

SHAP 值在 XGBoost MRI 模型中鉴别 腮腺恶性肿瘤与 Warthin 瘤的价值

张潇涵¹, 朱翰林², 韩志江³, 杨海⁴

(1. 杭州师范大学附属萧山医院放射科, 浙江 杭州 311200; 2. 杭州市第九人民医院放射科, 浙江 杭州 311225;
3. 杭州市第一人民医院放射科, 浙江 杭州 310006; 4. 浙江省武义县中医院放射科, 浙江 金华 321200)

[摘要] 目的: 探讨 SHAP 值在极端梯度提升 (Extreme gradient boosting, XGBoost) MRI 模型中鉴别腮腺恶性肿瘤 (Malignant tumor, MT) 与 Warthin 瘤 (Warthin tumor, WT) 的价值。方法: 回顾分析经手术病理证实的 22 例 22 枚 MT 和 38 例 51 枚 WT 的 MRI 资料, 分析瘤体形态、位置和强化方式, 以及 T₁WI、T₂WI、FS-T₁WI、FS-T₂WI 影像征象, 经单因素分析筛选有统计学意义的征象, 纳入 XGBoost 模型, 使用受试者工作特征曲线下面积 (Area under the curve, AUC)、敏感度、特异度评价模型诊断效能。通过可解释机器学习模型 (Shapley additive explanations, SHAP) 值对模型进行分析, 明确各 MRI 征象在模型中的权重。结果: 22 枚 MT 和 51 枚 WT 中, 持续性强化 ($P<0.05$)、形态不规则 ($\chi^2=20.707, P<0.05$)、非尾叶 ($\chi^2=8.911, P<0.05$)、T₂WI 高信号 ($\chi^2=7.581, P<0.05$)、FS-T₁WI 等低信号 ($P<0.05$)、FS-T₂WI 高信号 ($\chi^2=9.016, P<0.05$) 对鉴别两者有统计学意义, 且均更常见于 MT 中。将 6 项单因素纳入 XGBoost 模型分析得出 AUC 为 0.847, 敏感度为 77.3%, 特异度为 92.2%; 6 种 MRI 征象绝对平均 SHAP 值为 0.21~0.99, 其中形态不规则权重最大。结论: 形态不规则、T₂WI 高信号、持续性强化、FS-T₂WI 高信号、非尾叶、FS-T₁WI 等低信号对 MT 和 WT 鉴别的权重存在差异, 对各 MRI 征象的准确识别, 有利于对两者精准诊断。

[关键词] 腮腺肿瘤; 腺淋巴瘤; 磁共振成像

[中图分类号] R739.8; R445.2 **[文献标识码]** A **[文章编号]** 1008-1062(2024)07-0462-05

DOI:10.12117/jccmi.2024.07.002

The value of SHAP value in differentiating parotid malignant tumor from Warthin tumor in XGBoost MRI model

ZHANG Xiao-han¹, ZHU Han-lin², HAN Zhi-jiang³, YANG Hai⁴

(1. Department of Radiology, Affiliated Xiaoshan Hospital, Hangzhou Normal University, Hangzhou 311200, China;

2. Department of Radiology, Hangzhou Ninth People's Hospital, Hangzhou 311225, China;

3. Department of Radiology, Hangzhou First People's Hospital, Hangzhou 310006, China;

4. Department of Radiology, Traditional Chinese Medicine Hospital of Wuyi County, Jinhua Zhejiang 321200, China)

Abstract: Objective: To explore the value of shapley additional explanations (SHAP) value in differentiating malignant tumor (MT) and Warthin tumor (WT) of parotid gland in extreme gradient boosting (XGBoost) MRI model. **Methods:** MRI data of 22 MT in 22 cases and 51 WT in 38 cases confirmed by surgery and pathology were retrospectively analyzed. The tumors' shape, location and enhancement mode, as well as T₁WI, T₂WI, FS-T₁WI and FS-T₂WI imaging signs were analyzed. The statistically significant signs were screened by univariate analysis and included in XGBoost model, using receiver operating characteristic curve area under the curve (AUC), sensitivity and specificity to evaluate the diagnostic efficiency of the model. The model was analyzed by SHAP value, and the weight of MRI signs in the model was made clear. **Results:** Among 22 MT and 51 WT, there were statistical differences in progressive reinforcement ($P<0.05$), irregular shape ($\chi^2=20.707, P<0.05$), non-caudal lobe ($\chi^2=8.911, P<0.05$), T₂WI high signal ($\chi^2=7.581, P<0.05$), FS-T₁WI equal or low signal ($P<0.05$) and FS-T₂WI high signal ($\chi^2=9.016, P<0.05$). The AUC was 0.847, the sensitivity was 77.3%, and the specificity was 92.2% when six single factors were included in XGBoost model. The absolute average SHAP value of six MRI signs was 0.21~0.99, among which the irregular shape had the largest weight. **Conclusion:** Irregular shape, T₂WI high signal, progressive reinforcement, FS-T₂WI high signal, non-caudal lobe and FS-T₁WI equal or low signal have different weights in the differential diagnosis of MT and WT. Accurate identification of MRI signs is conducive to accurate diagnosis of MT and WT.

Key words: Parotid Neoplasms; Adenolymphoma; Magnetic Resonance Imaging

Warthin 瘤 (Warthin tumor, WT) 约占腮腺肿瘤切除术^[3]; 腮腺恶性肿瘤 (Malignant tumor, MT) 的 15%~30%^[1], 且恶变率小于 1%^[2], 通常选择单纯 发病率约 3%~5%, 常与瘤体周围组织分界不清, 易

[收稿日期] 2023-11-15; [修回日期] 2023-12-23

[作者简介] 张潇涵 (1995-), 女, 山西临汾人, 住院医师。E-mail: zhx199588@126.com

[通信作者] 韩志江, 杭州市第一人民医院放射科, 310006。E-mail: hzj1022@zju.edu.cn

侵犯面神经,需进行全腮腺伴或不伴面神经切除术,术前精准诊断二者对手术方式的选择及预后至关重要。MRI 因高软组织分辨率及多参数扫描已成为鉴别 MT 和 WT 的重要影像学方式,且常规 MRI 序列较少受机型和扫描技术的限制,在临床上应用广泛,笔者团队^[4-5]和部分学者^[6]研究表明 T₁WI、T₂WI 鉴别 MT 和 WT 有较高的敏感度(55.9%~88.9%)和特异度(78.8%~89.7%)。本文是在前期研究基础上通过对肿瘤形态、强化方式、位置及信号特征构建极端梯度提升(Extreme gradient boosting, XGBoost)模型,探讨 MRI 征象对 MT 和 WT 的鉴别诊断价值,并通过可解释机器学习模型(Shapley additive explanations, SHAP)值对模型中不同 MRI 征象的权重进行量化,为二者鉴别提供可靠的诊断依据。

1 资料与方法

1.1 临床资料

回顾性分析 2018 年 3 月—2023 年 5 月我院诊治的腮腺病变患者 60 例 73 枚腮腺病灶的临床及影像资料。WT 中男 30 例,女 8 例,平均年龄(63.6±9.1)岁;MT 中男 15 例,女 7 例,平均年龄(53.7±18.2)岁。二者性别之比无统计学意义($\chi^2=0.861, P>0.05$),但年龄之比存在统计学意义($t=2.815, P<0.05$)。

73 枚腮腺病灶中 51 枚为 WT,22 枚为 MT(包括 7 枚腺泡细胞癌,6 枚鳞状细胞癌,5 枚淋巴上皮样癌,2 枚导管癌,1 枚多形性腺瘤,1 枚小圆细胞肉瘤)。

纳入标准:①经手术及病理证实 WT 和 MT;②经腮腺 MRI 检查及具有完整的临床资料。排除标准:①检查前患者已接受治疗;②图像质量差,存在伪影,无法评估。

本研究已通过杭州市第一人民医院伦理委员会批准(伦理号:ZN-20231025-0245-01),并免除知情同意。

1.2 检查方法

采用西门子 Magnetom Verio3.0T MRI 扫描仪,使用头颈联合线圈,患者仰卧,均行平扫及增强扫描。平扫序列:快速自旋回波(TSE T₁WI、TSE T₂WI)横断位扫描、频率预饱和脂肪抑制(FS-TSE T₁WI)横断位扫描及(FS-TSE T₂WI)冠状位扫描。增强扫描序列:经静脉高压注射 0.2 mmol/kg 钆喷酸葡胺(Gd-DTPA),速率 2 mL/s,行横断位增强图像(I 期和 II 期分别为 60 s 和 125 s)。成像参数:FOV 230 mm×230 mm, NEX 2~4,层厚 4.0 mm,层间距 0.8 mm,矩阵 320×320。TSE T₁WI (TR 500 ms, TE 20 ms)、TSE T₂WI (TR 3 700 ms, TE 96 ms)、FS-TSE T₁WI (TR 500 ms, TE 19 ms)、FS-TSE T₂WI (TR 3 200 ms,

TE 29 ms)。

1.3 图像分析法

由 1 名工作 4 年的放射科医师与 1 名工作 20 年的放射科医师在对病理结果不知情的情况下,在 PACS 终端观察病灶,当意见不一致时,再次观察病灶,直至达成共识,评估强化方式及信号强度均需手动绘制感兴趣区,面积大小约 20~30 mm²。MRI 征象的观察指标有:①形态:圆形、椭圆形为规则,余为形态不规则;②位置:瘤体中心位于腮腺浅叶下极 2 cm 内为位于尾叶^[7],余位于非尾叶;③强化方式:病灶增强 II 期/增强 I 期 MRI 信号强度比值(SIR II/I), SIR II/I ≥1 为持续性强化;SIR II/I <1 为非持续性强化;④信号特点:分别测量 T₁WI、T₂WI、FS-T₁WI 上瘤体与桥脑的 SIR, SIR>1 为高信号, SIR≤1 为低信号。测量瘤体信号强度均值时,需避开囊变坏死区域,测量桥脑信号强度均值时,尽量选择同层面桥脑横断层面。FS-T₂WI 信号:瘤内见类似血管流空影的极低信号影为低信号,余为高信号。

1.4 统计学方法

采用 SPSS 26.0 (版本 26.0) 和 R 语言 (版本 4.0.1) 软件进行统计分析。符合正态分布的连续变量以均数±标准差表示,两组间比较采用独立样本 *t* 检验;非正态连续变量以中位数(四分位数间距)表示,两组间比较采用 Wilcoxon 秩和检验;分类资料以构建百分比表示,用 Fisher's 精确检验或 χ^2 检验,以 $P<0.05$ 为差异有统计学意义。以病理为金标准评价各 MRI 征象鉴别 WT 和 MT 的敏感度、特异度,敏感度=真阳性例数/(真阳性例数+假阴性例数)×100%,特异度=真阴性例数/(真阴性例数+假阳性例数)×100%。

1.5 机器学习模型的构建及解释

将有统计学意义的单因素纳入 XGBoost 机器学习模型,诊断目标为 MT,绘制机器学习模型的 ROC 曲线并计算曲线下面积 (Area under the curve, AUC)、敏感度、特异度及准确度评估模型诊断效能。使用 Bootstrapping 方法 (n=2 000), 计算模型 AUC 的标准差。利用基于博弈论中的 SHAP 值对 XGBoost 模型中各变量进行解释和计算正/负向贡献度。

2 结果

2.1 MT 和 WT 各 MRI 征象的单因素分析

22 枚 MT 和 51 枚 WT 中,形态不规则、持续性强化、T₂WI 高信号、FS-T₁WI 等低信号、FS-T₂WI 高信号、非尾叶对 MT 和 WT 均具有统计学意义 ($P<0.05$), T₁WI 等低信号对二者无统计学意义 (表 1 和图 1,2)。

表 1 MT 和 WT 各 MRI 征象对比

	形态不规则		持续性强化		非尾叶		T ₂ WI 高信号		T ₁ WI 等低信号		FS-T ₂ WI 高信号		FS-T ₁ WI 等低信号	
	是	否	是	否	是	否	是	否	是	否	是	否	是	否
MT	16	6	11	11	12	10	15	7	8	14	21	1	8	14
WT	9	42	2	49	10	41	17	34	8	43	31	20	5	46
总计	25	48	13	60	22	51	32	41	16	57	52	21	13	60
AUC	0.775		0.730		0.675		0.674		/		0.673		0.663	
χ ²	20.707		/		8.911		7.581		/		9.016		/	
P	0.000		0.000		0.003		0.006		0.067		0.003		0.016	

注:χ²中“/”为 Fisher’s Exact Test for Count Data,余为 Pearson’s Chi-squared test。

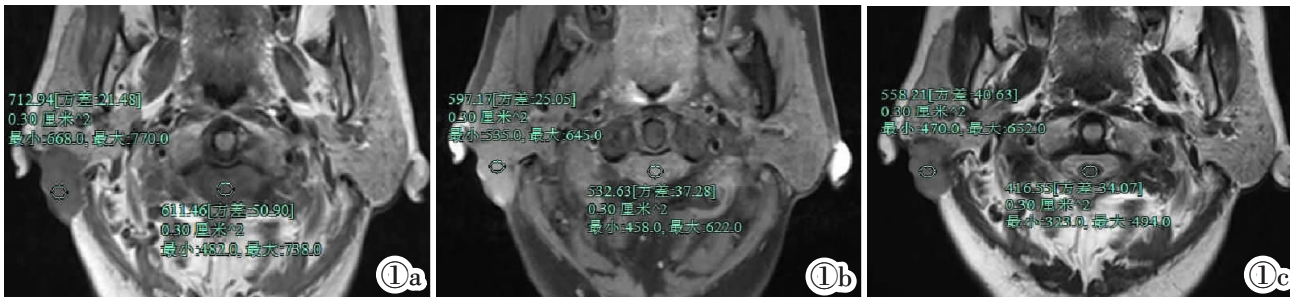


图 1 男,63 岁,导管癌。图 1a:T₁WI 横断位:瘤体信号为 712.94,同层面桥脑信号为 611.46,SIR=1.17,即 T₁WI 呈高信号。图 1b:FS-T₁WI 横断位:瘤体信号为 597.17,同层面桥脑信号为 532.63,SIR=1.12,即 FS-T₁WI 呈高信号。图 1c:T₂WI 横断位:瘤体信号为 558.21,同层面桥脑信号为 416.55,SIR=1.34,即 T₂WI 呈高信号。

Figure 1. Ductal carcinoma in a 63-year-old man. Figure 1a: T₁WI transverse position: the tumor signal is 712.94, the same level of pons signal is 611.46, SIR=1.17, so T₁WI shows high signal. Figure 1b: FS-T₁WI transverse position: the tumor signal is 597.17, the same level of pons signal is 532.63, SIR=1.12, so FS-T₁WI shows high signal. Figure 1c: T₂WI transverse position: the tumor signal is 558.21, the same level of pons signal is 416.55, SIR=1.34, so T₂WI shows high signal.

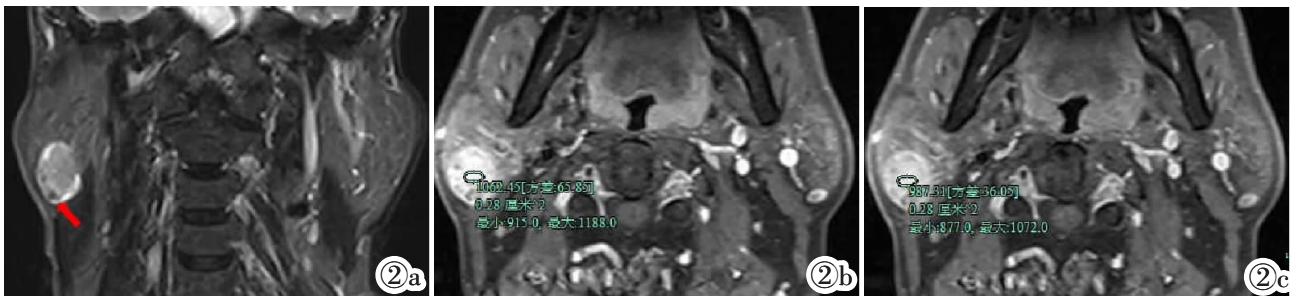


图 2 男,54 岁,WT。图 2a:FS-T₂WI 冠状位:瘤体位于尾叶,呈椭圆形,形态规则,红箭为其内极低信号影,即 FS-T₂WI 呈低信号。图 2b, 2c:增强 I 期信号为 1 062.45,增强 II 期信号为 987.31,SIR II/I=0.93,即该肿瘤呈非持续强化。

Figure 2. WT in a 54-year-old man. Figure 2a: The tumor is located in the caudate lobe, oval, regular shape, and the red arrow is the inner extremely low signal shadow, so the FS-T₂WI shows low signal. Figure 2b, 2c: Enhanced phase I signal is 1 062.45, enhanced phase II signal is 987.31, SIR II/I=0.93, so the tumor shows non-progressive reinforcement.

2.2 机器学习模型的构建

使用单因素分析存在差异的 MRI 征象构建 XGBoost 机器学习模型,该模型的 AUC 及 95%置信区间为 0.847(0.753~0.944),敏感度为 77.3%,特异度为 92.2%,准确度为 87.7%(图 3)。

2.3 SHAP 值分析解释

SHAP 值对 XGBoost 模型分析显示 MT 和 WT MRI 征象权重由大到小依次为形态不规则、T₂WI 高信号、持续性强化、FS-T₂WI 高信号、非尾叶、FS-T₁WI 等低信号。绝对平均 SHAP 值分别为 0.99、0.66、0.57、0.45、0.33、0.21。

所有 MRI 征象在 XGboost 模型中均为正向贡献。SHAP 摘要散点图中 x 轴为 SHAP 值,y 轴为 MRI 特征名称,点代表单个患者,红色代表 MRI 特征阳性,蓝色代表阴性。如瘤体形态不规则标记为红色的点,形态规则标记为蓝色的点(图 4),红色点位于右侧即为正向贡献。

3 讨论

年龄对 WT 和 MT 鉴别诊断的价值尚存在一定争议,如 Wu 等^[8]研究表明 WT 的平均年龄高于 MT,分别为 64.5 岁和 46.2 岁;而 Luna 等^[9]研究显示 WT

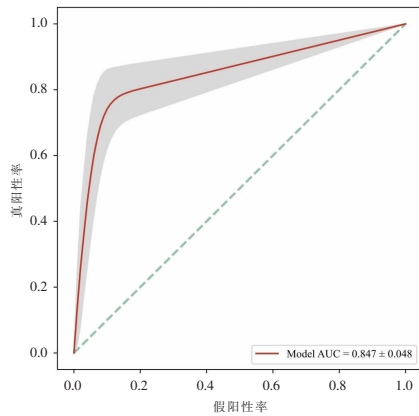


图3 XGBoost模型鉴别MT和WT的ROC曲线。

Figure 3. ROC curves of XGBoost model in differentiating MT and WT.

与高级别MT的发病年龄无统计学差异。本文示WT和MT的平均年龄分别为(63.6±9.1)岁和(53.7±18.2)岁,与Luna等和Wu等的研究存在差异,考虑与纳入的MT病理类型不一致有关。Luna等和Wu等的主要研究对象分别为未分化癌和淋巴瘤,本组MT中占比最高者为腺泡细胞癌,说明年龄鉴别WT和不同病理类型MT的诊断效能不稳定,亦说明MRI征象鉴别二者的重要价值。

腮腺良、恶性肿瘤不同生物学行为决定了瘤体形态的差异:WT为良性病变,呈膨胀性生长,其形态近似圆形;而MT细胞分裂速度不同,肿瘤生长方向不一致,其形态多不规则^[10]。麦文锋等^[11]研究中形态不规则诊断MT的敏感度和特异度为81.3%和54.3%。本文中形态不规则诊断MT的特异度(82.4%)明显高于麦文锋等的结果,但敏感度略低,为72.7%,考虑与本样本中31.8%的腺泡细胞癌有关,该肿瘤多呈圆形或椭圆形。本研究中SHAP分析显示形态不规则对MT和WT鉴别诊断的权重(0.99)最高,提示形态不规则对二者的鉴别诊断价值最高。

众所周知,自由水T₂WI呈高信号、T₁WI呈低信号,但当自由水中黏液蛋白含量增高时,T₂WI信号减低,T₁WI信号增高。WT富含黏液蛋白和淋巴滤泡组织,T₂WI呈低信号、T₁WI呈高信号^[6];而MT细胞密集,细胞外间隙狭窄导致自由水含量少,T₂WI及T₁WI均呈等低信号^[8,11-12]。在以MRI信号鉴别WT和MT方面,目前国内外已有较多报道,如Kato等^[12]统计了MT和WT的瘤体/腮腺信号比,结果显示T₁WI和T₂WI均难以鉴别二者;Wu等^[8]研究了MT和WT的T₂弛豫时间,结果亦表明二者无统计学差异。本研究结果中,T₂WI高信号可鉴别MT和WT,究其原因Kato等和Wu等研究中的MT以淋巴瘤为主,但本样本中的MT以腺泡细胞癌和鳞状细胞癌为

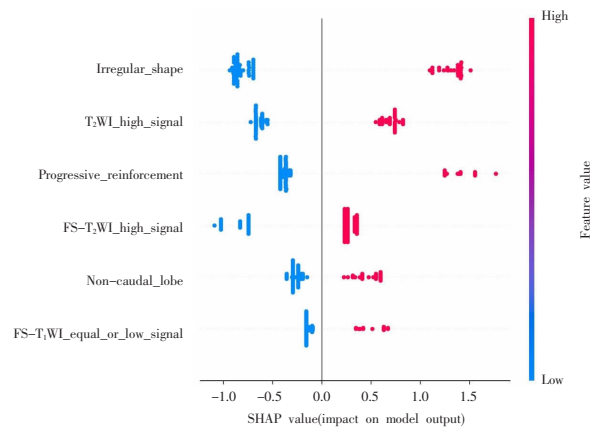


图4 SHAP摘要散点图。

Figure 4. SHAP summary scatterplot.

主,后两种肿瘤的T₂WI信号均高于淋巴瘤。SHAP值分析显示T₂WI高信号的权重(0.66)仅次于形态不规则,表明该征象对鉴别MT和WT具有一定的诊断价值。另外本研究结果提示FS-T₂WI高信号和FS-T₁WI等低信号亦均可鉴别MT和WT,但二种压脂序列的诊断权重均较低,分别为0.45和0.21,说明二者对MT和WT鉴别诊断效能有限。

肿瘤的强化模式在一定程度上反应了其内微血管情况,Gökce等^[13]及史旭波等^[14]研究表明MT和WT内微血管均密集,MT具有较低的细胞-间质比,导致造影剂流出速度慢,呈“平台型”强化,相反,WT细胞-间质比高,呈“流出型”强化,但两位学者均未探讨其对MT和WT的诊断效能。本研究显示持续性强化对诊断MT有较高的特异度(96.1%),但敏感度(50.0%)较低,说明仅以强化方式诊断MT的漏诊率较高,经SHAP分析表明持续性强化的权重(0.57)低于形态不规则和T₂WI高信号。肿瘤的好发位置与其来源密切相关,WT起源于腮腺内淋巴结或残存于邻近淋巴结内的异位涎腺组织,可想而知,WT好发于具有丰富淋巴组织的尾叶;MT病理类型繁多,可发生于腮腺任何位置,亦包括尾叶,故SHAP表明以位置鉴别MT和WT的权重(0.33)较低,该征象对二者的鉴别价值较低。

近年来,部分学者通过机器学习模型鉴别MT和WT,如Piludu等^[15]构建模型,分析肿瘤在T₂WI、T₁WI增强及ADC序列的组学特征,结果得出该模型诊断WT和MT的敏感度和特异度为87.5%和84.6%;Qi等^[16]研究T₁WI增强、FS-T₂WI及ADC序列,采用影像组学诺谟图鉴别WT和MT,其敏感度和特异度为93.3%和56.3%。本研究基于强化方式、形态、位置、信号等MRI征象构建的XGBoost模型,结果显示尽管敏感度(77.3%)略低于Piludu等和Qi等的研究,但特异度(92.2%)明显高于二者,说明通

过 XGBoost 模型可明显降低诊断 MT 的误诊率。此外,基于树模型的 XGBoost 算法能避免变量的冗余,增加模型的诊断效能,该模型诊断 MT 的 AUC 为 0.847,均高于各单一 MRI 征象的 AUC (0.663~0.775),表明多因素联合模型的诊断价值更高。

不足有以下两点:①本研究的 MT 样本量较少;②病例均来自同一医学中心,可能存在选择偏倚。

总之,本研究联合多种 MRI 征象构建的 XG-Boost 模型可有效鉴别 MT 和 WT,通过 SHAP 值可将不同 MRI 征象的权重可视化,充分掌握各 MRI 征象对二者诊断价值的差异,对临床个性化治疗决策的制定提供可靠依据。

[参考文献]

- [1]Zahran M, Alsedra S, Cope D, et al. The role of FNAC in the Diagnosis and Management of Warthin Tumour: Analysis of 74 cases[J]. *Int Arch Otorhinolaryngol*, 2021, 25(3): e379-e382.
- [2] 胡春锋,魏培英,舒艳艳,等. 常规 MRI 平扫信号强度比值在 Warthin 瘤和多形性腺瘤鉴别诊断中的价值 [J]. *中国临床医学影像杂志*, 2022, 33(11): 771-774.
- [3]Espinoza S, Felter A, Malinvaud D, et al. Warthin's tumor of parotid gland: Surgery or follow-up? Diagnostic value of a decisional algorithm with functional MRI [J]. *Diagn Interv Imaging*, 2016, 97(1): 37-43.
- [4]Wei PY, Shao C, Huan T, et al. Diagnostic value of maximum signal intensity on T₁-weighted MRI images for differentiating parotid gland tumours along with pathological correlation [J]. *Clin Radiol*, 2021, 76(6): 472.e419-472.e425.
- [5]Wei P, Shao C, Tian M, et al. Quantitative analysis and pathological basis of signal intensity on T₂-weighted MR images in benign and malignant parotid tumors[J]. *Cancer Manag Res*, 2021, 13: 5423-5431.
- [6]刘硕,史灵雪,于保婷,等. 磁共振 T₁ 序列对腮腺腺淋巴瘤与恶性肿瘤鉴别诊断价值的研究[J]. *临床放射学杂志*, 2023, 42(5): 739-742.
- [7]Monti S, Aiello M, Incoronato M, et al. DCE-MRI pharmacokinetic-based phenotyping of invasive ductal carcinoma: A radiomic study for prediction of histological outcomes [J]. *Contrast Media Mol Imaging*, 2018, 2018: 1-11.
- [8]Wu Q, Zhu LN, Jiang JS, et al. Characterization of parotid gland tumors using T₂ mapping imaging: Initial findings[J]. *Acta Radiol*, 2020, 61(5): 629-635.
- [9]Luna LP, Coffey W 3rd, Alvin MD, et al. Parotid Warthin's tumor: novel MR imaging score as diagnostic indicator [J]. *Clin Imaging*, 2022, 81: 9-14.
- [10]李星锐,赖宇林,文海洋,等. 多模态 MRI 影像特征与腮腺肿瘤良恶性判定的关系及病理结果对比[J]. *中国临床医学影像杂志*, 2021, 32(11): 777-780.
- [11]麦文锋,吴佳阳,李振宇,等. 腮腺良恶性肿瘤 MRI 征象 Logistic 回归模型的建立及其预测价值 [J]. *现代医用影像学*, 2023, 32(1): 13-18.
- [12]Kato H, Kanematsu M, Watanabe H, et al. Perfusion imaging of parotid gland tumours: Usefulness of arterial spin labeling for differentiating warthin's tumours [J]. *Eur Radiol*, 2015, 25(11): 3247-3254.
- [13]Gökce E, Beyhan M. Diagnostic efficacy of diffusion-weighted imaging and semiquantitative and quantitative dynamic contrast-enhanced magnetic resonance imaging in salivary gland tumors [J]. *World J Radiol*, 2023, 15(1): 20-31.
- [14]史旭波,黄贵生,刘扬虹,等. MR 动态增强扫描与扩散加权成像联合应用对腮腺良恶性肿瘤的诊断价值 [J]. *分子影像学杂志*, 2020, 43(2): 235-241.
- [15]Piludu F, Marzi S, Ravanelli M, et al. Mri-based radiomics to differentiate between benign and malignant parotid tumors with external validation[J]. *Front Oncol*, 2021, 11: 656918.
- [16]Qi J, Gao A, Ma X, et al. Differentiation of benign from malignant parotid gland tumors using conventional MRI based on radiomics nomogram[J]. *Front Oncol*, 2022, 12: 937050.