

doi:10.3969/j.issn.1005-3697.2024.11.002

❖ 食管癌放射治疗研究专题 ❖

# 人工勾画与三种自动勾画方法对食管癌放射治疗危及器官勾画结果的比较研究

肖剑波<sup>1</sup>, 杨明俊<sup>1</sup>, 陈旭澜<sup>1</sup>, 谢力<sup>1,2</sup>, 马代远<sup>1,2</sup>, 马强<sup>1</sup>, 张洪攀<sup>1,2</sup>, 皈燕<sup>1,2</sup>

(1. 川北医学院; 2. 川北医学院附属医院肿瘤科, 四川 南充 637000)

**【摘要】目的:** 通过比较和分析基于增强 CT 图像的人工手动勾画与三种自动勾画软件 [PVmed、United Imaging (UI)、MIM Maestro 7.1.4 (MIM)] 在食管癌放疗中危及器官 (OARs) 的勾画的精确度, 评估不同勾画方式在食管癌放射治疗靶区勾画中的优劣性及应用价值。**方法:** 回顾性分析在川北医学院附属医院放疗科接受根治性放疗的 20 例食管癌患者模拟西部 CT 定位图像, 先由高年资放疗医师在患者增强定位 CT 图像上手动勾画, 再用自动勾画软件进行勾画。分别计算不同勾画方式的豪斯多夫距离 (HD)、平均最小距离 (MDA)、形状相似性指数 (DSC) 及相似性 (JC), 并对勾画结果几何量化评价参数汇总并进行统计学分析。**结果:** 3 种自动勾画软件对心脏、肝脏、左右肺及脊髓勾画效果较好, 其 DSC 值均  $>0.8$ , 其中心脏、肝脏、肺等器官 JC 值均  $>0.8$ ; 对于心脏的勾画, 三组危及器官 HD 与 MDA 无明显差异, 但 MIM 要优于其它两者; 对于肝脏的勾画与心脏相反, PVmed、UI 要优于 MIM。虽然三种勾画软件在双肺的效果可观, 但其 HD 值比较: PVmed  $<$  MIM  $<$  UI, 分别为 (26.828, 28.371, 35.004), 其中 PVmed 优于 UI ( $P < 0.001$ ); 三种软件对于脊髓的勾画较人工勾画 DSC 无明显差异 (DSC 值均  $<0.7$ ,  $P = 0.239$ ), 但 JC 值均  $<0.8$ , 其中 HD 值比较: UI  $<$  PVmed  $<$  MIM; 而对于体积较小的甲状腺的勾画效果较差, 其 DSC 值  $<0.8$ 、JC 值  $<0.7$ , PVmed 和 MIM 均优于 UI, 且三种自动勾画软件 HD 与 MDA 有统计学意义 ( $P = 0.048$ ;  $P < 0.001$ ), 其中 PVmed 和 MIM 明显优于 UI。**结论:** 基于人工智能技术的放疗危及器官自动勾画软件 (PVmed、UI、MIM) 在食管癌 OARs 勾画对于心脏、肝脏、肺、脊髓等器官勾画准确性较高, 但对于甲状腺的勾画 UI、MIM 软件较 PVmed 精度需人工手动修改, 提高食管癌放疗靶区勾画的准确性。

**【关键词】** 食管癌; 放射治疗; 自动勾画; 危及器官

**【中图分类号】** R735.1; R730.44 **【文献标志码】** A

## Comparative study of the results of manual delineation and three methods for automatic segmentation of organs at risk in the esophageal cancer

XIAO Jian-bo<sup>1</sup>, YANG Ming-jun<sup>1</sup>, CHEN Xu-lan<sup>1</sup>, XIE Li<sup>1,2</sup>, MA Dai-yuan<sup>1,2</sup>, MA Qiang<sup>1</sup>, ZHANG Hong-pan<sup>1,2</sup>, GUI Yan<sup>1,2</sup>

(1. North Sichuan Medical College; 2. Department of Oncology, Affiliated Hospital of North Sichuan Medical College, Nanchong 637000, Sichuan, China)

**【Abstract】Objective:** By comparing and analyzing the accuracy of manual mapping based on enhanced CT images and three kinds of automatic mapping software [PVmed, United Imaging (UI), MIM Maestro 7.1.4 (MIM)] in the mapping of Organs at Risk (OARs) in esophageal cancer radiotherapy, to evaluate the advantages and disadvantages of different delineation methods in the delineation of targets for radiotherapy of esophageal cancer. **Methods:** The simulated CT images of 20 patients with esophageal cancer who received radical radiotherapy in the radiotherapy Department of the Affiliated Hospital of North Sichuan Medical College were retrospectively analyzed. The enhanced CT images were manually outlined by senior radiotherapy doctors and then outlined by automatic sketch software. The Hausdorff distance (HD), mean minimum distance (MDA), Dice Similarity Coefficient (DSC) and Jaccard Coefficient (JC) of different delineation methods were calculated, and the geometric quantitative evaluation parameters of the delineation results were summarized and analyzed statistically. **Results:** Three kinds of automatic delineation software had better delineation effect on the heart, liver, left and right lung and spinal cord. DSC all  $>0.8$ , and JC values of heart, liver, lung and other organs were all  $>0.8$ . For the delineation of the heart, there was no significant difference between HD and MDA, but MIM was better than the other two. For the delineation of the liver, PVmed and UI were superior to MIM. Although the results of the three types of mapping software were satisfactory in both lungs, their HD values PVmed  $<$  MIM  $<$  UI were (26.828, 28.371, 35.004), among which PVmed was superior to UI.

**基金项目:** 国家自然科学基金 (82271959); 四川省南充市市校科技战略合作专项 (20SXQT0254)

**作者简介:** 肖剑波 (1998 -), 男, 硕士研究生。E-mail: luck\_xiao1998@163.com

**通讯作者:** 皈燕。E-mail: guiyuan\_luck@163.com

( $P < 0.001$ ). There was no significant difference in the delineation of spinal cord by DSC compared with manual delineation (all DSC  $< 0.7$ ,  $P = 0.239$ ), but JC value was all  $< 0.8$ , and HD value (UI  $< PVmed < MIM$ ). However, for the small thyroid, the mapping effect was poor, DSC  $< 0.8$ , JC  $< 0.7$ , PVmed and MIM were superior to UI, and the three kinds of automatic mapping software HD and MDA had statistical significance ( $P = 0.048$ ,  $P < 0.001$ ), among which PVmed and MIM were significantly superior to UI. **Conclusion:** The automatic delineation software of radiotherapy endangered organs based on artificial intelligence technology (PVmed, UI, MIM) has higher accuracy for delineation of heart, liver, lung, spinal and other organs of esophageal cancer OARs, but the delineation accuracy of thyroid gland UI and MIM software needs manual modification compared with PVmed. Improve the accuracy of the delineation of radiotherapy targets for esophageal cancer.

**【Key words】** Esophageal cancer; Radiation therapy; Automatic outlining; Organ endangerment

放射治疗作为恶性肿瘤治疗的一种重要手段,约有70%的肿瘤患者在病程的不同时期需要进行放射治疗。近年来随着计算机技术和调强技术的发展,如调强放射治疗(intensity modulated radiotherapy, IMRT)利用不同强度的笔形束实现了肿瘤靶区和邻近重要组织分别接受不同剂量照射,精准勾画肿瘤靶区和OARs的轮廓是实现精准放射治疗的前提和保障<sup>[1]</sup>。OARs勾画在放疗计划的优化设计及评估环节起到关键作用。虽然放疗医师在CT图像上手动勾画靶区和OARs仍是目前放疗的“金标准”,但靶区层次较多导致勾画过程耗时长,包含大量简单重复性工作,降低了临床诊疗效率<sup>[2]</sup>,延长了病人等待治疗时间,而自动勾画技术的引入有望解决这些难题。近年来,随着人工智能技术在放疗医学领域的发展,MIM、OnQ、ABAS等自动勾画软件被广泛报道<sup>[3-5]</sup>。目前放射治疗作为食管癌最主要的治疗方式之一,在食管癌不同分期中均有重要临床意义。因食管生理结构的独特性,食管肿瘤靶区周围毗邻心脏、肺、脊髓、甲状腺等OARs,所以在放射治疗的同时也要保护正常组织,注意OARs的剂量。而自动勾画技术自应用以来,为临床靶区勾画提供了帮助的同时也反映出诸多问题。因此自动勾画技术在食管癌OARs的精确度及应用效率更值得关注。本文探究基于增强CT的手动勾画与三种自动勾画方法在胸部放疗中OARs的勾画的精确性及应用价值。

## 1 资料与方法

### 1.1 一般资料

选取2023年5月至2023年9月川北医学院附属医院收治的20例胸部肿瘤患者作为研究对象,统一取仰卧位,双手上举,在飞利浦CT模拟定位机下增强扫描,CT扫描层厚3 mm。将所有病例增强定位CT图像分别导入Monaco5.0计划系统、PVmed、UI和MIM自动勾画系统。

### 1.2 勾画工具

手动勾画工具:Elekta公司Monaco5.0计划系统中的勾画模块。由同一个放疗医师使用Monaco5.0计划系统中的勾画模块对选定的OARs进行

勾画,勾画完成后由另外两名高级职称放疗医师进行确认。所有病例所勾画的OARs为双肺、肝脏、心包及心脏、脊髓、甲状腺,勾画标准参考RTOG 1106勾画指南<sup>[6]</sup>。(1)肺:勾画时用肺窗,左右肺分别勾画,勾画时要包括所有膨胀的、塌陷的、纤维化的气肿性的组织及肺门外的小血管,但应该减去原发肺肿瘤、肺门、气管/支气管。(2)肝脏:沿肝脏外缘勾画整个肝脏,胆囊不应勾画在内。(3)心包勾画:从主动脉弓上缘1~2层开始勾画,止于横膈膈顶心尖部,心包包括心脏在内。(4)心脏勾画:肺动脉干分叉下一层开始勾画,止于横膈膈顶心尖部。(5)脊髓:勾画扫描范围内的所有层面,椎管的骨性内缘。(6)甲状腺:软组织窗勾画,位于中下颈部、甲状软骨下方、气管的两旁上达甲状软骨中部,下抵第6气管软骨环。

自动勾画软件:PVmed、United Imaging(UI组)和MIM Maestro 7.1.4(MIM组)。

### 1.3 评估参数

采用MIM自动勾画软件,获取手动勾画与三种不同自动勾画软件勾画结果的豪斯多夫距离(HD)、平均最小距离(MDA)、形状相似性指数(DSC)及相似性(JC),以评价自动勾画OAR轮廓的准确性<sup>[7-8]</sup>。(1)HD(hausdorff distance, HD)用于度量体积轮廓中2个轮廓(A和B)之间的最大失配程度;二维平面中,HD为(A和B)2个子集之间最近点所OAR有距离中的最大值,对轮廓之间个别部位的差异尤为敏感,其值越小,代表勾画效果越优<sup>[9]</sup>。(2)MDA(mean minimum distance, MDA)是代表每个体素从一个结构(A)表面到另一个结构(B)表面的平均最短距离的几何参数<sup>[9]</sup>,理想情况下MDA=0。见公式(1)-(3)。

$$HD(A, B) = \max_{\substack{a \in \partial A \\ b \in \partial B}} \{d(a, \partial B), d(b, \partial A)\} \quad (1)$$

$$DSC(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (2)$$

$$MDA(A, B) = \frac{1}{2} \left[ \frac{\sum_{a \in \partial A} d(a, \partial B)}{|\partial A|} + \frac{\sum_{b \in \partial B} d(b, \partial A)}{|\partial B|} \right] \quad (3)$$

(3) DSC 重合性 (dice similarity coefficient, DSC) 自动勾画与手工勾画的 OAR 的重合性, 用于度量标准体积重叠, 通过计算 2 个轮廓 (A 和 B) 之间相互重叠部分的体积与 2 个轮廓体积之和的比值来描述轮廓之间的重合程度, 其值越大代表勾画效果越优。见公式 (4)。其中,  $V_{auto}$  为自动勾画 OAR 体积,  $V_{manual}$  为手工勾画 OARs 体积, DSC 值范围 0 ~ 1,

DSC 越接近 1, 表示自动勾画与手工勾画的重合性越好。Fung 等<sup>[10]</sup>认为  $DSC > 0.7$  表示重合度较好。(4) 相似性 (jaccard coefficient, JC) 自动勾画与手工勾画的 OAR 的相似性系数<sup>[11]</sup>。见公式 (5)。

$$DSC = \frac{2 \times (V_{auto} \cap V_{manual})}{V_{auto} + V_{manual}} \quad (4)$$

$$JC = \frac{|V_{auto} \cap V_{manual}|}{|V_{auto} \cup V_{manual}|} = \frac{|V_{auto} \cap V_{manual}|}{|V_{auto}| + |V_{manual}| - |V_{auto} \cap V_{manual}|} \quad (5)$$

其中, JC 范围为 0 到 1, JC 值越大, 表明自动勾画与手工勾画相似性度越高。

### 1.4 统计学分析

采用 SPSS. 26.0 软件进行统计分析。计量资料采用 M (IQR) 表示, 组间比较使用 Kruskal-Wallis H 非参数检验。P < 0.05 为差异有统计学意义。

## 2 结果

### 2.1 三种自动勾画结果的几何量化评价参数比较

2.1.1 三种自动勾画结果的 HD 比较 三个勾画组勾画的右肺及甲状腺 HD 有统计学差异  $P < 0.05$ , 心脏、肝脏、左肺及脊髓 HD 无统计学差异 ( $P > 0.05$ )。见表 1。

表 1 三种软件自动勾画危及器官的 HD 比较 [M (IQR)]

| 危及器官 | PVmed         | UI             | MIM           | 克鲁斯卡尔-沃利斯 H(K) | P 值    |
|------|---------------|----------------|---------------|----------------|--------|
| 心脏   | 21.283(9.53)  | 20.149(10.82)  | 18.472(11.35) | 1.478          | 0.478  |
| 肝脏   | 20.230(9.64)  | 21.814(11.96)  | 25.379(30.26) | 3.753          | 0.163  |
| 左肺   | 20.918(10.61) | 27.133(9.76)   | 27.075(11.98) | 4.101          | 0.129  |
| 右肺   | 26.828(8.29)  | 35.004(6.00)   | 28.371(9.65)  | 16.984         | <0.001 |
| 脊髓   | 13.762(14.22) | 12.939(10.04)  | 14.287(14.21) | 0.174          | 0.917  |
| 甲状腺  | 9.329(4.08)   | 11.102(234.62) | 11.563(5.41)  | 6.062          | 0.048  |

在右肺的勾画参数 HD 比较中 PVmed 与 UI 有差异 ( $P < 0.001$ ), PVmed 优于 UI。在甲状腺的勾画参数 HD 比较中 UI 与 PVmed ( $P = 0.001$ ) 及 MIM ( $P = 0.002$ ) 有的差异, PVmed 优于 MIM, 二者明显优于 UI。见表 2 及图 1-2。

2.1.2 三种自动勾画结果的 MDA 比较 结果显示, 三个勾画组勾画甲状腺 MDA 有统计学差异 ( $P < 0.05$ ), 心脏、肝脏、双肺及脊髓 MDA 无统计学差异 ( $P > 0.05$ )。见表 3。

在甲状腺的勾画参数 MDA 比较中, UI 与 PVmed ( $P = 0.006$ ) 及 MIM ( $P = 0.019$ ) 均有差异, 且 PVmed

优于 MIM, 二者明显优于 UI。见表 4 及图 3。

表 2 三种软件自动勾画危及器官的 HD 事后检验 LSD 多重比较

| 危及器官 | 组 1   | 组 2   | P 值    |
|------|-------|-------|--------|
| 心脏   |       | PVmed | 0.462  |
|      | MIM   | UI    | 0.806  |
|      | PVmed | UI    | 0.328  |
| 肝脏   |       | PVmed | 0.614  |
|      | MIM   | UI    | 0.916  |
|      | PVmed | UI    | 0.690  |
| 左肺   |       | PVmed | 0.526  |
|      | MIM   | UI    | 0.931  |
|      | PVmed | UI    | 0.585  |
| 右肺   |       | PVmed | 0.213  |
|      | MIM   | UI    | 0.005  |
|      | PVmed | UI    | <0.001 |
| 脊髓   |       | PVmed | 0.300  |
|      | MIM   | UI    | 0.365  |
|      | PVmed | UI    | 0.895  |
| 甲状腺  |       | PVmed | 0.908  |
|      | MIM   | UI    | 0.002  |
|      | PVmed | UI    | 0.001  |

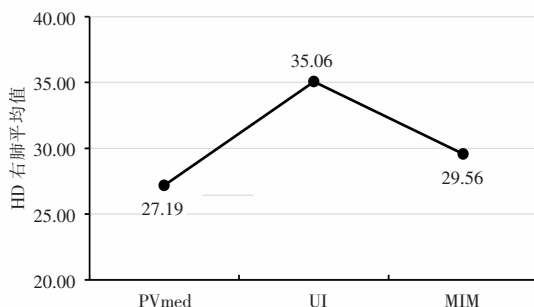


图 1 右肺 HD 平均值

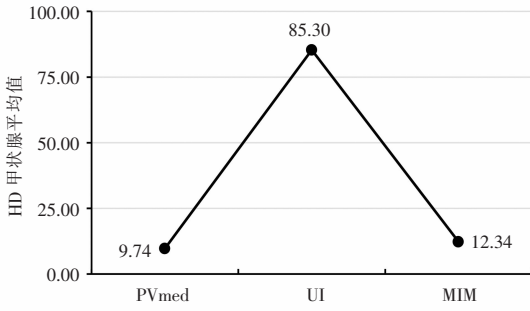


图 2 甲状腺 HD 平均值

表 3 三种软件自动勾画危及器官的 MDA [ (M(IQR)) ]

| 危及器官 | PVmed       | UI          | MIM         | 克鲁斯卡尔-沃利斯 H(K) | P 值    |
|------|-------------|-------------|-------------|----------------|--------|
| 心脏   | 2.738(1.01) | 2.738(1.01) | 2.247(1.48) | 2.273          | 0.321  |
| 肝脏   | 1.159(0.31) | 1.159(0.31) | 1.943(1.26) | 2.297          | 0.289  |
| 左肺   | 1.272(0.39) | 1.272(0.39) | 1.118(0.48) | 4.148          | 0.126  |
| 右肺   | 1.225(0.56) | 1.225(0.56) | 1.139(0.68) | 3.102          | 0.212  |
| 脊髓   | 1.084(0.36) | 1.084(0.36) | 0.906(0.3)  | 1.847          | 0.397  |
| 甲状腺  | 0.696(0.26) | 0.696(0.26) | 1.559(0.79) | 23.723         | <0.001 |

表 4 三种软件自动勾画危及器官的 MDA 事后检验 LSD 多重比较

| 危及器官 | 组 1   | 组 2   | P 值   |
|------|-------|-------|-------|
| 心脏   |       | PVmed | 0.386 |
|      | MIM   | UI    | 0.768 |
|      | PVmed | UI    | 0.565 |
| 肝脏   |       | PVmed | 0.908 |
|      | MIM   | UI    | 0.936 |
|      | PVmed | UI    | 0.972 |
| 左肺   |       | PVmed | 0.538 |
|      | MIM   | UI    | 0.696 |
|      | PVmed | UI    | 0.316 |
| 右肺   |       | PVmed | 0.750 |
|      | MIM   | UI    | 0.470 |
|      | PVmed | UI    | 0.686 |
| 脊髓   |       | PVmed | 0.213 |
|      | MIM   | UI    | 0.218 |
|      | PVmed | UI    | 0.990 |
| 甲状腺  |       | PVmed | 0.652 |
|      | MIM   | UI    | 0.019 |
|      | PVmed | UI    | 0.006 |

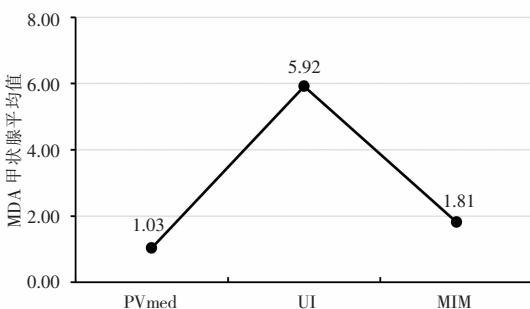


图 3 甲状腺 MDA 平均值

2.1.3 三种自动勾画结果的 DSC 比较 三个勾画组勾画的甲状腺 DSC 有统计学差异 ( $P < 0.05$ ), 心脏、肝脏、左右肺及脊髓 DSC 无统计学差异 ( $P > 0.05$ )。见表 5。

表 5 三种软件自动勾画危及器官的 DSC [  $\bar{M}(IQR)$  ]

| 危及器官 | PVmed       | UI          | MIM         | 克鲁斯卡尔-沃利斯 H(K) | P 值    |
|------|-------------|-------------|-------------|----------------|--------|
| 心脏   | 0.897(0.03) | 0.907(0.04) | 0.920(0.06) | 2.971          | 0.226  |
| 肝脏   | 0.954(0.02) | 0.954(0.02) | 0.925(0.03) | 2.595          | 0.263  |
| 左肺   | 0.945(0.02) | 0.955(0.02) | 0.954(0.02) | 5.543          | 0.063  |
| 右肺   | 0.955(0.02) | 0.956(0.02) | 0.962(0.02) | 2.774          | 0.250  |
| 脊髓   | 0.813(0.06) | 0.816(0.07) | 0.835(0.04) | 2.866          | 0.239  |
| 甲状腺  | 0.840(0.05) | 0.765(0.07) | 0.700(0.09) | 33.022         | <0.001 |

LSD 事后检验分析,在甲状腺的勾画参数 DSC 比较中, MIM Maestro 7.1.4 与 United Imaging ( $P = 0.001$ ) 及 PVmed ( $P < 0.001$ ) 有差异, PVmed 及 UI 明显优于 MIM, PVmed 与 UI 无统计学差异 ( $P = 0.098$ )。见表 6 及图 4。

表 6 三种软件自动勾画危及器官的 DSC 事后检验 LSD 多重比较

| 危及器官 | 组 1   | 组 2   | P 值    |
|------|-------|-------|--------|
| 心脏   |       | PVmed | 0.328  |
|      | MIM   | UI    | 0.745  |
|      | PVmed | UI    | 0.513  |
| 肝脏   |       | PVmed | 0.572  |
|      | MIM   | UI    | 0.593  |
|      | PVmed | UI    | 0.975  |
| 左肺   |       | PVmed | 0.204  |
|      | MIM   | UI    | 0.932  |
|      | PVmed | UI    | 0.235  |
| 右肺   |       | PVmed | 0.416  |
|      | MIM   | UI    | 0.483  |
|      | PVmed | UI    | 0.910  |
| 脊髓   |       | PVmed | 0.173  |
|      | MIM   | UI    | 0.208  |
|      | PVmed | UI    | 0.917  |
| 甲状腺  |       | PVmed | <0.001 |
|      | MIM   | UI    | 0.001  |
|      | PVmed | UI    | 0.098  |

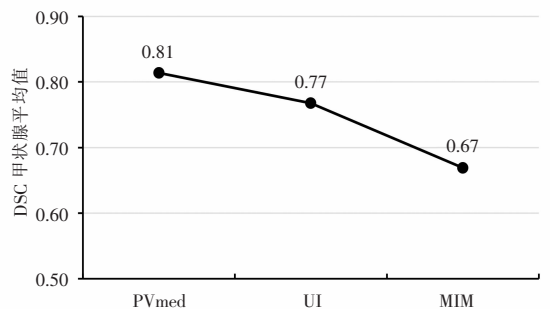


图 4 甲状腺 DSC 平均值

2.1.4 三种自动勾画结果的 JC 比较 JC 及统计学分析结果显示,三个勾画组勾画的甲状腺 JC 有统计学差异( $P < 0.05$ ),心脏、肝脏、左右肺及脊髓 JC 无统计学差异( $P > 0.05$ )。见表 7。

表 7 三种软件自动勾画危及器官的 JC [(M(IQR))]

| 危及器官 | PVmed       | UI          | MIM         | 克鲁斯卡尔-沃利斯 H(K) | P 值    |
|------|-------------|-------------|-------------|----------------|--------|
| 心脏   | 0.813(0.06) | 0.829(0.06) | 0.851(0.1)  | 2.971          | 0.226  |
| 肝脏   | 0.912(0.03) | 0.912(0.04) | 0.861(0.06) | 2.595          | 0.263  |
| 左肺   | 0.896(0.03) | 0.915(0.04) | 0.913(0.03) | 5.543          | 0.063  |
| 右肺   | 0.914(0.04) | 0.916(0.04) | 0.927(0.04) | 2.774          | 0.250  |
| 脊髓   | 0.685(0.08) | 0.689(0.1)  | 0.717(0.06) | 2.866          | 0.239  |
| 甲状腺  | 0.724(0.07) | 0.620(0.09) | 0.539(0.1)  | 33.022         | <0.001 |

在甲状腺的勾画参数 JC 比较中,MIM 与 UI( $P < 0.001$ )及 PVmed( $P < 0.001$ )有差异,PVmed 与 UI 有统计学差异( $P = 0.026$ ),PVmed 优于 UI,二者均优于 MIM。见表 8 及图 5。

表 8 三种软件自动勾画危及器官的 JC 事后检验 LSD 多重比较

| 危及器官 | 组 1   | 组 2   | P 值    |
|------|-------|-------|--------|
| 心脏   |       | PVmed | 0.277  |
|      | MIM   | UI    | 0.697  |
|      | PVmed | UI    | 0.484  |
| 肝脏   |       | PVmed | 0.349  |
|      | MIM   | UI    | 0.372  |
|      | PVmed | UI    | 0.964  |
| 左肺   |       | PVmed | 0.193  |
|      | MIM   | UI    | 0.941  |
|      | PVmed | UI    | 0.219  |
| 右肺   |       | PVmed | 0.41   |
|      | MIM   | UI    | 0.479  |
|      | PVmed | UI    | 0.906  |
| 脊髓   |       | PVmed | 0.152  |
|      | MIM   | UI    | 0.210  |
|      | PVmed | UI    | 0.854  |
| 甲状腺  |       | PVmed | <0.001 |
|      | MIM   | UI    | <0.001 |
|      | PVmed | UI    | 0.026  |

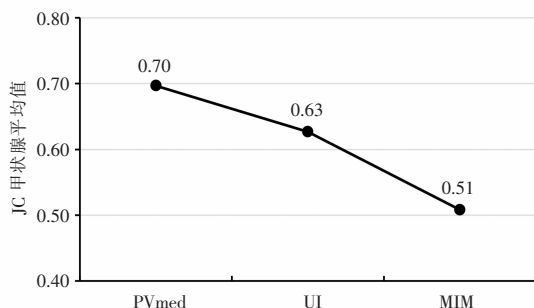


图 5 甲状腺 JC 平均值

### 3 讨论

随着放疗技术日新月异的发展,放疗逐渐进入鼻咽癌、肺癌、食管癌、宫颈癌、直肠癌等不同癌种的治疗选择中,其中鼻咽癌更是可以通过放疗达到良好的 5 年生存率<sup>[12-13]</sup>。中国作为食管癌大国,尤其是川东北地区作为食管癌高发地区,放射治疗在食管癌诊治过程中起着重要作用。临床评估 OARs 剂量的前提的经放疗医师手动勾画 OARs 轮廓,但该步骤极为耗时且可重复性差,在临床工作中可因医师的主观能动性导致勾画偏差。自动勾画软件随着计算机算法的介入不断更新,虽然减少了临床医生的工作量,节省了病人等待治疗的时间、提高了靶区及 OARs 勾画的一致性和规范性,但自动化软件在 OARs 的勾画精度仍需探究。

目前国际上已广泛报道基于深度学习的自动勾画研究<sup>[14-15]</sup>,证明 OARs 的自动勾画结合放疗医师的仔细检查,能在不影响治疗质量的前提下,基于算法的轮廓可以安全应用与临床实践。在 Vaassen 等<sup>[16]</sup>的研究中,经过评估 235 例患者勾画实践对比,使用自动勾画轮廓的总平均时间约能节省 12% (食道)至 77% (右肺),并且其 DSC 能达到 0.46 (食道)至 0.98 (右肺)。而在 Zabel 等<sup>[17]</sup>的研究中,前列腺患者的自动勾画中可以节省约 8.5 min,从而可以实现资源的重新分配,将更多的时间用于患者的治疗及计划的监测中去。

本研究对三种自动勾画软件在食管癌 OARs 勾画的实际应用效果进行探究,通过比较 HD、MDA、DSC 及 JC 四个参数评价三种自动勾画软件的准确性,其中 DSC 是一项评价自动勾画与手动勾画 OARs 重合性的重要指标,从计算得出的 DSC 参数来看,三个自动软件对于结构比较大、形状轮廓比较规则的器官,如心脏、肝脏、肺及脊髓,DSC 均  $> 0.8$ 。而对于体积较小的器官,如甲状腺勾画效果不太理想( $P < 0.001$ )。考虑手动勾画与自动勾画间的微小偏差可引起 DSC 较大波动,因此,进一步对比三者 JC、HD 与 MDA 值。经比较三种自动勾画软件 JC 值在心脏、肝脏、肺及脊髓等无统计学差异( $P > 0.05$ )。对于心脏勾画研究显示,MIM 的 4 项参考指标结果显示均明显优于其它两者,结合心脏图层分析,PVmed 与 UI 勾画上界较 MIM 更高,同时更容易出现动脉血管干扰,导致勾画轮廓适形度下降,而 MIM 软件与人工勾画从右肺动脉下壁开始勾画。肝脏勾画结果显示, PVmed 与 UI 均要优于 MIM,主要原因可能是肝脏与肺的分界不清导致 MIM 在自动勾画上的下降。根据检查结果显示右肺的 JC 值

最佳,均在 0.914 ~ 0.927 之间,进一步分析双肺 HD 可见 PVmed 要明显优于 UI 及 MIM,同时 PVmed 与 UI 具有统计学差异 ( $P < 0.001$ ),分析其原因可能是 PVmed 在识别左右肺门及支气管要优于 UI;而在甲状腺的自动勾画中 HD 与 MDA 值:PVmed < MIM < UI。三种自动勾画软件 JC 值:PVmed > UI > MIM,进一步分析 MIM 软件自动勾画甲状腺的效果较差的原因可能在于 MIM 以 Atlas 数据库建立模板,病例数据为 30 例,通过与图谱进行相应匹配后自动勾画轮廓,由于图谱库数据限制、小体积器官解剖的多样性导致 MIM 在甲状腺的自动勾画效率差于其他两种自动勾画软件。PVmed、United Imaging 勾画则是基于神经网络深度学习,大数据更多,通过对数据集的训练实现快速准确的勾画<sup>[18]</sup>。

在实际应用中,虽然自动勾画软件的 OARs 轮廓并不能直接应用于放疗计划的设计,仍需临床医生仔细审核、确定。但自动勾画软件可显著提高医师胸部肿瘤 OARs 勾画效率,有相关研究<sup>[19]</sup>证实可缩短约 60% 的耗时。自动勾画技术的快速发展在很大程度上提高了 OARs 的勾画效率,减轻了放疗医生和物理师的工作量<sup>[20]</sup>。Kawamura 等<sup>[21]</sup>研究说明,使用缩短分割实践是自动勾画系统的主要优势,但其精确性仍是当前阶段的一大挑战。

综上,直接将 PVmed、UI 及 MIM 的胸部危及器官自动勾画结果应用于计划设计中时,总体安全可靠,但仍需关注部分小体积器官,如甲状腺,在计划设计前需对这些器官进行人工修改及验证,以避免发生剂量偏差现象。

## 参考文献

- [1] 吴哲,庞亚,明智,等. 人工智能技术在鼻咽癌放疗危及器官自动勾画中的应用研究[J]. 实用肿瘤学杂志,2021,35(2):137-141.
- [2] Kosmin M, Ledsam J, Romera-Paredes B, et al. Rapid advances in auto-segmentation of organs at risk and target volumes in head and neck cancer[J]. Radiotherapy and Oncology, 2019, 135:130-140.
- [3] Velker VM, Rodrigues GB, Dinniwell R, et al. Creation of RTOG compliant patient CT-atlases for automated atlas based contouring of local regional breast and high-risk prostate cancers[J]. Radiation Oncology (London, England), 2013, 8:188.
- [4] 谷珊珊,田娟秀,王运来,等. 基于 MIM 软件模版数据库的病例数对危及器官自动勾画的探讨[J]. 中国医学装备, 2018, 15(3):1-4.
- [5] 张玉海,李月敏. 人工智能在肿瘤放射治疗中的研究进展[J]. 实用肿瘤学杂志, 2019, 33(6):570-574.
- [6] Yang J, Veeraraghavan H, van Elmpt W, et al. CT images with expert manual contours of thoracic cancer for benchmarking auto-segmentation accuracy[J]. Medical Physics, 2020, 47(7):3250-3255.

- [7] MacKay K, Bernstein D, Glocker B, et al. A review of the metrics used to assess auto-contouring systems in radiotherapy[J]. Clinical Oncology, 2023, 35(6):354-369.
- [8] Jameson MG, Holloway LC, Vial PJ, et al. A review of methods of analysis in contouring studies for radiation oncology[J]. Journal of Medical Imaging and Radiation Oncology, 2010, 54(5):401-410.
- [9] Mohammadi R, Shokatian I, Salehi M, et al. Deep learning-based auto-segmentation of organs at risk in high-dose rate brachytherapy of cervical cancer[J]. Radiotherapy and Oncology, 2021, 159:231-240.
- [10] Fung NTC, Hung WM, Sze CK, et al. Automatic segmentation for adaptive planning in nasopharyngeal carcinoma IMRT: time, geometrical, and dosimetric analysis[J]. Medical Dosimetry, 2020, 45(1):60-65.
- [11] Ju Z, Wu Q, Yang W, et al. Automatic segmentation of pelvic organs-at-risk using a fusion network model based on limited training samples[J]. Acta Oncologica (Stockholm, Sweden), 2020, 59(8):933-939.
- [12] Tham IWK, Lu JJ. Treatment of early stage nasopharyngeal carcinoma: conventional versus new radiation therapy technologies[J]. Journal of Radiation Oncology, 2012, 1(2):99-106.
- [13] 孔维佳,文卫平,王德辉,等. 鼻咽癌手术治疗探讨[J]. 临床耳鼻咽喉头颈外科杂志, 2023, 37(7):507-518.
- [14] Hasibul Hoque SM, Pirrone G, Matrone F, et al. Clinical use of a commercial artificial intelligence-based software for autocontouring in radiation therapy: geometric performance and dosimetric impact[J]. Cancers, 2023, 15(24):5735.
- [15] Ginn JS, Gay HA, Hilliard J, et al. A clinical and time savings evaluation of a deep learning automatic contouring algorithm[J]. Medical Dosimetry, 2023, 48(1):55-60.
- [16] Vaassen F, Hazelaar C, Vaniqui A, et al. Evaluation of measures for assessing time-saving of automatic organ-at-risk segmentation in radiotherapy[J]. Physics and Imaging in Radiation Oncology, 2019, 13:1-6.
- [17] Zabel WJ, Conway JL, Gladwish A, et al. Clinical evaluation of deep learning and atlas-based auto-contouring of bladder and rectum for prostate radiation therapy[J]. Practical Radiation Oncology, 2021, 11(1):e80-e89.
- [18] 胡兴刚,王嫻,张扬,等. 基于深度学习算法的自动勾画系统在头颈部危及器官勾画精度的研究[J]. 中国医学物理学杂志, 2024, 41(5):548-553.
- [19] 高文超,王军良,郑倩倩,等. 不同图像分割软件用于胸部危及器官自动勾画的临床应用[J]. 中国医疗设备, 2023, 38(7):50-54, 85.
- [20] Rong Y, Chen Q, Fu Y, et al. NRG oncology assessment of artificial intelligence deep learning-based auto-segmentation for radiation therapy: current developments, clinical considerations, and future directions[J]. International Journal of Radiation Oncology, Biology, Physics, 2024, 119(1):261-280.
- [21] Kawamura M, Kamomae T, Yanagawa M, et al. Revolutionizing radiation therapy: the role of AI in clinical practice[J]. Journal of Radiation Research, 2024, 65(1):1-9.

(收稿日期:2024-07-04

修回日期:2024-07-20)