

基于高分辨率 T2WI 的影像组学对直肠癌 EGFR 表达的预测价值

胡友强¹, 罗敏¹, 董小娟³, 邹龙权², 谢刚¹, 张远林¹, 刘文军¹, 姜萍¹

(自贡市第四人民医院, 1. 放射科; 2. 病理科, 四川 自贡 643000; 3. 自贡市第三人民医院放射科, 四川 自贡 643020)

【摘要】目的: 探讨基于高分辨率 T2WI 的影像组学对直肠癌 EGFR 表达状态的预测价值。**方法:** 回顾性分析经术后病理确诊且在接受治疗前行 MRI 检查的 208 例直肠癌患者的临床及影像资料, 根据 EGFR 表达水平不同将患者分为阳性组和阴性组。在高分辨率 T2WI 图像上勾画病灶的三维容积兴趣区 (VOI) 并提取影像组学特征, 将 208 例患者分为训练集 ($n = 145$) 和测试集 ($n = 63$), 并对特征进行降维, 将降维后的特征建立支持向量机 (SVM)、逻辑回归 (LR)、随机森林 (RF) 及线性判别分析 (LDA) 四种分类器学习模型, 分别绘制训练集和测试集的受试者工作特征 (ROC) 曲线, 并获得曲线下面积 (AUC)。**结果:** 208 例患者中, EGFR 阳性表达 99 例 (47.6%)。二元 Logistic 回归分析显示低分化和淋巴结转移是 EGFR 阳性表达的独立危险因素 ($P < 0.05$)。训练集与测试集的患者在性别、年龄、TN 分期及分化程度差异均无统计学意义 ($P > 0.05$)。4 种影像组学模型均有一定的预测效能, 其中 SVM 模型训练集与测试集的诊断效能均为最高, 在训练集和测试集中的 AUC 分别为 0.803、0.725。**结论:** 基于高分辨率 T2WI 图像构建的影像组学模型对直肠癌 EGFR 表达状态具有一定预测价值。

【关键词】 直肠肿瘤; 影像组学; 磁共振成像; EGFR

【中图分类号】 R445 **【文献标志码】** A

High-resolution T2-weighted image-based radiomics in the preoperative EGFR expression of rectal cancer

HU You-qiang¹, LUO Min¹, DONG Xiao-juan³, ZOU Long-quan², XIE Gang¹, ZHANG Yuan-lin¹, LIU Wen-jun¹, JIANG Ping¹

(1. Department of Radiology; 2. Department of Pathology, Zigong Fourth People's Hospital, Zigong 643000; 3. Department of Radiology, Zigong Third People's Hospital, Zigong 643020, Sichuan, China)

【Abstract】Objective: To investigate the value of high-resolution T2-weighted (T2WI) image-based radiomics in the preoperative EGFR expression of rectal cancer. **Methods:** This was a retrospective analysis of 208 patients with a pathology confirmed rectal cancer diagnosis who underwent surgery and high-resolution T2WI magnetic resonance imaging (MRI) before surgery. Patients were divided into two groups: high EGFR expression and low EGFR expression. The volume of interest (VOI) were drawn based on high-resolution T2WI images and then radiomics features were extracted. The samples were divided into a training set ($n = 145$) and a test set ($n = 63$) for machine learning. Four machine learning classifier models were established: support vector machine (SVM), logistic regression (LR), random forests (RF), and linear discriminant analysis (LDA). Receiver operating characteristic (ROC) curves and the area under the curve (AUC) of the training set and test set were obtained. **Results:** A total of 208 patients were enrolled, including 99 patients with high EGFR expression (47.6%). Logistic regression analysis showed that low differentiation and lymph node metastasis were independent risk factors of EGFR expression ($P < 0.05$). There was no significant difference of gender, age, TN stage and differentiation degree between training set and testing set ($P > 0.05$). All the four radiomics models had certain diagnostic efficacy. SVM algorithm had the highest diagnostic accuracy, the AUCs in the training and testing groups were 0.803 and 0.725, respectively. **Conclusion:** The high-resolution T2WI image-based radiomics could be a valuable preoperative test for predicting EGFR expression of rectal cancer.

【Key words】 Rectal neoplasms; Radiomics; Magnetic resonance imaging; EGFR

结直肠癌是消化系统中常见的恶性肿瘤之一, 其发病率在世界范围内高居恶性肿瘤的第三位^[1]。

在我国,结直肠癌中有近一半为直肠癌,且近年来死亡率呈上升趋势,多数患者入院就诊时肿瘤已表现为局部进展期癌^[2],我国抗击直肠癌的形势依然十分严峻。表皮生长因子受体(epidermal growth factor receptor, EGFR)作为一种跨膜糖蛋白受体,是表皮生长因子受体家族信号通路中的重要成员,具有酪氨酸激酶活性,活化后可激活下游通路,促进血管生成,诱导细胞增殖^[3]。在许多肿瘤中均见到 EGFR 表达升高,EGFR 表达与肿瘤的发生、发展有关。随着靶向 EGFR 药物西妥昔单抗的应用,有研究^[4-5]显示,转移性结直肠癌化疗联合西妥昔单抗靶向治疗的总体生存和无进展生存均显著高于单纯化疗。因此,EGFR 表达状态对直肠癌靶向治疗及预后评估具有重要价值,但目前尚缺乏无创性预测 EGFR 表达状态的可行办法。

影像组学的快速发展使 MRI 图像的定量化信息挖掘成为可能。目前,影像组学技术已广泛应用于肿瘤的精准预测、早期诊断及预后评估等方面^[6]。目前,影像组学在直肠癌术前 TN 分期、远处转移及新辅助化疗疗效评估方面已取得诸多成果^[7-9]。也有学者^[10]开始探索影像组学在直肠癌生物学特征评估方面的价值,但对于直肠癌 EGFR 表达状态的预测的研究尚未见报道。因此,本研究试图建立并验证基于高分辨率 T2WI 的影像组学特征,以无创性预测直肠癌 EGFR 表达状态。

1 资料与方法

1.1 一般资料

回顾性分析 2018 年 6 月至 2020 年 6 月自贡市第四人民医院收治的 208 例直肠癌患者的临床及影像资料。纳入标准:(1)术前 1 个月内行 MRI 检查,且有完整临床及 MRI 资料;(2)术后病理确诊为直肠腺癌;(3)病理资料中包含 EGFR 表达状态。排除标准:(1)MRI 检查前行抗肿瘤治疗(新辅助化疗、放疗或者靶向治疗等);(2)合并其他恶性肿瘤;(3)MRI 图像质量不佳,无法进行影像组学特征提取。其中年龄(67.00 ± 10.20)岁。本研究经自贡市第四人民医院伦理委员会批准。

1.2 仪器与方法

采用西门子 3.0 T MRI 成像系统。所有研究对象检查前禁食时间 > 4 h,并在扫描前用开塞露(20 mL 甘油)清洁肠道。直肠 MR 扫描主要序列包括:T1WI(横轴面)、脂肪抑制 T2WI(矢状面)、高分辨率 T2WI(斜横轴面)、DWI 及增强扫描。高分辨率 T2WI 扫描平面垂直于病灶处肠管的长轴。扫描参数:TR 5 000 ms,TE 113 ms,FOV 160 mm,矩阵 320

× 320,层厚 3 mm,无层间距,无脂肪抑制技术。

1.3 免疫组化

病理学检查与免疫组化结果均由我院病理科医生完成,所有直肠癌肿瘤标本放置在 4% 中性甲醛固定,经石蜡包埋后连续切片,随后进行免疫组化染色。EGFR 阳性评定标准:以脂肪组织为对照,在胞质及胞膜上棕黄色分布为阳性,若标本中阳性细胞占比 ≥ 5%,为 EGFR 阳性,反之则为阴性。

1.4 图像分割及特征提取

通过图像存档和通信系统上获取 MRI 原始图像,将轴位高分辨率 T2WI 序列图像导入 3D slicer 软件(<https://www.slicer.org>),由一名 5 年以上腹部放射学工作经验的医师沿肿瘤组织边缘逐层手动勾画,由另一名具备 10 年工作经验的高年资医师审核,共同确认得到病灶感兴趣容积(volume of interest, VOI)。见图 1。应用 Radiomics 函数包提取影像组学特征。上述提取的影像组学特征,分为以下两类:(1)基于特征类,包括一阶统计量特征(first order statistic)、形状特征(shape-based)、纹理特征(texture)等;(2)基于过滤器类,包括小波分析(Wavelets)、拉普拉斯变换(Laplacian transforms)、对数特征(logarithm)、指数特征(exponential)等。最终,每例患者都提取了 1 223 个影像组学特征。

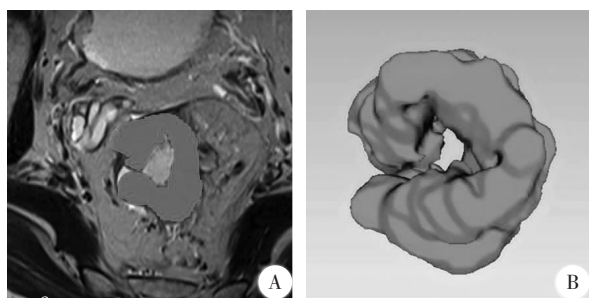


图 1 直肠癌病灶分割

患者,男性,55 岁,直肠癌 T4AN0M0 期。A. 基于高分辨 T2WI,勾画直肠癌病灶 ROI; B. 基于单层 ROI 重建得到直肠癌 VOI。

1.5 构建影像组学模型

采用 Fe Ature Explorer (Version 0.3.6) 分析影像组学特征:(1)数据预处理,以上采样法(Up Sampling)平衡样本,并以 Z-score 法对影像组学特征数据进行标准化处理;(2)特征降维,应用 Pearson 相关系数(pearson correlation coefficients, PCC)分析影像组学特征的相关性,如两个特征之一的 PCC > 0.99,则随机删除一个;(3)构建模型,建立模型之前,采用方差分析或递归特征消除方法筛选出基于高分辨率 T2WI 最优影像组学特征,分别建立支持向量机(SVM)、逻辑回归(LR)、随机森林(RF)及线性判别分析(LDA)四种分类器学习模型。为确定

模型的超参数(如特征的数量),在训练数据集上应用 5 折交叉验证法,于训练集进行训练,于测试集评估其效能。

1.6 统计学分析

采用 SPSS 23.0 软件对数据进行统计学分析。计量资料以($\bar{x} \pm s$)表示,服从正态分布的数据组间比较采用独立样本 t 检验,非正态分布的数据组间比较则采用 Mann-Whitney U 检验;计数资料以[n (%)表示,组间比较采用独立样本 χ^2 检验。以术后免疫组化结果为金标准,应用 Fe Ature Explorer 软件评估影像组学模型的效能,采用受试者工作特征(ROC)曲线、曲线下面积(AUC)、准确率、敏感度和特异度等共同评价,影响直肠癌 EGFR 表达的影响因素用 Logistic 回归分析。对观察者内和观察者

间获得病灶 VOI 的一致性采用组内相关系数(intra-class correlation coefficient, ICC)进行评价,ICC > 0.8 为一致性良好。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 一般资料

208 例患者中,EGFR 阳性表达 99 例(47.6%)。EGFR 阳性组与阴性组间患者性别、年龄差异均无统计学意义($P > 0.05$),而 T 分期、淋巴结转移、分化程度差异有统计学意义($P < 0.05$)。见表 1。二元 Logistic 回归分析显示低分化和淋巴结转移是 EGFR 阳性表达的独立危险因素($P < 0.05$)。见表 2。训练集与测试集的患者在性别、年龄、TN 分期及分化程度差异均无统计学意义($P > 0.05$)。见表 3。

表 1 EGFR 阳性组与阴性组一般资料比较[$\bar{x} \pm s, n$ (%)]

组别	性别		年龄(岁)	分化		T 分期		淋巴结转移	
	男	女		低	中高	T1~2	T3~4	无	有
EGFR 阳性($n=99$)	51(51.52)	48(48.48)	65.4 ± 11.0	19(19.19)	80(80.81)	20(20.20)	79(79.80)	49(49.49)	50(50.51)
EGFR 阴性($n=109$)	69(63.30)	40(36.70)	66.8 ± 11.3	9(8.26)	100(91.74)	36(33.03)	73(66.97)	74(67.89)	35(32.11)
t/χ^2 值	2.953	0.902	0.902	5.325		4.338		7.264	
P 值	0.086	0.368	0.368	0.021		0.037		0.007	

表 2 影响直肠癌 EGFR 表达指标二元 Logistic 回归分析

变量	OR 值	95% CI	P 值
T 分期	0.573	0.295 ~ 1.114	0.101
淋巴结转移	0.341	0.188 ~ 0.619	<0.001
分化程度	0.295	0.121 ~ 0.719	0.007

2.2 影像组学特征

观察者内和观察者间基于高分辨率 T2WI 图像勾画直肠癌病灶 VOI 的一致性良好,ICC 值为 0.90 (95% CI:0.81 ~ 0.95)。共提取 1 223 个影像组学特征,最终筛选出 6 个最优影像组学特征。见表 4。

表 3 训练集与测试集患者及病灶一般资料比较[$\bar{x} \pm s, n$ (%)]

组别	性别		年龄	分化		T 分期		淋巴结转移	
	男	女		低	中高	T1~T2	T3~T4	无	有
训练集($n=145$)	83(57.24)	62(42.76)	66.2 ± 11.4	19(13.19)	126(86.90)	41(28.28)	104(71.72)	84(57.93)	61(42.07)
测试集($n=63$)	37(58.73)	26(41.27)	65.9 ± 10.5	9(14.29)	54(85.71)	15(23.81)	48(76.19)	39(61.94)	24(38.09)
t/χ^2 值	0.40	0.931	0.931	0.053		0.445		0.287	
P 值	0.842	0.348	0.348	0.818		0.505		0.592	

表 4 基于高分辨率 T2WI 最终筛选出的影像组学特征

序号	特征
1	Original-shap-Maximum2DDiameterSlice
2	Original-glrlm-LargeAreaEmphasis
3	log-sigma-glszm-SizeZoneNonUniformityNormalized
4	Original-glrlm-LargeDependenceHighGrayLevelEmphasis
5	Original-shap-Maximum3DDiameter
6	log-sigma-glszm-ZoneEntropy

2.3 分类器模型预测表现

4 种影像组学模型均有一定的预测效能。见表 5。ROC 曲线显示,SVM 模型训练集与测试集的诊断效能均为最高:训练集 AUC 为 0.803 (95% CI: 0.720 ~ 0.874),准确率、敏感度、特异度分别为 77.9%、82.6%、73.7%;测试集 AUC 为 0.725 (95% CI:0.587 ~ 0.849),准确率、敏感度、特异度分别为 71.4%、70.0%、72.7%。见图 2。

表 5 4 种分类器影像组学模型评估直肠癌 EGFR 表达的效能

分组模型	训练集				测试集			
	AUC 值(95% CI)	准确率	敏感度	特异度	AUC 值(95% CI)	准确率	敏感度	特异度
SVM	0.803(0.720~0.874)	0.779	0.826	0.737	0.725(0.587~0.849)	0.714	0.700	0.727
LR	0.791(0.710~0.863)	0.752	0.681	0.816	0.696(0.556~0.824)	0.667	0.867	0.485
RF	0.785(0.703~0.856)	0.752	0.710	0.790	0.700(0.569~0.827)	0.668	0.733	0.606
LDA	0.802(0.718~0.870)	0.752	0.797	0.711	0.704(0.571~0.832)	0.681	0.733	0.686

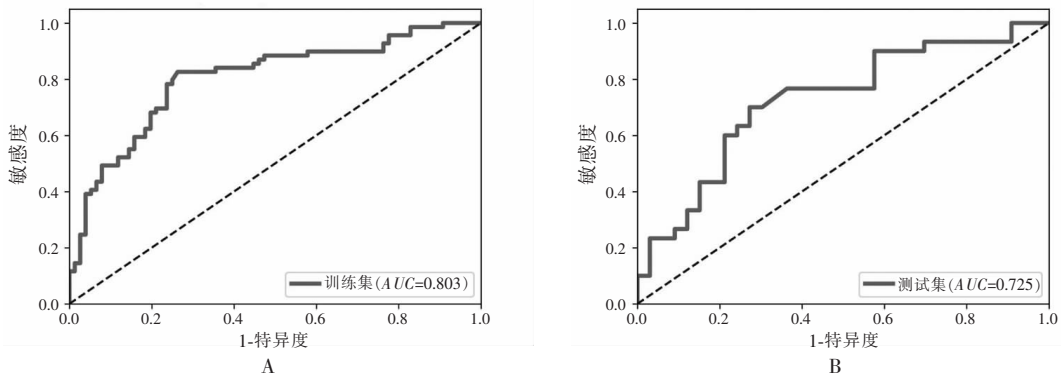


图 2 SVM 模型评估直肠癌 EGFR 表达的 ROC 曲线

A. 训练集;B. 测试集

3 讨论

EGFR 作为表皮生长因子受体家族信号通路中的重要成员,可促进肿瘤血管生成,诱导细胞增殖^[3],对直肠癌靶向治疗及预后评估具有重要价值,但目前尚缺乏无创性预测 EGFR 表达状态的可行办法。本研究基于高分辨率 T2WI 图像用直肠癌 VOI 提取特征,再通过 4 种分类器建立相应影像组学模型用于直肠癌患者 EGFR 表达的预测研究。结果显示,4 种影像组学模型均有一定的预测效能,SVM 模型训练集与测试集的诊断效能均为最高。

EGFR 广泛分布于器官组织中,在许多肿瘤中均见到 EGFR 表达升高。本研究中 EGFR 阳性表达率高达为 47.6%,表明 EGFR 表达与直肠癌的发生密切相关。有学者^[11]在非小细胞肺癌的研究中发现,在癌前病变阶段,患者的癌基因激活 C-erbB-1 基因,随后 EGFR 过度表达,引起细胞异常分裂及增殖,导致不典型增生进而产生癌变。本研究结果显示,低分化和淋巴结转移是 EGFR 阳性表达的独立危险因素,这与焦俊霞等^[12]研究基本一致,表明 EGFR 过度表达与直肠癌淋巴结转移和预后有关。

影像组学在直肠癌的研究主要集中在术前 TN 分期、远处转移及新辅助化疗疗效评估方面^[7-9]。对于直肠癌 EGFR 表达状态的预测的研究尚未见报道。本研究 4 种影像组学模型中,SVM 模

型训练集与测试集的诊断效能均为最高。SVM 是机器学习中最受关注的、性能良好的二分类器。它的思想是能够正确划分训练数据集,且在研究样本中找出几何间隔最大的超平面^[13]。王楠等^[14]应用 4 中分类方法进行建模,对直肠癌术前 T 分期进行预测,结果显示 SVM 是效果最优的算法,其 AUC 高达 0.968。本研究中 4 种分类方法,AUC 均达到 0.75 以上,准确率也都高于 0.71,这与 Lambin 等^[15]提出的观点一致,好的影像组学特征应用不同分类器计算都应该获得较好的结果。因此,我们研究基于 MRI 数据开发肿瘤本身隐藏的数据信息来预测直肠癌的 EGFR 表达状态是有价值的。

本研究存在以下不足:(1)由于直肠癌肿瘤组织形态不规则,本研究采用手动分割逐层勾画兴趣区,该方法工作量大,不适用于大规模数据处理,基于深度学习的直肠癌兴趣区自动分割方法值得未来进一步研究。(2)本研究仅选择基于高分辨率 T2WI 图像进行评估,未来研究可加入多种 MRI 序列、多种影像学方法及实验室指标,构建多模态影像组学模型,更综合、更全面地探究术前预测直肠癌 EGFR 表达的问题。

综上,基于高分辨率 T2WI 图像的影像组学模型对直肠癌术前 EGFR 表达状态预测的评估具有一定价值,从而有助于临床医生为直肠癌患者制定个体化治疗方案和预后评估。

参考文献

[1] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA: a Cancer Journal for Clinicians, 2021, 71(3): 209-249.

[2] 王锡山. 中美结直肠癌流行病学特征对比及防控策略分析[J]. 中华结直肠疾病电子杂志, 2019, 8(1): 1-5.

[3] Hynes NE, Lane HA. ERBB receptors and cancer: the complexity of targeted inhibitors[J]. Nature Reviews Cancer, 2005, 5(5): 341-354.

[4] Van Cutsem E, Lenz HJ, Köhne CH, et al. Fluorouracil, leucovorin, and irinotecan plus cetuximab treatment and RAS mutations in colorectal cancer[J]. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology, 2015, 33(7): 692-700.

[5] Rachiglio AM, Lambiase M, Fenizia F, et al. Genomic profiling of KRAS/NRAS/BRAF/PIK3CA wild-type metastatic colorectal cancer patients reveals novel mutations in genes potentially associated with resistance to anti-EGFR agents[J]. Cancers, 2019, 11(6): 859.

[6] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. European Journal of Cancer (Oxford, England; 1990), 2012, 48(4): 441-446.

[7] Hou M, Zhou L, Sun J. Deep-learning-based 3D super-resolution MRI radiomics model: superior predictive performance in preoperative T-staging of rectal cancer[J]. European Radiology, 2023, 33(1): 1-10.

[8] Liu X, Meng X, Zhang H, et al. Predicting distant metastasis and chemotherapy benefit in locally advanced rectal cancer[J]. Nature Communications, 2020, 11(1): 4308.

[9] Feng L, Liu X, Li C, et al. Development and validation of a radiopathomics model to predict pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a multicentre observational study[J]. The Lancet Digital Health, 2022, 4(1): e8-e17.

[10] Meng X, Xia W, Xie P, et al. Preoperative radiomic signature based on multiparametric magnetic resonance imaging for noninvasive evaluation of biological characteristics in rectal cancer[J]. European Radiology, 2019, 29(6): 3200-3209.

[11] Zhu YJ, Qu X, Zhan DD, et al. Specific gene co-variation acts better than number of concomitant altered genes in predicting EGFR-TKI efficacy in non-small-cell lung cancer[J]. Clinical Lung Cancer, 2021, 22(1): e98-e111.

[12] 焦俊霞, 叶晓霞. EGFR、P53、Ki-67 在结直肠癌中表达及临床病理学意义[J]. 中国老年学杂志, 2020, 40(10): 2065-2068.

[13] Anaissi A, Goyal M, Catchpole DR, et al. Ensemble feature learning of genomic data using support vector machine[J]. PLoS One, 2016, 11(6): e0157330.

[14] 王楠, 王远军, 廉朋. 基于影像组学的直肠癌术前 T 分期预测[J]. 波谱学杂志, 2022, 39(1): 43-55.

[15] Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine[J]. Nature Reviews Clinical Oncology, 2017, 14(12): 749-762.

(收稿日期: 2023-07-12

修回日期: 2023-09-30)

(上接第 343 页)

[7] 张吉超, 董跃福, 牟志芳, 等. 骨关节炎患者在不同步态角度下膝关节内部生物力学变化的有限元分析[J]. 中国组织工程研究, 2022, 26(9): 1357-1361.

[8] 杨晓露, 张津沁, 胡小卫, 等. 膝骨关节炎患者步态中膝关节三维运动特征[J]. 中国康复, 2022, 37(2): 90-94.

[9] Yamagata M, Taniguchi M, Tateuchi H, et al. The effects of knee pain on knee contact force and external knee adduction moment in patients with knee osteoarthritis[J]. Journal of Biomechanics, 2021, 123: 110538.

[10] Lee SP, Farrokhi S, Kent JA, et al. Comparison of clinical and biomechanical characteristics between individuals with lower limb amputation with and without lower back pain: a systematic review and meta-analysis[J]. Clinical Biomechanics (Bristol, Avon), 2023, 101: 105860.

[11] 赵功赫, 曲峰, 杨辰, 等. 躯干负重对步态的影响及相应补偿策略的试验研究[J]. 天津体育学院学报, 2016, 31(6): 524-529.

[12] Silva MDC, Perriman DM, Fearon AM, et al. Effects of neuromuscular gait modification strategies on indicators of knee joint load in people with medial knee osteoarthritis: a systematic review and meta-analysis[J]. PLoS One, 2022, 17(9): e0274874.

[13] Bishnoi A, Lee R, Hu Y, et al. Effect of treadmill training interventions on spatiotemporal gait parameters in older adults with neurological disorders: systematic review and meta-analysis of randomized controlled trials[J]. International Journal of Environmental

Research and Public Health, 2022, 19(5): 2824.

[14] 赵功赫, 曲峰. 女性单侧负重步态的躯干和骨盆运动学特征研究[J]. 天津体育学院学报, 2020, 35(4): 439-445.

[15] 阮家诚, 蒯声政, 王大平. 步态及影像学在膝骨性关节炎严重程度评估中的应用进展[J]. 中国康复医学杂志, 2020, 35(10): 1274-1279.

[16] 杨洋, 王亚平, 徐诚, 等. 身背负重对士兵行军步态影响试验研究[J]. 兵工学报, 2016, 37(11): 2050-2057.

[17] Doyle E, Doyle TLA, Bonacci J, et al. The effectiveness of gait retraining on running kinematics, kinetics, performance, pain, and injury in distance runners: a systematic review with meta-analysis[J]. The Journal of Orthopaedic and Sports Physical Therapy, 2022, 52(4): 192-A5.

[18] DeJong P, Hatamiya NS, Barkley LC. Running gait analysis and biomechanics[J]. Current Sports Medicine Reports, 2022, 21(4): 107-108.

[19] Anderson LM, Martin JF, Barton CJ, et al. What is the effect of changing running step rate on injury, performance and biomechanics? A systematic review and meta-analysis[J]. Sports Medicine-Open, 2022, 8(1): 112.

[20] Ma J, Chen X, Xin J, et al. Overall treatment effects of aquatic physical therapy in knee osteoarthritis: a systematic review and meta-analysis[J]. Journal of Orthopaedic Surgery and Research, 2022, 17(1): 190.

(收稿日期: 2023-09-27

修回日期: 2023-12-02)