷酸多态性位点($P < 5 \times 10^{-8}$)作为工具变量,用逆方差加权法作为主要的分析方法,并采用 MR-Egger 回归和加权中位数法作为补充。比值比(odd ratio,OR)值和 95%置信区间(confidence interval,CI)用于评价 2 型糖尿病和肾癌之间是否存在因果关系。采用 MR 多效性检测水平多效性,采用留一法检测敏感性。**结果:**从 UKB 数据库男性和女性肾癌数据集中分别提取了 152 个 SNP 作为工具变量。MR-Egger 回归、加权中位数和逆方差加权方法得到 OR 值和 95%CI:女性 1.27(95%CI:1.04~1.55),差异有统计学意义($P = 0.0035$);男性 1.04(95%CI:0.98~1.10),差异无统计学意义($P = 0.33$)。mr_pleiotropy_test 显示水平多效性不显著。在进行留一敏感性分析后,没有观察到任何特定 SNP 位点对结果的显著影响,表明结果具有高度的稳定性和可靠性。**结论:**双样本孟德尔随机化分析显示 2 型糖尿病与女性肾癌可能存在因果关联。然而,需要进一步验证更大的样本量和来自不同种族背景的个人数据,以加强研究结果。

[关键词] 女性肾癌;2 型糖尿病;孟德尔随机化分析;因果推断

DOI:10.13201/j.issn.1001-1420.2024.11.011

[中图分类号] R737.11 **[文献标志码]** A

Causal relationship between type 2 diabetes and renal cell carcinoma based on a two-sample Mendelian randomization study

WU Siyu CHEN Xiaonan

(Department of Urology, Shengjing Hospital, China Medical University, Shenyang, 110000, China)

Corresponding author: CHEN Xiaonan, E-mail: chenxn@cmu.edu.cn

Abstract Objective: To investigate the potential causal relationship between type 2 diabetes and renal cell carcinoma by a two-sample Mendelian randomization method. **Methods:** Single nucleotide polymorphisms(SNPs) associated with type 2 diabetes($P < 5 \times 10^{-8}$) were screened from the GWAS Catalog database as instrumental variables. The inverse variance weighting method was used as the main analysis method, and MR-Egger regression and weighted median method were used to supplement the explanation. Odd ratio(OR) values and 95% confidence interval(CI) were used to evaluate the causal relationship between type 2 diabetes and renal cell carcinoma. Level pleiotropy was tested by mr_pleiotropy_test, and sensitivity was tested by leave-one-out method. **Results:** From the UKB database male and female renal cell carcinoma data sets, 152 SNPs were extracted as instrumental variables, respectively. The MR-Egger regression, weighted median, and inverse variance weighting methods were used to obtain OR value and 95%CI: female 1.27(95%CI:1.04-1.55) with statistical significance($P = 0.0035$), male 1.04(95%CI:0.98-1.10) without statistical significance($P = 0.33$). Mr_pleiotropy_test showed that horizontal pleiotropy was not statistically significant. After leave-one-out sensitivity analysis was performed, no significant effect of any specific SNP locus was observed, indicating a high degree of stability and reliability of the results. **Conclusion:** The two-sample Mendelian randomization analysis showed a possible causal association between type 2 diabetes and renal cell carcinoma in women. However, further validation with larger sample sizes and data from individuals with different ethnic backgrounds is needed to strengthen the findings.

Key words renal cell carcinoma in women; type 2 diabetes mellitus; Mendelian randomization analysis; causal inference

流行病学统计学分析表明,2 型糖尿病(type 2

diabetes mellitus,T2DM)可能是肾癌患者进展和死亡的独立危险因素^[1]。回顾性分析发现糖尿病与肾癌病理分期升高相关^[2],合并 T2DM 的肾癌

¹中国医科大学附属盛京医院泌尿外科(沈阳,110000)
通信作者:陈小楠,E-mail:chenxn@cmu.edu.cn

患者总生存期更短,复发和死亡风险更高^[3]。一项性别分层队列研究发现,T2DM 只与女性肾癌风险增高独立相关,与男性无关^[4]。这些研究提示,T2DM 可能是肾癌的危险因素,并且这种风险大小在男性和女性中可能不同。然而,关于因果关系的结论不能仅基于观察性研究中存在的关联,因为它受到许多混杂因素和反向因果关系的影响,没有足够的说服力^[5]。为了排除过多的混杂偏倚,我们使用孟德尔随机化(Mendelian randomization,MR)探索 T2DM 与肾癌的因果关系。作为一种强有力的因果推断方法,MR 利用遗传变异(SNP)作为工具变量(IVs),识别风险因素与疾病之间的因果关系^[6]。

1 资料与方法

如图 1 所示,研究设计的示意图和 MR 的 3 个关键假设如下:基因变异必须是与感兴趣的暴露相关的有效工具;遗传变异独立于已知的结果危险因素;暴露与结果之间的关联仅由感兴趣的暴露介导,没有多效性,也就是说,没有其他将遗传变异与结局联系起来因果途径。

1.1 数据来源

用于本研究的数据集描述见表 1。T2DM 的 GWAS 数据来源于 GWAS Catalog 数据库中样本

量较大的糖尿病 GWAS 数据集 GCST90018926,其中包括 38 841 例 T2DM 和 451 248 例正常对照,共有 24 167 560 个 SNP。肾癌的 GWAS 数据来源于英国生物银行(UKB)数据库第 2 轮数据汇总,包含肾癌 774 例和 10 125 例正常对照,共有 13 791 467 个 SNP。以上所有人群均为欧洲血统,以尽量减少因人口异质性造成的潜在偏差。本分析中包含的数据已公开可得,并已在原始 GWAS 中获得伦理批准(No:21/NW/0157)和知情同意。

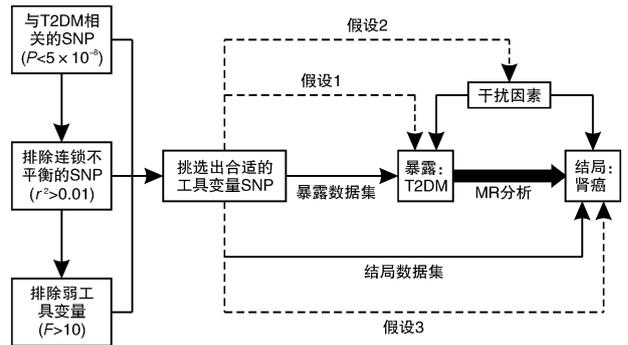


图 1 MR 研究设计示意图

表 1 数据来源

数据库	数据集	描述	样本量/例	人群	包含 SNP 数量/个	病例来源	
暴露数据集	Gwas Catalog	GCST90018926	T2DM 患者	490 089	欧洲	24 167 560	根据 ICD10 的疾病分类招募参与者
结局数据集	UKB	20001_1034	男性肾癌患者 女性肾癌患者	7 119 3 780	英国	13 791 467	参与者的登记数据和自我报告的信息(由经验丰富的护士访谈验证)

1.2 工具变量的筛选

我们从暴露的 GWAS 中提取了达到全基因组显著性的 SNP($P < 5 \times 10^{-8}$),以确保与危险因素的关联是稳健的(假设 1)。并排除了所有连锁不平衡的 SNP(在 10 000 kb 窗口内, $r^2 > 0.001$),以确保统计独立性。然后将这些 SNP 与结局 GWAS 进行匹配和协调。随后,在 PhenScanner V2 数据库(<http://www.phenoscanner.medschl.cam.ac.uk/>,于 2024 年 7 月 23 日获取)中分析与所选 SNP 相关的潜在混杂因素,重点排除相应表型与结果有相关意义的 SNP。我们计算了工具变量的强度,以防止弱工具偏差。 $F = \frac{R^2}{1-R^2} \times \frac{N-k-1}{k}$,其中 R^2 表示 SNP 解释的暴露方差, k 表示 SNP 的数量, N 表示样本量。 R^2 值是依据公式 $\frac{\beta^2}{\beta^2 + SE^2} \times N$ 算的,如果相关的 F 统计量大于 10,则认为足以减轻潜在偏差的影响,可以被纳入^[7]。

通过 UKB 数据库提取结局信息,从结局中获得满足假设的 SNP。将包含上述工具变量的暴露和结局数据集进行合并,并删除回文 SNP。最后剩余的 SNP 是关于暴露的最终工具变量。

1.3 因果效应估计

采用随机效应模型下的逆方差加权方法(IVW)进行初步分析。IVW 方法首先通过将 SNP-结局关联除以 SNP-暴露关联来估计每个 SNP 的 Wald 比值,然后将这些比值合并到固定效应 meta 分析中,其中每个比值由 SNP-结局关联方差的倒数加权得到^[8]。IVW 的前提是所有遗传变异总体上满足 MR 假设,或者所有 SNP 都是有效的工具变量。除此之外,另外 4 种孟德尔随机化方法——即 MR-egger、加权中位数、简单模式和加权模式——被用作次要方法辅助分析,以评估结果的稳健性。MR-Egger 假设 IVs 效应与结果之间的关系是线性的,即效应可以用线性回归模型来描述。该方法不仅能够估计因果效应,还能够检测和

调整水平多效性的偏倚,特别适用于存在相当数量的水平多效性 IVs 时,MR-Egger 将总效应分为直接(多效性)效应和间接(因果)效应来进行评估^[9]。加权中位数方法基于中位数的估计,允许不超过 50% 的 IVs 是无效的,因此在存在一定比例无效 IVs 的情况下仍能提供稳健的因果估计。加权模式方法则基于 IV 效应的模式值,假设大多数 IVs 的效应一致,在此前提下提供可靠结果。简单模式方法依赖于 IVs 效应的频率分布,其假设是大多数 IVs 效应接近于真实效应,因此在 IVs 效应分布较为集中的情况下更有效。简单模式和加权模式的稳健性有限,在实际研究中常用于其他分析方法的辅助证明。所有分析均使用 R 软件(4.4.1)中的“TwoSampleMR”包和“MR-PRESSO”包进行。

1.4 可靠性评价

水平多效性是指遗传变异是否通过暴露以外的途径显著影响结果。IVW 法提供的结论最准确,但是对水平多效性较敏感。因此,我们使用 MR-Egger 截距法和漏斗图^[10]作为可靠性评价。MR-Egger 方法通过在回归模型中引入一个截距项,来评估 IVs 效应与结果之间的关系,从而区分因果效应与多效性效应。如果 MR-Egger 截距偏离零, $P < 0.05$ 表示存在水平多效性,说明所选工具变量通过其他途径显著影响结果变量,这违反了图 1 中所示的假设 2 和假设 3。 $P > 0.05$ 表明除暴露外其他途径对结果变量的影响不显著^[9]。漏斗图也可以识别水平多效性,表现为图形不对称。此外,通过依次去除每个 SNP,使用 leave-one-out 方法来识别可能影响 MR 估计的异常值。由于暴露和结局分别来自不同的数据库,我们采用 Cochran's Q 值来评估可能存在的人群不同等异质性。使用 mr_heterogeneity 软件包对符合所有假设的 SNP 进行了 Cochran's Q 检验,以评估个体遗传变异之间的异质性^[11]。如果 Cochran's Q 检验结果 $P > 0.05$,则结果是异质的,表明暴露与结果的关系不受别的干扰因素的影响。Leave-one-out 分析作为一种敏感性分析,可以检测是否有异常值的 SNP。另外,我们通过检查漏斗图中的不对称性来验证结果是否稳定。若不稳定,可以通过 MR-PRESSO

方法识别异常值,并评估异常值对结果的影响。

2 结果

2.1 工具变量信息

经过剔除不显著及连锁不平衡的 SNP 后,暴露数据集 GCST90018926 剩下 183 个 SNP,经过与女性肾癌数据集匹配后得到 180 个 SNP,剔除检测到的 28 个回文 SNP,纳入共 152 个 SNP。SNP 位点 F 值分布范围 29.60 ~ 1064.32,中位数 47.58,表明不存在弱工具变量。相同步骤处理男性肾癌数据集后得到的 SNP 数量相同, F 值分布范围 11.50 ~ 1100.59,中位数 27.81。

2.2 双样本孟德尔随机化结果

加权中位数方法允许不超过 50% 的无效 IVs,而 MR-Egger 方法允许所有 IVs 无效。因此,在不存在水平多效性的情况下,IVW 法的结果更可靠。如果只有 IVW 法的 P 值具有统计学意义,结果仍可视作阳性^[12]。尽管单独的 MR-Egger、加权中位数、加权模式和简单模式结果不被视为具有统计学意义的有力证据,但 IVW 的结果在以上分析方法中得到验证时,结论更具说服力。

我们的统计结果显示,使用 IVW 和 MR-Egger 法,基因预测的 T2DM 与女性肾癌有统计学意义相关(表 2),并且其他方法的效应 b 值方向都与 IVW 的 b 值方向相同,结果存在统计学意义($P < 0.05$),可以认为 T2DM 是女性肾癌的危险因素。

值得注意的是,虽然加权中位数、加权模式和简单模式方法的 P 值均大于 0.05(分别为 0.075 1、0.062 6 和 0.349 3),但它们的结果并不削弱 IVW 和 MR-Egger 方法的显著性。这 3 种方法对无效 IVs 的稳健性更强,因此结果更为保守。另一方面,尽管本研究中使用的 IVs 满足与暴露强关联的要求($F > 10$),但有相当一部分 F 统计量低于 100,这可能是导致加权中位数、加权模式和简单模式没有得到阳性结果的原因。综上,尽管有 3 项分析方法的 P 值大于 0.05,但 IVW 和 MR-Egger 方法的显著结果仍然能够支持暴露与结果之间存在显著因果关系的结论。

另外,我们的统计结果显示,T2DM 与男性肾癌没有统计学意义的因果关系(表 3)。

表 2 女性肾癌患者中 5 种 MR 分析方法结果

方法	b	SE	P 值	OR(95%CI)	Q	Q P 值
IVW	0.24	0.10	0.003 5	1.27(1.04~1.55)	132.90	0.85
MR-Egger	0.42	0.03	0.037 1	1.05(1.44~1.61)	132.24	0.85
加权中位数	0.27	0.09	0.075 1	1.31(1.10~1.56)		
加权模式	0.38	0.15	0.062 6	1.46(1.09~1.96)		
简单模式	0.36	0.04	0.349 3	1.43(1.32~1.55)		

表 3 男性肾癌患者中 5 种 MR 分析方法结果

方法	<i>b</i>	SE	<i>P</i> 值	OR (95% CI)
IVW	0.04	0.03	0.33	1.04(0.98~1.10)
MR-Egger	0.19	0.05	0.65	1.21(1.10~1.33)
加权中位数	-0.01	0.04	0.11	0.99(0.92~1.07)
加权模式	-0.06	0.19	0.18	0.94(0.65~1.37)
简单模式	-0.11	0.11	0.17	0.99(0.80~1.23)

2.3 可靠性分析结果

女性患者 T2DM 与肾癌 MR 效应的散点图见图 2。leave-one-out 敏感性分析(图 3)显示,没有单个 SNP 不成比例的影响暴露和结局之间的正向因果关系,表明之前由 IVW 法得出的效应 OR 值相对稳健可靠。图 2 中的 MR-Egger 的截距项 Intercept 为 -1.63×10^{-5} , 接近于 0, 标准误为 2.01×10^{-5} , $P=0.42$, 表明没有水平多效性, 图 4 异质性检验漏斗图也佐证了这一点。并且, 漏斗图和 Cochran's Q 检验结果(表 2)都证明作为工具变量的这些 SNP 之间不存在异质性。

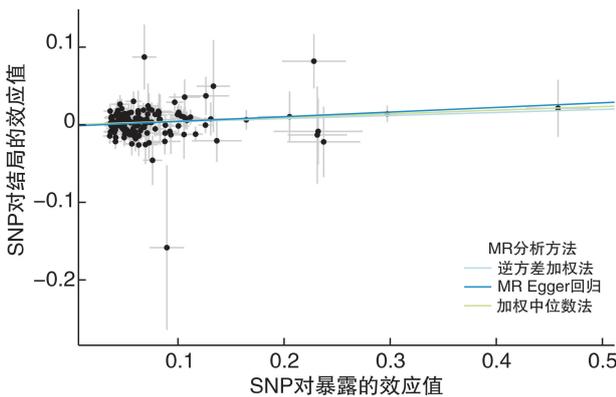


图 2 女性患者中工具变量的 MR 效应散点图

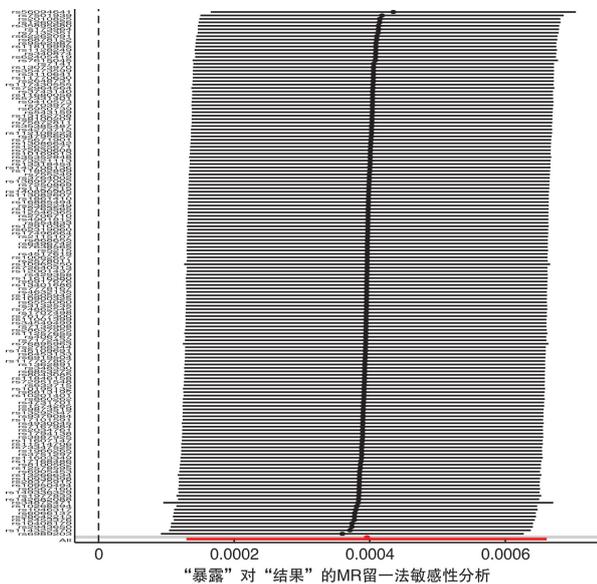


图 3 leave-one-out 敏感性分析结果

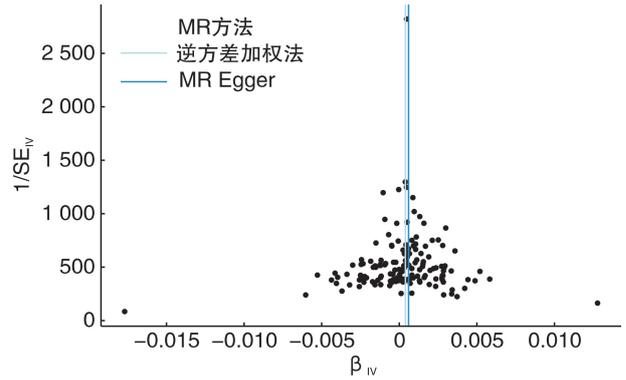


图 4 异质性检验漏斗图

3 讨论

我们的双样本孟德尔随机化分析首次尝试在 T2DM 和肾癌中进行性别分层研究, 从基因水平发现了 T2DM 和女性肾癌之间的因果关系, 而这种关系在男性中不存在。这一点可能从性别特异性的肾癌组织分型得到解释。既往研究证明, 男性和女性肾癌患者在多种亚型(比如乳头状、嫌色细胞癌和易位性肾癌)上的组织学分布不同。尤其在肾透明细胞癌和乳头状肾细胞癌中, 性别特异性基因表达明显^[13]。这可能是导致 T2DM 在 2 种不同性别的肾癌人群中因果效应不同的原因之一。提示后续研究可以着力于将肾癌结局进行更细致的分型, 在基因分层分析中我们可能会得到更多阳性结果。另外, 反向关联(肾癌→糖尿病)也只在女性人群中观察到: 美国的一项回顾性研究发现, 有肾癌病史的女性患糖尿病的比例高于男性^[14]。这些结果似乎表明, 女性中肾癌和糖尿病的关联比男性中更密切和显著, 这一现象背后的原因还有待我们深入研究。

从病理层面来说, 我们已经知道特定部位的肾损伤可导致某些类型的肾癌。例如, 肾癌可发生在先前受损的近端小管(S1/S2 段)^[15]。糖尿病可能从血糖和循环胰岛素浓度 2 个方面来参与肾损伤, 继而导致肾癌的发生和发展^[16-17]。糖尿病和肥胖患者近端小管代谢负荷增加, 这可能会加剧导致肾癌的皮质损伤。另外, 糖尿病与肾癌有相似的分子背景, 比如蛋白激酶 B(Akt)/哺乳动物雷帕霉素靶蛋白(mTOR)通路的过度活化^[18]。Akt 激酶负责 TSC2 蛋白的磷酸化或失活, 从而激活负责生长、增殖的途径^[19]。糖尿病患者 Akt 活性升高, 导致 TSC2 蛋白浓度降低, mTOR 活性升高, 这些过程都会刺激肾小管上皮细胞的增殖, 导致肾癌发生^[20]。这些研究为我们进一步探索 T2DM 促肾癌发生的机制提供了潜在的方向。

本研究的局限性在于, UKB 数据库肾癌样本量较少, 且缺少亚洲人群。当有效样本量较少时, SNP 遗传性估计可能存在偏差。另外, 如前文所

述,缺少肾癌具体分型的 GWAS 数据。如果能得到每一种临床病理分型的数据,进行基因分层分析,或许能帮助我们更全面地厘清 T2DM 与肾癌的因果关联。

综上所述,我们的发现为糖尿病和肾癌的研究提供了流行病学调查基础,T2DM 与肾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,还需继续探究 T2DM 导致肾癌发生的作用机制。在抗糖尿病药物可能对癌症治疗有益的背景下^[21-22],这些研究有望为糖尿病和肾癌患者的临床监测和用药管理提供理论指导,改善远期生存率。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 施旭,冯德超,李登雄,等. 代谢综合征和肾癌的研究进展[J]. 临床泌尿外科杂志,2022,37(9):712-717.
- [2] Pruthi DK, Miller G, Ankerst DP, et al. Diabetes, obesity, and pathological upstaging in renal cell carcinoma: results from a large multi-institutional consortium [J]. J Urol, 2023, 210(5):750-762.
- [3] Vavallo A, Simone S, Lucarelli G, et al. Pre-existing type 2 diabetes mellitus is an independent risk factor for mortality and progression in patients with renal cell carcinoma [J]. Medicine (Baltimore), 2014, 93(27):e183.
- [4] Wang S, Lo Galbo MD, Blair C, et al. Diabetes and kidney cancer risk among post-menopausal women: the Iowa women's health study[J]. Maturitas, 2021, 143:190-196.
- [5] Gagnon E, Daghlas I, Zagkos L, et al. Mendelian randomization applied to neurology: promises and challenges[J]. Neurology, 2024, 102(4):e209128.
- [6] Levin MG, Burgess S. Mendelian randomization as a tool for cardiovascular research: a review[J]. JAMA Cardiol, 2024, 9(1):79-89.
- [7] Sanderson E, Spiller W, Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization [J]. Stat Med, 2021, 40(25):5434-5452.
- [8] Mounier N, Kutalik Z. Bias correction for inverse variance weighting Mendelian randomization [J]. Genet Epidemiol, 2023, 47(4):314-331.
- [9] Lin ZT, Pan I, Pan W. A practical problem with Egger regression in Mendelian randomization[J]. PLoS Genet, 2022, 18(5):e1010166.
- [10] Lu ZC, Chen YS, Tang ZC, et al. Basal metabolic rate and the risk of urolithiasis: a two-sample Mendelian randomization study[J]. World J Urol, 2024, 42(1):235.
- [11] Li N, Wang Y, Wei P, et al. Causal effects of specific gut microbiota on chronic kidney diseases and renal function—a two-sample Mendelian randomization study [J]. Nutrients, 2023, 15(2):360.
- [12] Yang WW, Yang YJ, He L, et al. Dietary factors and risk for asthma: a Mendelian randomization analysis [J]. Front Immunol, 2023, 14:1126457.
- [13] Pallauf M, Ged Y, Singla N. Molecular differences in renal cell carcinoma between males and females [J]. World J Urol, 2023, 41(7):1727-1739.
- [14] Habib SL, Prihoda TJ, Luna M, et al. Diabetes and risk of renal cell carcinoma [J]. J Cancer, 2012, 3:42-48.
- [15] Peired AJ, Lazzeri E, Guzzi F, et al. From kidney injury to kidney cancer [J]. Kidney Int, 2021, 100(1):55-66.
- [16] Wang L, Xu R, Kaelber DC, et al. Glucagon-like peptide 1 receptor agonists and 13 obesity-associated cancers in patients with type 2 diabetes [J]. JAMA Netw Open, 2024, 7(7):e2421305.
- [17] Ghareghomi S, Arghavani P, Mahdavi M, et al. Hyperglycemia-driven signaling bridges between diabetes and cancer [J]. Biochem Pharmacol, 2024, 229:116450.
- [18] Das A, Reis F. mTOR signaling: new insights into cancer, cardiovascular diseases, diabetes and aging [J]. Int J Mol Sci, 2023, 24(17):13628.
- [19] Liu BM, Paudel S, Flowers WL, et al. Uterine histotroph and conceptus development: III. adrenomedullin stimulates proliferation, migration and adhesion of porcine trophectoderm cells via AKT-TSC2-MTOR cell signaling pathway [J]. Amino Acids, 2023, 55(6):743-756.
- [20] Zhan SZ, Bai XJ, Zhao YQ, et al. TGFB1 promotes proliferation and epithelial-mesenchymal transition in renal cell carcinoma through PI3K/AKT/mTOR/HIF-1 α pathway [J]. Cancer Cell Int, 2024, 24(1):265.
- [21] Garczorz W, Kosowska A, Francuz T. Antidiabetic drugs in breast cancer patients [J]. Cancers (Basel), 2024, 16(2):299.
- [22] Ahmmed R, Hossen MB, Ajadee A, et al. Bioinformatics analysis to disclose shared molecular mechanisms between type-2 diabetes and clear-cell renal-cell carcinoma, and therapeutic indications [J]. Sci Rep, 2024, 14(1):19133.

(收稿日期:2024-08-20)