

• 论 著 •

乳腺癌脑转移差异基因的生物信息学分析*

吴双丽¹, 徐敬宣², 邢 龙^{2△}

(青岛滨海学院附属医院; 1. 神经内科; 2. 肿瘤内科, 山东 青岛 266500)

[摘要] **目的** 利用生物信息学方法筛选乳腺癌脑转移瘤差异表达的核心基因。**方法** 从 GEO 数据库下载 GSE125989 数据集, 运用 GEO 在线分析工具 GEO2R 分析数据集的差异基因。利用在线工具 DAVID 对差异表达基因进行 GO 功能注释和 KEGG 通路富集分析, 将获得的差异表达基因数据输入 STRING 数据库并通过 Rstudio 进行可视化。利用 Rstudio 的 CytoHubba 插件筛选核心基因。利用 Kaplan-Meier Plotter 在线工具验证核心基因与总生存期的关系。**结果** 经 GEO2R 工具筛选出 22 277 个基因, 进一步筛选集中获得差异表达基因 74 个, GO 功能注释发现, 差异表达基因主要在细胞外基质组织等方面显著富集; KEGG 通路主要在蛋白质消化吸收等通路富集。通过 CytoHubba 插件以度筛选出 PPI 网络中排名前 10 位的差异表达基因为核心基因。生存分析显示 DCN、ELN 与患者总生存期具有显著的相关性。**结论** 本研究共筛选出 2 个与乳腺癌脑转移预后相关的核心基因, 分别为 DCN、ELN。

[关键词] 乳腺癌脑转移; 差异表达基因; 预后; 生物信息学方法

DOI: 10. 3969/j. issn. 1009-5519. 2023. 18. 002 **中图分类号:** R741. 02; R563

文章编号: 1009-5519(2023)18-3065-05 **文献标识码:** A

Bioinformatics analysis of differential genes in breast cancer brain metastases*

WU Shuangli¹, XU Jingxuan², XING Long²

(1. Department of Neurology; 2. Department of Oncology, Affiliated Hospital of Qingdao Binhai University, Qingdao, Shandong 266500, China)

[Abstract] **Objective** To screen the differentially expressed core genes in breast cancer brain metastases by bioinformatics. **Methods** The GSE125989 dataset was downloaded from Gene Expression Omnibus (GEO) database, and the differential genes of the dataset were analyzed by GEO2R, an online analysis tool of GEO. The online tool DAVID was used to perform GO functional annotation and KEGG pathway enrichment analysis for differentially expressed genes. The obtained differentially expressed gene data were input into Search Tool for Recurring Instances of Neighbouring Genes (STRING) database and visualized by Rstudio. The core genes were screened using CytoHubba plugin of Rstudio. Kaplan-Meier Plotter online tool was used to verify the relationship between core genes and overall survival. **Results** A total of 22 277 genes were screened by GEO2R, and 74 differentially expressed genes were further selected. GO functional annotation showed that differentially expressed genes were significantly enriched in extracellular matrix tissues. The KEGG pathway was mainly enriched in protein digestion and absorption pathways. The top 10 differentially expressed genes in PPI network were selected as the core genes by CytoHubba plugin. Survival analysis showed that DCN and ELN were significantly correlated with overall survival. **Conclusion** In this study, two core genes related to the prognosis of breast cancer brain metastasis have been screened, namely DCN and ELN.

[Key words] Breast cancer brain metastases; Differentially expressed gene; Prognosis; Bioinformatics method

* 基金项目: 山东省医药卫生科技发展计划项目(202203100054); 青岛西海岸新区 2020 年度科技惠民项目(2020-44)。

作者简介: 吴双丽(1988—), 本科, 主治医师, 主要从事神经系统疾病的基础与临床研究。△ 通信作者, E-mail: xinglong0307@163.com。

脑转移是一个毁灭性的恶性肿瘤并发症,影响约 20% 的癌症患者^[1]。乳腺癌是女性最常见的恶性肿瘤,近年来随着新疗法的引入,乳腺癌的治疗发生了巨大变化。但是由于患者生存期的延长及血脑屏障的存在,脑转移的发病率持续上升,这给肿瘤学专家带来了挑战,同时也导致患者生存时间和生存质量显著下降^[2]。从历史上来看,将患有脑转移的乳腺癌患者排除在临床试验之外,使得治疗选择更加有限。低存活率及较高的致残率迫使人们寻找新的治疗方法,以提高治疗成功的概率。近年来,分子靶向药物在众多恶性肿瘤的治疗中取得了骄人的成绩,但分子靶向治疗在乳腺癌脑转移中研究尚少,因此研究乳腺癌脑转移发生、发展的分子机制,进而制定有针对性的治疗策略将造福众多乳腺癌脑转移患者。微阵列技术已被用于众多恶性肿瘤的致癌基因研究中,发现了许多具有重要科研及临床价值的核心基因,本研究通过对乳腺癌脑转移组织基因芯片进行生物信息学分析,以期发现具有与患者预后相关核心基因。

1 材料与方法

1.1 基因芯片数据集选取 以“breast cancer”和“brain metastases”为检索词在线数据库 GEO (<http://www.ncbi.nlm.nih.gov/geo>) 上检索基因芯片数据集,从检索出的数据集中选取 GSE1259892 (平台:GPL571HG-U133A_2] Affymetrix Human Genome U133A 2.0 Array) 进行研究,该芯片集中有乳腺癌组织及乳腺癌脑转移组织样本各 16 例。

1.2 方法

1.2.1 筛选差异基因 分别通过在线工具 GEO2R (<http://www.ncbi.nlm.nih.gov/geo/geo2r>) 选出芯片集中乳腺癌脑转移组织与乳腺癌组织的差异表达基因 (DEGs)。筛选条件为 $|\logFC| > 2$ 和校正 $P < 0.05$ 。以 $\logFC > 2$ 、 $\logFC < -2$ 作为上下调基因的筛选标准。

1.2.2 DEGs 的 KEGG 和 GO 富集分析 通过在线

数据库 DAVID (<https://david.ncifcrf.gov/home.jsp>) 对筛选得到的 DEGs 进行 GO (Gene Ontology) 功能富集分析和 KEGG (Kyoto Encyclopedia of Genes and Genomes) 通路富集分析,以校正 $P < 0.05$ 为差异有统计学意义。GO 功能富集分析主要包括生物过程 (BP)、细胞成分 (CC)、和分子功能 (MF)。

1.2.3 PPI 网络构建及核心基因筛选 使用在线工具 STRING 构建 DEGs 的蛋白互作网络 (PPI), 并利用 CytoHubba 插件进一步对差异基因编码蛋白相互作用进行可视化处理,利用 CytoHubba 插件筛选度 (Degree) 位于前 10 位的为核心基因。

1.2.4 核心基因生存分析 将 10 个核心基因导入在线工具 Kaplan-Meier Plotter (<http://kmplot.com/analysis/index.php?p=background>) 选择 OS 进行生存分析,以 $P < 0.05$ 为条件筛选出与预后明显相关的核心基因。

2 结果

2.1 筛选差异表达基因 将芯片结果标准化后,鉴定出 GSE125989 中共筛选出 72 个 DEGs,其中上调基因 6 个,下调基因 66 个。见表 1。

2.2 差异基因 GO 及 KEGG 富集分析 GO 功能注释显示 BP 在细胞外基质组织、胶原原纤维组织、细胞黏附等方面显著富集;CC 主要在胞外区、细胞外基质、细胞外的外来体等方面显著富集;MF 主要在细胞外基质结构成分、细胞外基质具有抗拉强度的结构成分、胶原结合区等方面显著富集。见图 1A。KEGG 通路主要在蛋白质消化吸收、ECM 受体的交互作用、癌症中的蛋白聚糖等方面富集 (图 1B)。

2.3 PPI 蛋白互作网络构建 STRING 数据库构建 DEGs 的 PPI 网络,并利用 CytoHubba 插件以进一步对 DEGs 编码蛋白相互作用进行可视化处理,上、下调的基因被标记为红色和绿色 (图 2)。利用 CytoHubba 插件以度来筛选 PPI 网络中相互作用排名前 10 位的 DEGs 作为核心基因。见图 3。

表 1 差异表达基因

基因表达情况	基因
上调	JMJD6 OPA1 ZNF226 GFAP GFAP AQP4
下调	DPT CILP POSTN ADH1B JCHAIN MFAP5 GREM1 LUM GAS1 COL6A3 GREM1 POSTN AEBP1 IGKC IGF1 AC-TA2 IGKC COL1A1 VCAN COL3A1 COL14A1 COL1A1IGK MFAP5 COL15A1 RUNX1T1 FABP4 COL3A1 COL1A2 ELN SGCD FBN1 IGK SPON1 VCAN COL3A1 HSPG2 IGK CTSK MFAP4 ADIPOQ COL14A1 IGF1 COL1A2 CTGF CD3E DPT COL1A1 CXCL12 COL5A2 SULF1 LRRC15 IGHG1 COL6A1 HTRA1 IGF1 CCL19 THBS2 DCN SPARC MMP2 F2RL2 TPSB2 SERPINF1 SPARC COL8A2

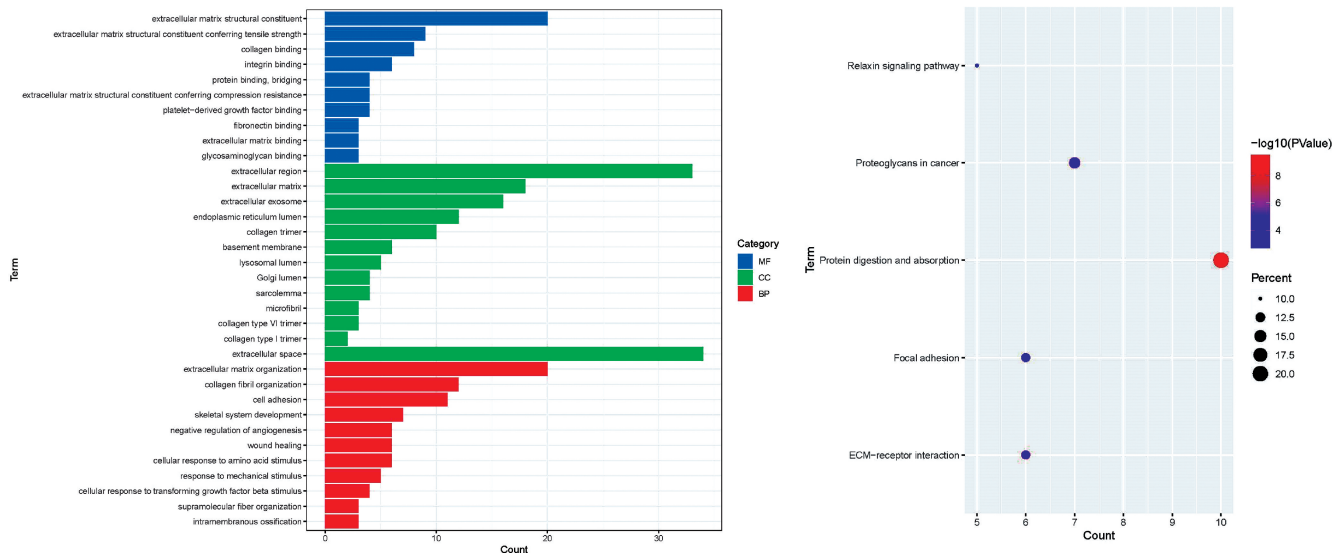


图 1 差异基因 GO 和 KEGG 富集分析

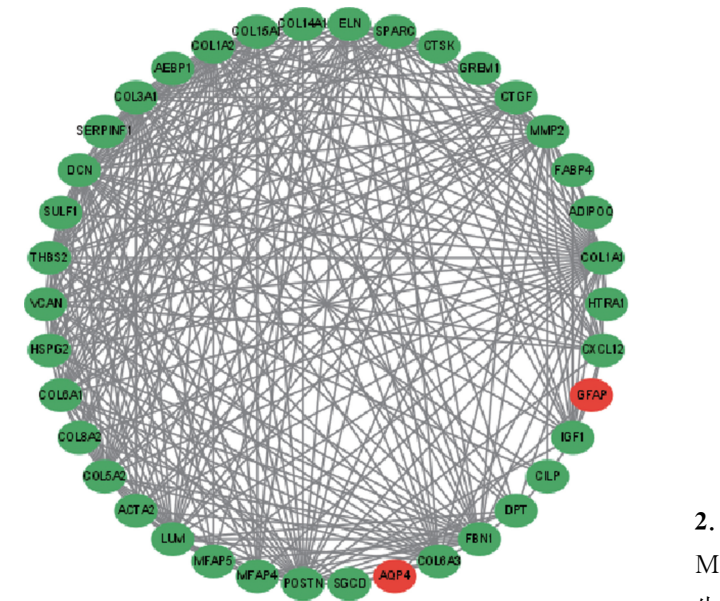


图 2 PPI 网络交互图

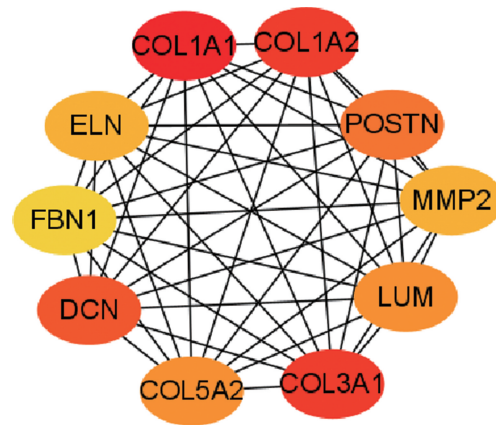


图 3 核心基因 PPI 网络交互图

2.4 核心基因生存分析 使用在线工具 Kaplan-Meier Plotter 进行生存分析,得到核心基因与患者总生存期(OS)关系,其中 DCN、ELN 与患者 OS 具有显著相关性。见图 4。

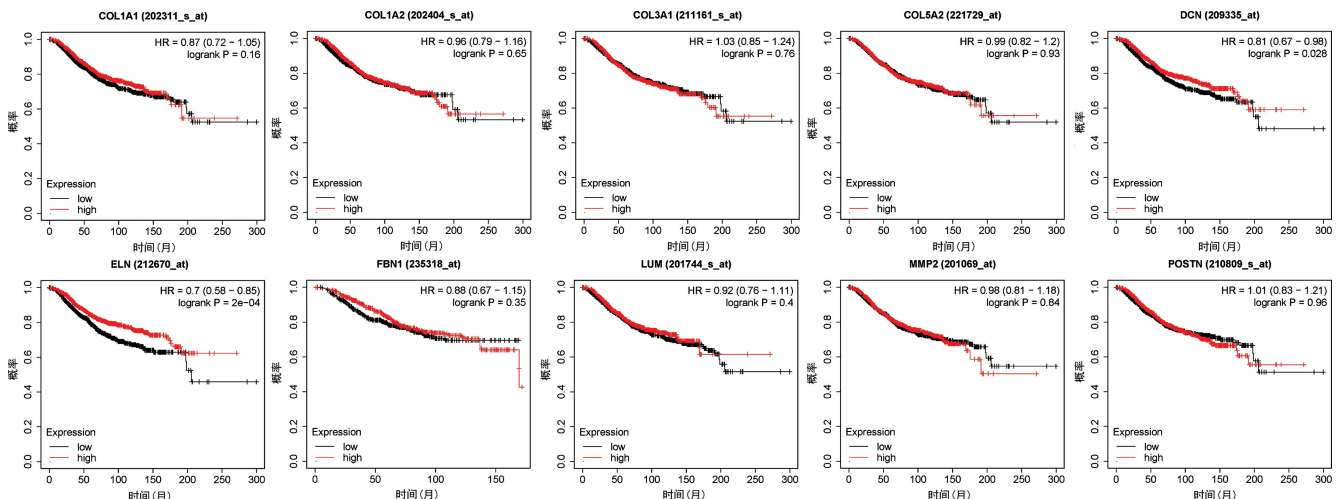


图 4 核心基因生存分析

3 讨 论

目前,脑转移瘤的确切发生率尚不清楚,但由于现有治疗手段的进步使得许多癌症患者生存期更长、预后更好;另外,监测成像技术的使用、人口老龄化、人群健康意识提高及肿瘤学家的认识水平提高,脑转移瘤的发生率正在逐步上升,脑转移的发生率在实体肿瘤患者中约占 30%。此外,尸检队列的数据表明,脑转移瘤的真实发生率在某些原发癌患者中较高^[3-4]。恶性肿瘤中最易出现脑转移瘤的前 3 位是肺癌、乳腺癌和黑色素瘤,约占全部脑转移瘤的 67%~80%^[5]。流行病学结果显示,小细胞肺癌或非小细胞肺癌患者在诊断时有较高的脑转移率,而黑色素瘤患者则有较高的脑转移风险^[6]。尽管脑转移在儿童中很少见,但在儿童神经母细胞瘤、肉瘤、肾母细胞瘤、黑色素瘤和生殖细胞肿瘤患者中都表现出较高的脑转移的倾向^[7-8]。

DCN(decorin)一方面可抑制肿瘤细胞的增殖、转移,另一方面可抑制肿瘤组织中的血管生成^[9]。有研究发现,肺腺癌组织中 DCN 表达显著降低,进一步细胞增殖实验表明,对肺腺癌细胞系 A549 和 H157 细胞的 DCN 进行过表达处理,细胞系的增殖速率可被显著抑制^[10]。HU 等^[11]研究发现,DCN 在炎性乳腺癌患者的肿瘤组织中表达显著下调,在 IBC 异种移植小鼠模型中,过表达 DCN 可通过破坏 E-cadherin 的稳定性和抑制 EGFR/ERK 信号通路减弱肿瘤细胞的侵袭及迁移能力。ZHAO 等^[12]通过将一种基于 DCN 蛋白研制的溶瘤毒素(rAd. DCN),通过对免疫活性乳腺肿瘤模型瘤内和静脉给予 rAd. DCN,发现与静脉输注相比,瘤内注射 rAd. DCN 可产生更高水平的转基因表达,诱发更强的肿瘤溶瘤作用,而静脉输注可导致肺部病毒积聚量增加,可起到防止肿瘤肺转移的作用,肺组织中 DCN 的表达增加了 Th1 细胞因子的表达,如白细胞介素(IL)-2、IL-12 和肿瘤坏死因子 α ,降低了 Th2 细胞因子的表达,如转化生长因子 β 和 IL-6。此外,rAd. DCN 治疗可诱导强烈的全身炎症反应和上调 CD8⁺ T 淋巴细胞,提示 rAd. DCN 在抑制乳腺癌细胞生长的同时可抑制肺转移的发生。另外,还有学者研究发现,在三阴性乳腺癌小鼠模型中,在高剂量照射下肿瘤周围肌肉比常规照射下分泌更多的 DCN,提示肿瘤周围肌肉分泌的 DCN 可能是超分割放疗能更有效抑制 TNBC 生长的原因之一^[13]。此外,DCN 表达上调可减弱三阴性乳腺癌的肺转移。

ELN 编码的蛋白为弹性蛋白,是弹性纤维的主要

成分。在许多恶性肿瘤中,重建的 ECM 表达高水平的弹性蛋白,这可能对肿瘤的生长有积极或消极的影响。弹性蛋白与其他 ECM 组分的交联以及控制这一过程的酶都对肿瘤的发生有影响。弹性蛋白酶的中性粒细胞弹性蛋白酶,是肿瘤入侵和转移的关键驱动因素^[14]。LI 等^[15]研究发现,弹性蛋白是结直肠癌肿瘤发生的关键因素,通过对 2 个微阵列数据集中分析 ELN 基因在结直肠癌患者肿瘤组织、邻近非肿瘤结肠组织及健康对照组中的表达情况,结果发现,与癌旁非肿瘤组织和健康对照组相比,结直肠癌患者肿瘤中 ELN 基因表达增加,提示 ELN 在结直肠癌肿瘤发展和微环境中发挥着重要作用。亦有学者研究结果提示,ELN 的表达与膀胱癌^[16]、胶质瘤^[17]患者的预后有关。

综上所述,通过对乳腺癌脑转移组织与乳腺癌组织之间的差异表达基因进行生物信息学分析,发现 DCN、ELN 2 个与乳腺癌脑转移患者预后相关的核心基因,由于肿瘤位置特殊性,能选用的数据集及样本量虽较少,但仍可为乳腺癌脑转移的预后和治疗提供一定方向。

参考文献

- [1] SACKS P, RAHMAN M. Epidemiology of brain metastases[J]. Neurosurg Clin N Am, 2020, 31(4):481-488.
- [2] IWAMOTO T, NIIKURA N, OGIYA R, et al. Distinct gene expression profiles between primary breast cancers and brain metastases from pair-matched samples[J]. Sci Rep, 2019, 9(1): 13343.
- [3] TSUKADA Y, FOUAD A, PICKREN JW, et al. Central nervous system metastasis from breast carcinoma. Autopsy study[J]. Cancer, 1983, 52(12):2349-2354.
- [4] SAMPSON J H, CARTER JH J R, FRIEDMAN A H, et al. Demographics, prognosis, and therapy in 702 patients with brain metastases from malignant melanoma[J]. J Neurosurg, 1998, 88(1):11-20.
- [5] NAYAK L. Epidemiology of brain metastases[J]. Curr Oncol Rep, 2012, 14:48-54.
- [6] CAGNEY D N, MARTIN A M, CATALANO P J, et al. Incidence and prognosis of patients with brain metastases at diagnosis of systemic

- malignancy: a population-based study[J]. *Neuro Oncol*, 2017, 19(11):1511-1521.
- [7] SUKI D, KHOURY ABDULLA R, DING M, et al. Brain metastases in patients diagnosed with a solid primary cancer during childhood: Experience from a single referral cancer center[J]. *J Neurosurg Pediatr*, 2014, 14(4):372-385.
- [8] GOLDMAN S, ECHEVARRÍA M E, FANG USARO J, et al. Pediatric brain metastasis from extraneural malignancies: A review[J]. *Cancer Treat Res*, 2007, 136:143-168.
- [9] YANG Q, WANG S, WANG Y, et al. Decorin-loaded poly lactic-co-glycolic acid nanoparticles modified by anti-alpha fetoprotein antibody: preparation, proliferation inhibition and induced apoptosis effects on HepG2 cells in vitro[J]. *J Pharm Pharmacol*, 2017, 69(6):633-641.
- [10] YAN Y, XU Z, QIAN L, et al. Identification of CAV1 and DCN as potential predictive biomarkers for lung adenocarcinoma[J]. *Am J Physiol Lung Cell Mol Physiol*, 2019, 316(4):L630-L643.
- [11] HU X, VILLODRE ES, LARSON R, et al. Decorin-mediated suppression of tumorigenesis, invasion, and metastasis in inflammatory breast cancer[J]. *Commun Biol*, 2021, 4(1):72.
- [12] ZHAO H, WANG H, KONG F, et al. Oncolytic adenovirus rAd. DCN inhibits breast tumor growth and lung metastasis in an immune-competent orthotopic xenograft model[J]. *Hum Gene Ther*, 2019, 30(2):197-210.
- [13] YU Q, XIN K, MIAO Y, et al. Anti-tumor responses to hypofractionated radiation in mice grafted with triple negative breast cancer is associated with decorin induction in peritumoral muscles[J]. *Acta Biochim Biophys Sin(Shanghai)*, 2018, 50(11):1150-1157.
- [14] WANG Y, SONG E C, RESNICK M B. Elastin in the tumor microenvironment[J]. *Adv Exp Med Biol*, 2020, 1272:1-16.
- [15] LI J, XU X, JIANG Y, et al. Elastin is a key factor of tumor development in colorectal cancer[J]. *BMC Cancer*, 2020, 20(1):217.
- [16] ZHU T J, WANG H, MA T, et al. Identification of immune-related genes as prognostic factors in bladder cancer[J]. *Sci Rep*, 2020, 10(1):19695.
- [17] IAN W, CHEN K, YAN G, et al. A novel prognostic tool for glioma based on enhancer RNA-regulated immune genes[J]. *Front Cell Dev Biol*, 2022, 9:798445.

(收稿日期:2023-02-21 修回日期:2023-06-03)

(上接第 3064 页)

- [8] 顾海, 吴迪, 韩光曙, 等. 我国区域远程会诊服务平台构建研究[J]. *中国卫生政策研究*, 2019, 12(7):65-69.
- [9] SOLIMINI R, BUSARD? F P, GIBELLI F, et al. Ethical and legal challenges of telemedicine in the era of the COVID-19 pandemic[J]. *Medicina*, 2021, 57(12):1314.
- [10] 张贤凤. 医联体内 2 型糖尿病双向转诊模式效果评价[J]. *山西医药杂志*, 2019, 48(17):2170-2172.
- [11] 王浩, 朱芮, 仝武军. 疗养院适应军队“双向转诊”模式改革的做法[J]. *解放军医院管理杂志*, 2016, 23(9):834-835.
- [12] 袁姣, 叶霖, 陈训军, 等. 武汉协和京山医院四级医共体医院建设实践与成效分析[J]. *中国医院*, 2021, 25(11):49-51.
- [13] 刘博, 齐玥, 郭幽燕, 等. 某三甲医院就诊者互联网医疗接受认可程度及服务使用维度分析[J]. *中国医院管理*, 2021, 41(9):69-73.

(收稿日期:2023-08-16 修回日期:2023-09-18)