

DOI:10.3969/j.issn.1005-202X.2024.03.001

医学放射物理

不同级别肿瘤中心医师对鼻咽癌调强放疗靶区和危及器官勾画差异比较

陈美宁^{1,2}, 刘懿梅², 彭应林², 谢秋英³, 石锦平³, 黄荣³, 赵充², 邓小武², 周美娟¹

1. 南方医科大学公共卫生学院放射医学系/国家药监局化妆品安全评价重点实验室/广东省热带病研究重点实验室, 广东 广州 510515; 2. 中山大学肿瘤防治中心/华南肿瘤学国家重点实验室/广东省鼻咽癌诊治研究重点实验室, 广东 广州 510060; 3. 佛山市第一人民医院放疗科, 广东 佛山 528010

【摘要】目的:评估不同级别肿瘤中心医师对鼻咽癌调强放疗计划靶区和危及器官勾画的个体化差异,为多中心放疗临床试验的质量控制提供数据参考。**方法:**随机抽取12例不同TMN分期的鼻咽癌患者,3名同一市级肿瘤中心医师以手工方式勾画每例患者的靶区(GTVnx)和危及器官(OAR)。以国家区域级肿瘤中心放疗专家手工修改和确认的靶区(GTVnx)和OAR勾画结构作为标准勾画。采用绝对体积差异比(ΔV_{diff})、体积最大/最小比(MMR)、离散系数(CV)和体积相似性指数(DSC)分别比较不同级别肿瘤中心(市级与国家区域级)医师之间和市级肿瘤中心3名医师之间的器官勾画差异,并进一步比较不同级别肿瘤中心医师对不同TMN分期的鼻咽癌放疗计划器官的勾画差异。**结果:**不同级别肿瘤中心(市级与国家区域级)医师之间勾画GTVnx体积差异明显,3名医生中的最大 ΔV_{diff} 、MMR和CV分别为 $97.23\% \pm 83.45\%$ 、 2.19 ± 0.75 和 0.31 ± 0.14 ,平均DSC <0.7 。同时,他们之间对于左右视神经、视交叉、脑垂体等小体积OAR勾画差异也较大,平均MMR >2.8 ,CV >0.37 ,DSC <0.51 ;但对于脑干、脊髓、左右眼球、左右下颌骨等大体积OAR的勾画差异相对较小,平均的 $\Delta V_{diff} < 42\%$,MMR <1.55 ,DSC >0.7 。相比于不同级别肿瘤中心医师之间的勾画差异,市级肿瘤中心3名医生勾画差异有一定的减小。另外,不同级别肿瘤中心医师之间对不同分期的鼻咽癌靶区勾画也存在差异,相对早期患者(I或II期)的靶区勾画,他们之间在对晚期患者(III或IV期)靶区勾画的差异更小,平均 ΔV_{diff} 和DSC值分别为($98.31\% \pm 67.36\%$ vs $69.38\% \pm 72.61\%$, $P < 0.05$)和(0.55 ± 0.08 vs 0.72 ± 0.12 , $P < 0.05$)。**结论:**不同级别肿瘤中心医师勾画鼻咽癌放疗计划靶区和OAR存在差异,尤其对早期患者靶区(GTVnx)和小体积OAR的勾画。为确保多中心临床试验的准确性,建议对不同级别医师进行统一培训并审核其勾画结果,以减少差异对治疗结果的影响。

【关键词】鼻咽癌;调强放射治疗;器官勾画;绝对体积差异;相似性系数;离散系数

【中图分类号】R739.62

【文献标志码】A

【文章编号】1005-202X(2024)03-0265-08

Comparison of interobserver variations in delineation of target volumes and organs-at-risk for intensity-modulated radiotherapy of nasopharyngeal carcinoma among physicians from different levels of cancer centers

CHEN Meining^{1,2}, LIU Yimei², PENG Yinglin², XIE Qiuying³, SHI Jinping³, HUANG Rong³, ZHAO Chong², DENG Xiaowu², ZHOU Meijuan¹

1. Department of Radiology, School of Public Health, Southern Medical University/Key Laboratory of Cosmetics Safety Evaluation, National Medical Products Administration/Key Laboratory of Tropical Diseases Research, Guangzhou 510515, China; 2. Sun Yat-sen University Cancer Center/State Key Laboratory of Oncology in South China/Key Laboratory of Diagnosis and Treatment of Nasopharyngeal Cancer in Guangdong Province, Guangzhou 510060, China; 3. Department of Radiation Oncology, the First People's Hospital of Foshan, Foshan 528010, China

Abstract: Objective To assess inter-observer variations (IOV) in the delineation of target volumes and organs-at-risk (OAR) for intensity-modulated radiotherapy (IMRT) of nasopharyngeal carcinoma (NPC) among physicians from different levels

【收稿日期】2023-10-21

【基金项目】国家重点研发计划(2022YFC2402304);国家自然科学基金(12005316);广州市科技计划项目(202206010154, 202206010180)

【作者简介】陈美宁,初级技师,研究方向:肿瘤放射治疗物理学,E-mail: chenmn1@sysucc.org.cn

【通信作者】周美娟,博士,教授,研究方向:放射生物学,E-mail: 80758598@qq.com

of cancer centers, thereby providing a reference for quality control in multi-center clinical trials. **Methods** Twelve patients with NPC of different TMN stages were randomly selected. Three physicians from the same municipal cancer center manually delineated the target volume (GTVnx) and OAR for each patient. The manually modified and confirmed target volume (GTVnx) and OAR delineation structures by radiotherapy experts from the regional cancer center were used as the standard delineation. The absolute volume difference ratio (ΔV_{diff}), maximum/minimum volume ratio (MMR), coefficient of variation (CV), and Dice similarity coefficient (DSC) were used to compare the differences in organ delineation among physicians from different levels of cancer centers and among the 3 physicians from the same municipal cancer center. Furthermore, the IOV of GTVnx and OAR among physicians from different levels cancer centers were compared across different TMN stages. **Results** Significant differences in the delineation of GTVnx were observed among physicians from different levels of cancer centers. Among the 3 physicians, the maximum values of ΔV_{diff} , MMR, and CV were $97.23\% \pm 83.45\%$, 2.19 ± 0.75 , and 0.31 ± 0.14 , respectively, with an average DSC of less than 0.7. Additionally, there were considerable differences in the delineation of small-volume OAR such as the left and right optic nerves, chiasm, and pituitary, with average $\text{MMR} > 2.8$, $\text{CV} > 0.37$, and $\text{DSC} < 0.51$. However, relatively smaller differences were observed in the delineation of large-volume OAR such as the brainstem, spinal cord, left and right eyeballs, and left and right mandible, with average $\Delta V_{\text{diff}} < 42\%$, $\text{MMR} < 1.55$, and $\text{DSC} > 0.7$. Compared with the differences among physicians from different levels cancer centers, the differences among the 3 physicians from the municipal cancer center were slightly reduced. Furthermore, there were also differences in the delineation of target volumes for NPC among physicians from different levels cancer centers, depending on the staging of the disease. Compared with the delineation of target volumes for earlier stage patients (stages I or II), the differences among physicians in the delineation of target volumes for advanced stage patients (stages III or IV) were smaller, with average ΔV_{diff} and DSC of $98.31\% \pm 67.36\%$ vs $69.38\% \pm 72.61\%$ ($P < 0.05$) and 0.55 ± 0.08 vs 0.72 ± 0.12 ($P < 0.05$), respectively. **Conclusion** There are differences in the delineation of GTVnx and OAR in radiation therapy for NPC among physicians from different levels of cancer centers, especially in the delineation of target volume (GTVnx) and small-volume OAR for early-stage patients. To ensure the accuracy of multicenter clinical trials, it is recommended to provide unified training to physicians from different levels of cancer centers and review their delineation results to reduce the effect of differences on treatment outcomes.

Keywords: nasopharyngeal carcinoma; intensity modulated radiotherapy; organ segmentation; absolute volume difference ratio; Dice similarity coefficient; coefficient of variation

前言

鼻咽癌(Nasopharyngeal Carcinoma, NPC)发病分布具有明显的地域性,有研究显示40%的NPC发生在中国,放射治疗是NPC最主要的治疗手段^[1]。调强放射治疗(Intensity-Modulated Radiotherapy, IMRT)能在提高肿瘤照射剂量的同时保护好危及器官(Organs-at-Risk, OAR)^[2-4],但其必要条件是实现靶区和OAR体积轮廓的精准勾画^[4-6]。然而,由于NPC的解剖学结构特点,肿瘤的生长、浸润、转移等特征,其靶区形状复杂、相邻OAR众多,不同医生对同一器官的勾画存在较大的差异,这些差异不仅会影响下游治疗计划的优化和计划的质量,并且会影响疗效的评估和分析的结果^[6-8]。近年来系统的NPC靶区及OAR勾画指南的制定以及智能勾画软件的发展和临床应用在一定程度上减少了医师间OAR的勾画差异^[9-12],但患者解剖结构勾画的准确性高度依赖于医师的影像学水平和临床训练的经验,不同级别肿瘤中心医师对指南的主观理解存在较大分歧,且不同肿瘤中心训练自动勾画的模板也参差不齐。因此,

本研究通过对比分析不同级别肿瘤中心医师在勾画NPC放疗计划靶区和OAR体积结构方面的差异,评估医师间的勾画差异程度(Interobserver Variation, IOV)。这一研究为开展多中心临床试验的放疗器官勾画提供了质量控制的重要参考依据。

1 资料与方法

1.1 研究对象

随机选取2017年5月至2018年10月在中山大学肿瘤防治中心放射治疗科行IMRT的NPC患者12例,其中男10例,女2例;年龄35~82岁,中位年龄50岁。入组患者I、II、III和IV分期分别为1、5、4和2例。根据第7版国际抗癌联盟/美国癌症联合委员会(UICC/AJCC)分期手册进行TMN分期,一般临床资料见表1。

1.2 模拟定位图像的获取

每例患者均取仰卧位,以真空袋和头颈部面罩固定后行螺旋CT(Somatom Sensation Open, Siemens AG)平扫和增强扫描。扫描条件为电压140 kV,电流280 mAs,扫描和重建层的层厚均为3 mm,扫描范

表1 患者一般资料
Table 1 General information of patients

编号	性别	年龄/岁	T分期	N分期	M分期	总分期
1	男	35	T1	N0	M0	I
2	女	35	T1	N1	M0	II
3	男	49	T2	N0	M0	II
4	男	82	T2	N0	M0	II
5	男	60	T2	N1	M0	II
6	男	40	T2	N1	M0	II
7	男	59	T2	N2	M0	III
8	女	36	T2	N2	M0	III
9	男	55	T3	N1	M0	III
10	男	59	T3	N2	M0	III
11	男	50	T4	N2	M0	IVa
12	男	59	T4	N3	M0	IVa

围从头顶到器官隆突水平。获得的患者定位CT图像(平扫+增强)分别传输至本地。

1.3 图像勾画

1.3.1 市级肿瘤中心医师勾画 将每例患者的平扫CT图像、增强CT图像和MR图像一并拷贝复制成4份,将其中3份分别发送给同一家市级医院的3名放射治疗医师,由他们参考增强CT和MR图像,在平扫CT图像上进行手工勾画。市级肿瘤中心的3名医师在全盲状态下参考指南^[4]对每例患者的靶区和脊髓、脑干、晶体、视神经、腮腺、颞叶等OAR进行手动勾画,一共得到36套独立的结构轮廓,3组市级肿瘤中心医师勾画的器官体积分别命名为 V_{M1} 、 V_{M2} 和 V_{M3} 。所有勾画都在Eclipse放疗计划系统(Version 13.5, Varian, USA)上完成。

1.3.2 国家区域级肿瘤中心专家的标准勾画 采用ABAS自动勾画软件(Version2.01, Elekta AB, Stockholm, Sweden)对12例患者的CT图像进行OAR的预勾画,预勾画后导入Monaco治疗计划系统(Version 5.10, ElektaAB),由一名主治医师对每例患者的靶区体积和OAR进行手工勾画和修改。然后,由3名专家共同对每例患者的靶区和OAR进行讨论、修改、审核和确认,最终获得12例患者的靶区与OAR勾画的体积结构,并作为本研究的标准勾画体积(V_s)。

1.4 勾画差异大小的评价指标

(1)绝对体积差异(ΔV_diff),反映不同医师勾画体积与标准勾画体积的绝对体积差别大小:

$$\Delta V_diff = \frac{|V_M - V_s|}{V_s} \times 100\%$$

(1)

其中, V_M 为市级肿瘤中心医师勾画器官的个体化体积, V_s 为其标准勾画体积。

(2)体积相似性指数(Dice Similarity Coefficient, DSC),反映不同医师独立勾画器官的重合性,DSC值越接近1,表示两种勾画的体积重合性越好^[13-14]。在图像匹配评价中建议配准DSC>0.7为重合度较好的标准^[15]。DSC公式为:

$$DSC = \frac{2(V_1 \cap V_2)}{|V_1| + |V_2|}$$

(2)

其中, V_1 和 V_2 分别为两名不同医师勾画的OAR体积, $(V_1 \cap V_2)$ 为 V_1 和 V_2 的交集体积。

(3)体 积 最 大/最 小 比 (Maximum/Minimum Ratio, MMR):反映不同医师独立勾画器官的最大体积和最小体积之间的差异程度^[6],MMR公式为:

$$MMR = \frac{V_{max}}{V_{min}}$$

(3)

其中, V_{max} 和 V_{min} 分别为所评价器官的体积最大值和最小值。MMR反映不同勾画者之间的最大差异,数值上越接近1越好。

(4)离散系数(Coefficient of Variation, CV):反映不同医师对各个器官体积勾画的差异(离散)程度^[6,16],CV越大,代表OAR勾画的差异性越大。CV公式为:

$$CV = \frac{V_{std}}{V_{ave}}$$

(4)

其中, V_{std} 和 V_{ave} 分别为所评价器官体积的标准差和平均值。

1.5 勾画差异大小分析

(1)不同级别肿瘤中心医师之间靶区和OAR的勾画差异大小:分别比较3名市级肿瘤中心的医师所勾画器官体积(V_M)与国家区域级肿瘤中心专家勾画的器官标准体积(V_s)的差异,以 V_s 为标准,分别计算不同级别肿瘤中心医师勾画器官的 ΔV_diff 和DSC。另外,还比较了不同级别肿瘤中心医师勾画之间的MMR和CV。

(2)市级肿瘤中心医师之间靶区和OAR的勾画差异大小:采用MMR、CV和两两医师勾画各器官体积的DSC来比较3名市级肿瘤中心医师之间勾画各器官体积(V_M)的差异大小。

(3)不同级别肿瘤中心医师之间对不同分期患者靶区和OAR的勾画差异大小:根据患者的TMN分期,将总分期为I或II期的患者定为早期患者,将总分期III或IV期的患者定为晚期患者。以DSC和 ΔV_diff 指标比较不同分期下不同级别肿瘤中心医师之间勾画IOV大小。

1.6 统计学方法

使用 SPSS 22.0 软件 (IBM SPSS Inc., Chicago, USA)对4组器官勾画体积先行正态性检验,符合正态分布的使用配对 *t* 检验,不符合正态分布的使用 Wilcoxon 秩检验;对3组不同级别肿瘤中心专家勾画器官的 DSC 行正态性检验,符合正态分布的采用方差分析法,否则使用 Firedman 检验;对两组不同分期患者的 DSC 和 ΔV_diff 做多因素的方差分析,以分期为主效应; $P<0.05$ 为差异有统计学意义。

2 结果

2.1 不同级别肿瘤中心医师之间勾画靶区和 OAR 的体积大小差异比较

不同级别肿瘤中心医师之间勾画靶区 GTVnx 的体积大小差异较大,3名市级肿瘤中心医师勾画的靶区体积与标准 GTVnx 体积之间的差异均有统计学意义 ($P<0.05$),平均 MMR 和 CV 分别 2.19 ± 0.75 和 0.31 ± 0.14 ,最大者与标准 GTVnx 体积平均相差接近 1 倍,最大 ΔV_diff 为 $97.23\%\pm83.45\%$ 。同时,不同级别肿瘤中心医师之间在勾画左右视神经、视交叉、脑垂体等小体积且边界不清晰的 OAR 时,表现出了更大的体积大小差异,平均 MMR >2.8 ,平均 CV >0.37 。而对于脑干、脊髓、左右眼球、左右下颌骨等大体积 OAR 的勾画,其体积大小差异相对较小,平均 $\Delta V_diff<42\%$,平均 MMR <1.55 ,其中脑干的体积差异最小,最小 ΔV_diff 仅为 $6.71\%\pm6.44\%$ 。详见表 2。

表2 不同级别肿瘤中心医师之间勾画靶区和 OAR 的体积差异比较
Table 2 Comparison of volume variation in GTVnx and OAR delineation among physicians from different levels of cancer centers

器官	标准勾画/cm ³	$\Delta V_diff/\%$			MMR	CV
		市级医生 1	市级医生 2	市级医生 3		
靶区	57.20±53.70	59.74±45.59*	94.58±76.46*	97.23±83.45*	2.19±0.75	0.31±0.14
左侧颞叶	88.75±5.93	36.28±13.55*	19.86±9.29*	39.78±12.89*	1.76±0.33	0.26±0.09
右侧颞叶	94.93±8.16	41.12±15.73*	28.13±12.25*	44.56±14.49*	1.94±0.45	0.31±0.12
左侧下颌骨	36.05±6.19	18.1±12.69*	35.88±16.24*	40.82±20.08*	1.43±0.19	0.16±0.05
右侧下颌骨	36.46±6.99	15.81±8.97	35.67±13.04*	41.38±18.16*	1.54±0.26	0.19±0.07
脑干	25.77±4.49	6.71±6.44	9.77±6.7*	10.55±8.77	1.19±0.12	0.07±0.04
左侧腮腺	24.9±6.73	31.89±15.27*	10.08±8.16*	28.53±18.29*	1.65±0.33	0.22±0.09
右侧腮腺	24.3±6.8	34.62±18.14*	13.33±7.43	33.16±20.89*	1.80±0.43	0.26±0.12
脊髓	14.73±5.51	28.00±17.12	19.88±13.60	21.54±12.68	1.46±0.15	0.17±0.05
左侧眼球	6.88±1.11	35.63±20.68*	34.24±16.71*	36.44±25.31*	1.43±0.25	0.14±0.06
右侧眼球	6.68±0.97	40.72±20.87*	36.18±20.79*	32.75±31.26*	1.47±0.28	0.16±0.06
左颞颌关节	1.68±0.74	92.09±82.26*	47.15±52.15	74.97±86.20	2.08±0.92	0.27±0.16
右颞颌关节	1.58±0.79	147.77±97.60*	71.69±58.66	114.88±105.04*	2.62±0.82	0.39±0.12
视交叉	0.56±0.14	44.38±34.54	41.43±27.6	39.24±27.69*	2.88±2.02	0.37±0.21
左侧视神经	0.19±0.1	164.17±163.22*	191.67±178.73*	113.89±112.33*	3.26±1.52	0.42±0.20
右侧视神经	0.21±0.11	105.3±117.04*	176.52±155.94*	131.06±123.05*	3.15±1.24	0.45±0.16
脑垂体	0.14±0.07	47.22±86.7	147.22±147.51*	38.89±44.00	3.18±1.54	0.51±0.26
左侧晶体	0.11±0.05	58.33±46.87*	58.33±66.86	54.14±49.81	1.90±0.57	0.39±0.23
右侧晶体	0.12±0.04	54.17±49.81*	66.67±49.24	58.33±51.49	1.92±0.29	0.36±0.11

与标准勾画相比,* $P<0.05$

2.2 不同级别肿瘤中心医师之间勾画靶区和 OAR 的 DSC 比较

不同级别肿瘤中心医师之间勾画靶区不仅在体积大小上有较大差异,其在形状结构上也差异较大,

3名市级肿瘤中心医师所勾画的 GTVnx 平均 DSC 值均 <0.7 。同时,对于视神经、视交叉、脑垂体等小体积 OAR 勾画比较,其 DSC 也较小,平均 DSC <0.51 ,其中视交叉的 DSC 最小,3名医师中最小 DSC 值仅为

0.12±0.15。而对于脑干、脊髓、左右眼球、左右下颌骨等大体积器官勾画,不同级别肿瘤中心医师之间勾画的形状结构具有较好的一致性,平均DSC>0.7。另外,3名市级肿瘤中心医师在左右颞叶、左右腮腺、视交叉和左下颌骨等器官勾画的DSC具有统计学意义($P<0.05$),详见图1和表3。

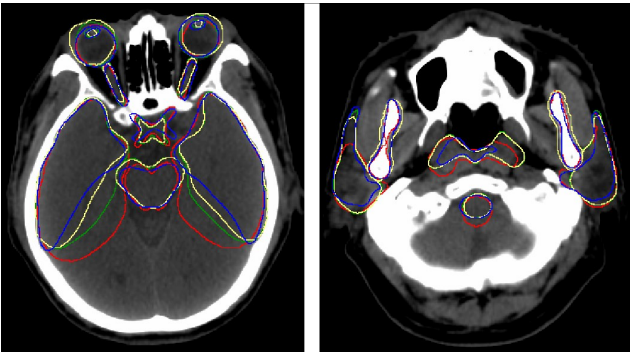


图1 3名市级肿瘤中心医师与国家区域级肿瘤中心标准结构的勾画差异比较图

Figure 1 Comparison of segmentation differences between 3 physicians from the same municipal cancer center and the standard structure of national regional cancer center

红色为标准器官结构轮廓;蓝色、绿色、黄色分别为市级肿瘤中心3名医生勾画的器官结构轮廓

2.3 3名市级肿瘤中心医师之间勾画靶区和OAR体积的差异性比较

相比于不同级别肿瘤中心医师之间的勾画差异,市级肿瘤中心3名医生勾画差异有一定的减小。其中脑干、左右眼球、左右下颌骨、左右腮腺、脊髓、左右颞叶以及左右颞颌关节均有较好的体积结构相似性($DSC>0.7$),且3名医师之间的平均CV和MMR也相对变小,分别小于0.26和1.71。但是,对边界不清晰的小体积器官如左右视神经、视交叉、脑垂体等器官的勾画,3名市级医生仍存在较大的差异,平均CV和MMR分别大于0.3和2.1。详见表4。

2.4 不同级别肿瘤中心医师之间对不同分期患者靶区和OAR的勾画差异大小

不同级别肿瘤中心医师之间对不同分期患者靶区勾画存在差异,但在体积大小上没有统计学意义($P>0.05$)(图2a)。而对于靶区的 ΔV_{diff} 和DSC值之间比较显示,相对早期患者(I或II期)的靶区勾画,他们之间在对晚期患者(III或IV期)靶区勾画的差异更小,平均 ΔV_{diff} 和DSC值分别为98.31%±67.36% vs 69.38%±72.61% ($P<0.05$) 和 0.55±0.08 vs 0.72±0.12 ($P<0.05$)(图2b)。而对于早期和晚期患者的OAR勾画比较,发现大多数器官之间没有差别($P>0.05$),

表3 不同级别肿瘤中心医师之间勾画靶区和OAR的DSC比较

Table 3 Comparison of DSC in GTVnx and OAR delineation among physicians from different levels of cancer centers

器官	市级医生1	市级医生2	市级医生3	P值
靶区	0.65±0.12	0.64±0.14	0.62±0.14	0.854
左侧颞叶	0.73±0.08	0.77±0.05	0.69±0.06	0.012
右侧颞叶	0.69±0.09	0.74±0.05	0.66±0.09	0.031
左侧下颌骨	0.86±0.06	0.83±0.05	0.80±0.06	0.041
右侧下颌骨	0.80±0.11	0.82±0.04	0.81±0.06	0.407
脑干	0.84±0.03	0.85±0.03	0.85±0.03	0.368
左侧腮腺	0.69±0.07	0.83±0.05	0.74±0.1	0.001
右侧腮腺	0.66±0.07	0.83±0.04	0.71±0.11	0.015
脊髓	0.74±0.05	0.73±0.05	0.74±0.05	0.502
左侧眼球	0.83±0.06	0.83±0.06	0.82±0.07	0.662
右侧眼球	0.82±0.06	0.83±0.06	0.81±0.09	0.344
左颞颌关节	0.69±0.09	0.74±0.05	0.66±0.09	0.809
右颞颌关节	0.48±0.15	0.56±0.17	0.51±0.17	0.515
视交叉	0.31±0.25	0.17±0.19	0.12±0.15	0.039
左侧视神经	0.47±0.21	0.46±0.19	0.51±0.18	0.788
右侧视神经	0.48±0.22	0.47±0.21	0.48±0.22	0.995
脑垂体	0.32±0.31	0.48±0.17	0.33±0.28	0.185
左侧晶体	0.58±0.16	0.65±0.16	0.62±0.12	0.453
右侧晶体	0.62±0.21	0.68±0.11	0.70±0.11	0.556

仅脊髓的 ΔV_{diff} 以及颞叶等几个器官的DSC具有一定的统计学意义($P<0.05$)(图2c和d)。

3 讨论

不同医师之间勾画NPC靶区和OAR体积存在差异,且体积差异直接影响放疗计划的设计和优化,并可能影响疗效和临床评价^[17-18]。尽管已经发布了NPC放疗计划器官勾画指南,但在实际的临床操作中,由于NPC靶区及其周边解剖结构的复杂性,以及不同临床医生在影像学知识和临床经验上的个体差异,使得在勾画肿瘤靶区GTVnx时,经验不足的医师经常面临界定肿瘤浸润范围的挑战。这可能导致靶区体积的勾画过大或过小,从而影响放疗的精确性和治疗效果^[19-21]。本研究的结果进一步揭示了不同级别肿瘤中心的医师在勾画GTVnx体积时存在显著差异,平均DSC小于0.7,平均体积差异(ΔV_{diff})最大超过90%,且平均MMR高达2.19,部分患者的GTVnx体积勾画差异甚至达到3倍以上。这与Peng等^[20]在国内进行的多中心研究结果相似,他们发现不同肿瘤中心医师勾画同一例NPC计划靶区体积(PTVnx)时,平均MMR为2.41。

表 4 市级放疗中心不同医生之间勾画差异比较

Table 4 Comparison of segmentation differences among different physicians in municipal cancer center

器官	CV	MMR	DSC		
			市级医生 1 vs 市级医生 2	市级医生 1 vs 市级医生 3	市级医生 2 vs 市级医生 3
靶区	0.20±0.16	1.57±0.57	0.67±0.16	0.68±0.17	0.82±0.11*
左侧颞叶	0.19±0.06	1.40±0.13	0.78±0.04*	0.89±0.09	0.82±0.07*
右侧颞叶	0.17±0.07	1.38±0.15	0.78±0.05*	0.89±0.08	0.82±0.09*
左侧下颌骨	0.11±0.05	1.23±0.12	0.86±0.06*	0.84±0.06*	0.95±0.03
右侧下颌骨	0.17±0.10	1.43±0.33	0.82±0.09*	0.81±0.10*	0.95±0.03
脑干	0.05±0.02	1.11±0.05	0.88±0.03*	0.91±0.06	0.92±0.04
左侧腮腺	0.22±0.08	1.52±0.23	0.73±0.07*	0.83±0.13	0.82±0.11*
右侧腮腺	0.25±0.14	1.61±0.38	0.71±0.09*	0.84±0.13	0.81±0.14*
脊髓	0.14±0.06	1.32±0.16	0.77±0.09*	0.79±0.13*	0.86±0.09
左侧眼球	0.05±0.03	1.10±0.07	0.92±0.02*	0.89±0.04	0.89±0.04
右侧眼球	0.08±0.06	1.17±0.14	0.92±0.02	0.86±0.05	0.86±0.06
左颞颌关节	0.19±0.17	1.53±0.61	0.77±0.11*	0.85±0.19*	0.84±0.14
右颞颌关节	0.26±0.17	1.71±0.51	0.71±0.17*	0.83±0.21*	0.85±0.13
视交叉	0.37±0.32	2.53±2.13	0.26±0.23*	0.41±0.39	0.54±0.34
左侧视神经	0.33±0.19	2.12±1.13	0.54±0.19*	0.63±0.25	0.73±0.10*
右侧视神经	0.32±0.19	2.31±1.74	0.58±0.22*	0.68±0.22	0.69±0.22*
脑垂体	0.45±0.34	2.56±1.42	0.52±0.18*	0.56±0.27	0.51±0.22*
左侧晶体	0.23±0.21	1.66±0.55	0.68±0.11	0.67±0.14	0.70±0.12
右侧晶体	0.35±0.14	1.89±0.33	0.66±0.14	0.64±0.13*	0.65±0.101

*表示市级肿瘤中心 3 名医师中两两医师勾画的器官体积之间的差异具有统计学意义,即采用配对 *t* 检验或 Wilcoxon 秩检验, *P*<0.05

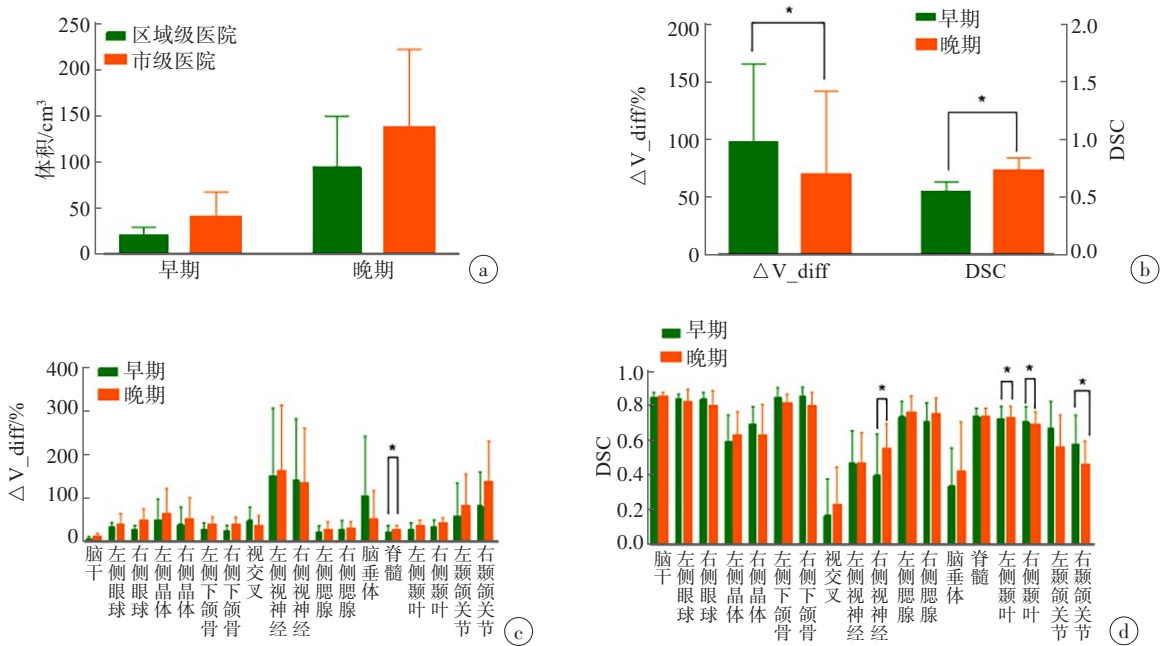


图 2 不同级别肿瘤中心医师之间对不同分期患者靶区和 OAR 的勾画差异大小对比图

Figure 2 Comparison of segmentation differences of GTVnx and OAR in patients of different stages among physicians at different cancer centers

图 2a 为 GTVnx 的体积大小差异对比图;图 2b 为 GTVnx 的 ΔV_{diff} 和 DSC 对比图;图 2c 和 d 为 OAR 的 ΔV_{diff} 和 DSC 对比图;*表示两者做多因素的方差分析后具有统计学意义,即 *P*<0.05

NPC周围涉及的关键器官和OAR数量众多,其解剖结构极为复杂。在日常的临床勾画实践中,不同经验的医师对于OAR勾画指南的主观理解可能存在差异,且他们在基于影像识别正常器官的能力上也有所不同。这些因素共同导致在勾画鼻咽癌的OAR时,不同医师之间仍存在显著的差异^[22-23]。本研究对不同级别肿瘤中心医师勾画NPC靶区与OAR的经验水平进行对比研究,结果显示在左右视神经、视交叉、脑垂体等边界不清晰的小体积OAR的勾画中,不同级别肿瘤中心医师的勾画存在明显的差异,平均MMR>2.8,平均CV>0.37且平均DSC<0.51。而对于脑干、脊髓、左右眼球、左右下颌骨等体积较大,边界相对清晰的器官,不同级别肿瘤中心医师之间具有较好的一致性,平均 ΔV_{diff} <42%,平均MMR<1.55,且平均DSC>0.7。这些发现进一步强调了在NPC放疗中,针对不同体积和边界清晰度的OAR,医师需要具备不同的勾画技巧和知识,以确保放疗的精确性和治疗效果。钱建军等^[22]分析了同一医院中低年资医师与高年资医师对16例NPC的IMRT计划OAR的勾画差异,结果显示脑干、颞叶、脊髓、腮腺、晶体、喉、视神经和视交叉的DSC值依次降低,这反映了器官体积越小,其勾画DSC越小,这与本研究的发现相似。Peng等^[20]指出小体积OAR勾画的DSC相对较低的原因在于即使是微小的勾画差异也可能导致明显的体积变化,且OAR的边界通常难以准确鉴别。另外,钱建军等^[22]进一步指出,低年资医师组在勾画OAR时的平均体积偏差值高于高年资医师组。但经过集中的勾画培训后,医师们再次勾画OAR的DSC值有所提高,同时低年资医师的平均体积偏差值也有所降低。同时,李丹明等^[23]也分析了同一肿瘤中心的4名医师在学习同一指南前后对NPC患者OAR的勾画差异,发现学习指南后,不同医师OAR勾画差异也有所减少。这为减少不同级别肿瘤中心医师之间勾画差异提供有效的解决思路。此外,本研究发现相比于不同级别肿瘤中心医师之间的勾画差异,市级肿瘤中心3名医生勾画差异有一定的减小,如在不同级别肿瘤中心所勾画右侧视神经平均DSC为 $0.47\pm0.21\sim0.48\pm0.22$,而市级肿瘤中心医师之间的平均DSC为 $0.58\pm0.22\sim0.68\pm0.22$,这可能是由于不同级别肿瘤中心之间缺乏详细、清晰的指南规定以及勾画一致性的培训所导致的。因此,为了进一步提高NPC放疗的精确性和治疗效果,加强医师的培训和指导,以及制定更为详细和清晰的勾画指南,显得尤为重要。

此外,本研究创新性地研究和分析了不同级别肿瘤中心医师之间勾画不同分期NPC患者靶区和OAR的勾画差异大小,结果显示,不同级别肿瘤中心医师之间对不同分期患者靶区(GTVnx)勾画存在差异,相对早期患者(I或II期)的靶区勾画,他们之间在对晚期患者(III或IV期)靶区的勾画差异更小,平均 ΔV_{diff} 和DSC值分别为 $98.31\%\pm67.36\%$ vs $69.38\%\pm72.61\%$ ($P<0.05$)和 0.55 ± 0.08 vs 0.72 ± 0.12 ($P<0.05$)。主要原因可能是因为市级肿瘤中心医生对早期NPC靶区边界判断不够自信,而趋向于勾画更大体积的靶区以防止漏靶,而晚期NPC患者由于肿瘤侵犯范围广,且多侵犯重要OAR,故无法进一步扩大靶区范围,以防止OAR出现严重放疗并发症。

本研究尚有一些不足,首先,本研究仅对比分析了两级肿瘤中心医师(区域级与市级)在勾画NPC放疗器官体积时的差异。为了更全面地为临床试验提供参考数据,未来的研究将纳入县级肿瘤中心医师的勾画数据,从而分析3级肿瘤中心医师(区域级-市级-县级)之间在放疗靶区和OAR勾画方面的差异。这将有助于推动放疗同质化勾画,为临床实践提供更有价值的参考。其次,本研究主要采用了DSC、 ΔV_{diff} 等几何评价参数来评估勾画精度。然而,这些参数在评价小体积OAR的勾画精度时具有一定的局限性。因此,未来的研究将探索更多体积轮廓参数差异,如平均表面距离和豪斯多夫距离,更全面地评估放疗靶区和OAR勾画差异。此外,还将考虑加入剂量学参数评价,更深入地探讨勾画差异对临床治疗效果的影响。

在开展不同级别多中心的放疗临床试验时,确保不同级别肿瘤中心医师对放疗靶区和OAR的勾画一致性是保证放疗剂量学数据准确性的先决条件^[24-25]。基于本研究结果,建议开展多中心临床试验时有必要对早期NPC患者的靶区和小体积OAR的勾画结果作统一的评估和修改,尽可能缩小个体勾画差异对临床试验结果的影响。

4 结 论

本研究显示,不同级别肿瘤中心医师在NPC放疗计划靶区和OAR勾画方面存在差异,尤其在早期患者靶区和小体积OAR勾画时。为确保多中心临床试验的准确性,建议对不同级别医师进行统一培训并审核勾画结果,以减少差异对治疗结果的影响,提高放疗的同质化水平。

【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 Cancer J Clin, 2021, 71(3): 209-249.
- [2] Mackie TR, Kapatoes J, Ruchala K, et al. Image guidance for precise conformal radiotherapy[J]. Int J Radiat Oncol Biol Phys, 2003, 56(1): 89-105.
- [3] Fang FM, Chien CY, Tsai WL, et al. Quality of life and survival outcome for patients with nasopharyngeal carcinoma receiving three-dimensional conformal radiotherapy vs. intensity-modulated radiotherapy-a longitudinal study[J]. Int J Radiat Oncol Biol Phys, 2008, 72(2): 356-364.
- [4] 国家癌症中心/国家肿瘤质控中心. 鼻咽癌靶区勾画和计划设计指南[J]. 中华放射肿瘤学杂志, 2022, 31(11): 971-991. National Cancer Center/National Cancer Quality Control Center. Guideline of target delineation and treatment planning for nasopharyngeal carcinoma [J]. Chinese Journal of Radiation Oncology, 2022, 31(11): 971-991.
- [5] Sun Y, Yu XL, Luo W, et al. Recommendation for a contouring method and atlas of organs at risk in nasopharyngeal carcinoma patients receiving intensity-modulated radiotherapy[J]. Radiother Oncol, 2014, 110(3): 390-397.
- [6] 彭应林, 孙文钊, 程皖琴, 等. 鼻咽癌调强计划靶区和危及器官不同勾画方式下个体化差异观察[J]. 中华放射肿瘤学杂志, 2019, 28(10): 762-766. Peng YL, Sun WZ, Cheng WQ, et al. Interobserver variations in the delineation of planning target volume and with organs at risk different contouring methods in intensity-modulated radiation therapy for nasopharyngeal carcinoma [J]. Chinese Journal of Radiation Oncology, 2019, 28(10): 762-766.
- [7] Nelms BE, Tomé WA, Robinson G, et al. Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer [J]. Int J Radiat Oncol Biol Phys, 2012, 82(1): 368-378.
- [8] Feng M, Demiroz C, Vineberg KA, et al. Normal tissue anatomy for oropharyngeal cancer: contouring variability and its impact on optimization[J]. Int J Radiat Oncol Biol Phys, 2012, 84(2): e245-e249.
- [9] Stapleford LJ, Lawson JD, Perkins C, et al. Evaluation of automatic atlas-based lymph node segmentation for head-and-neck cancer[J]. Int J Radiat Oncol Biol Phys, 2010, 77(3): 959-966.
- [10] 刘洋, 张烨, 易俊林. 自动分割技术在鼻咽癌靶区及危及器官勾画应用价值的研究[J]. 中华放射肿瘤学杂志, 2021, 30(9): 882-887. Liu Y, Zhang Y, Yi JL. The value of automatic segmentation of target volume and organs at risk for nasopharyngeal carcinoma[J]. Chinese Journal of Radiation Oncology, 2021, 30(9): 882-887.
- [11] Tao CJ, Yi JL, Chen NY, et al. Multi-subject atlas-based auto-segmentation reduces interobserver variation and improves dosimetric parameter consistency for organs at risk in nasopharyngeal carcinoma: a multi-institution clinical study[J]. Radiother Oncol, 2015, 115(3): 407-411.
- [12] Lin L, Dou Q, Jin YM, et al. Deep learning for automated contouring of primary tumor volumes by MRI for nasopharyngeal carcinoma[J]. Radiology, 2019, 291(3): 677-686.
- [13] 彭应林, 游雁, 韩非, 等. ABAS软件勾画OAR临床前测试重要性研究[J]. 中华放射肿瘤学杂志, 2016, 25(6): 609-614. Peng YL, You Y, Han F, et al. Commissioning of an atlas-based auto-segmentation software for application in organ contouring of radiotherapy planning[J]. Chinese Journal of Radiation Oncology, 2016, 25(6): 609-614.
- [14] Voet PW, Dirks ML, Teguh DN, et al. Does atlas-based autosegmentation of neck levels require subsequent manual contour editing to avoid risk of severe target underdosage? A dosimetric analysis[J]. Radiother Oncol, 2011, 98(3): 373-377.
- [15] Carillo V, Cozzarini C, Perna L, et al. Contouring variability of the penile bulb on CT images: quantitative assessment using a generalized concordance index[J]. Int J Radiat Oncol Biol Phys, 2012, 84(3): 841-846.
- [16] Geets X, Daisne JF, Arcangeli S, et al. Inter-observer variability in the delineation of pharyngo-laryngeal tumor, parotid glands and cervical spinal cord: comparison between CT-scan and MRI[J]. Radiother Oncol, 2005, 77(1): 25-31.
- [17] Gwynne S, Spezi E, Sebag-Montefiore D, et al. Improving radiotherapy quality assurance in clinical trials: assessment of target volume delineation of the pre-accrual benchmark case [J]. Br J Radiol, 2013, 86(1024): 20120398.
- [18] Peters LJ, O'Sullivan B, Giralt J, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: results from TROG 02.02[J]. J Clin Oncol, 2010, 28(18): 2996-3001.
- [19] 孙颖, 陶娟娟. 鼻咽癌正常组织勾画现状及改进[J]. 广东医学, 2014, 35(1): 5-6. Sun Y, Tao CJ. Current status and improvement of normal tissue delineation in nasopharyngeal carcinoma [J]. Guangdong Medical Journal, 2014, 35(1): 5-6.
- [20] Peng YL, Chen L, Shen GZ, et al. Interobserver variations in the delineation of target volumes and organs at risk and their impact on dose distribution in intensity-modulated radiation therapy for nasopharyngeal carcinoma[J]. Oral Oncol, 2018, 82: 1-7.
- [21] Liu X, Huang HX, Zhu CH, et al. Interobserver variations in target delineation in intensity-modulated radiation therapy for nasopharyngeal carcinoma and its impact on target dose coverage[J]. Technol Cancer Res Treat, 2023, 22: 15330338231169592.
- [22] 钱建军, 邢鹏飞, 陆雪官, 等. 鼻咽癌IMRT中OAR勾画差异所致剂量不确定性研究[J]. 中华放射肿瘤学杂志, 2014, 23(3): 239-243. Qian JJ, Xing PF, Lu XG, et al. Differences in delineation of organs at risk lead to dose uncertainties during intensity-modulated radiotherapy for nasopharynx carcinoma [J]. Chinese Journal of Radiation Oncology, 2014, 23(3): 239-243.
- [23] 李丹明, 穆庆霞, 孙新臣, 等. 鼻咽癌部分危及器官勾画差异对剂量分布的影响[J]. 临床肿瘤学杂志, 2015, 5: 445-449. Li DM, Mu QX, Sun XC, et al. Dosimetric impact of organs at risk delineation variability of inter-observer on patients with nasopharyngeal carcinoma [J]. Chinese Clinical Oncology, 2015, 5: 445-449.
- [24] Deeley MA, Chen A, Datteri R, et al. Comparison of manual and automatic segmentation methods for brain structures in the presence of space-occupying lesions: a multi-expert study[J]. Phys Med Biol, 2011, 56(14): 4557-4577.
- [25] Li XA, Tai A, Arthur DW, et al. Variability of target and normal structure delineation for breast cancer radiotherapy: an RTOG multi-institutional and multiobserver study[J]. Int J Radiat Oncol Biol Phys, 2009, 73(3): 944-951.

(编辑:谭斯允)