

瓦里安RapidPlan模型训练中统计离群值的处理及其剂量学影响

吴昊, 蒋璠, 岳海振, 张艺宝

北京大学肿瘤医院暨北京市肿瘤防治研究所放疗科, 恶性肿瘤发病机制及转化研究教育部重点实验室, 北京 100142

【摘要】目的:以统计图表和参数为依据, 分析处理瓦里安RapidPlan模型离群值, 并比较处理前后计划之间及其与人工优化的剂量学差异。**方法:**(1)选取80例直肠癌术前同步推量计划建立模型;(2)修正明显原因所致的离群值;(3)利用拟合曲线、Z值和库克距离寻找几何离群值, 检查原计划并分情况处理;(4)利用残差分布图、剂量体积直方图、学生化残差值和库克距离判断剂量离群值并酌情处理;(5)利用20例同类计划测试验证前后模型的优化效果。**结果:**统计学验证使模型构成计划中的离群参数最大值显著降低($P<0.05$)。相比原人工优化的测试计划, RapidPlan使用任一模型均能显著改善靶区均匀性指数($P<0.05$, 幅度2.7%~19.0%)和膀胱平均剂量($P<0.05$, 降幅12.8%~13.2%)。但利用验证前后模型分别优化的计划质量差异不大:肿瘤计划靶区(PGTV)均匀性指数相差0.5%, PGTV适形指数相差0.1%, 计划靶区适形指数相差0.5%, 股骨头和膀胱平均剂量分别相差0.3%和0.4%($P>0.05$)。**结论:**基于知识的RapidPlan计划可以改善靶区剂量均匀性和保护危及器官。依据统计学参数删改模型构成计划并不一定能取得更好的剂量学效果。

【关键词】RapidPlan; 基于知识的计划; 危及器官; 离群值; 剂量学评估

【中图分类号】R815; R730.55

【文献标志码】A

【文章编号】1005-202X(2016)07-0649-05

Statistical outlier processing and dosimetric effect in Varian RapidPlan model verification

WU Hao, JIANG Fan, YUE Hai-zhen, ZHANG Yi-bao

Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Radiotherapy, Peking University Cancer Hospital and Institute, Beijing 100142, China

Abstract: Objective Based on the statistical diagram and parameter, to analyze and process the outliers in Varian RapidPlan model, and to compare dosimetric differences before and after processing, and to compare the dosimetric differences between model optimization with the manual optimization. **Methods** Totally, 80 preoperative simultaneous integrated boosting plans for rectal cancer were selected to construct the model. The outliers caused by obvious reasons were firstly corrected. The geometric outliers were detected by using fitting curves, Z-score value and Cook's distance (CD), and the original plan were tested and processed according to different situations. The dosimetric outliers were judged and accordingly processed based on the residual plots, dose-volume histogram, studentized residual and CD. The optimization results of model before and after statistical verification were tested by 20 similar plans. **Results** The maximum outlier indicators were significantly reduced by statistical verification ($P<0.05$). Compared with the original manual optimization plan, any model of RapidPlan significantly improved the homogeneity index (HI) of target volume ($P<0.05$, with a range of 2.7%-19.0%) and reduced the mean dose of urinary bladder ($P<0.05$, with a range of 12.8%-13.2%). No significant differences were found in the quality of plans optimized by model before and after verification. The differences in the HI of planning gross target volume (PGTV), conformity index (CI) of PGTV, CI of planning target volume (PTV), mean dose of femoral head and urinary bladder were respectively 0.5%, 0.1%, 0.5%, 0.3% and 0.4% ($P>0.05$). **Conclusion** The knowledge-based RapidPlan improves the HI of target volume and the protection for organs at risk. The adjustment of model components based on the statistical indicators does not guarantee better dosimetric consequences.

Key words: RapidPlan; knowledge-based plan; organs at risk; outlier; dosimetric evaluation

【收稿日期】2016-02-10

【基金项目】国家自然科学基金(11505012, 81402535); 北京市医院管理局“青苗”计划专项(QML20151004); 质检公益性行业科研专项(201510001-02)

【作者简介】吴昊, 硕士, 高级工程师, 研究方向: 医学物理, E-mail: wuhao_bicr@sina.com

【通信作者】张艺宝, 博士, 高级工程师, 研究方向: 医学物理, E-mail: ybzhzhang66@163.com

前言

“基于知识”的放疗计划作为下一代计划系统的代表性技术之一^[1],不仅可以通过预测计划结果和参数^[2]实现(半)自动优化^[3],从而提高设计效率^[4]和进行计划质控^[5],同时可以避免因主观因素造成的欠佳计划和质量差异^[6-11]。瓦里安 Eclipse 计划在 V13.5 及之后版本中提供了基于知识的自动优化模块, RapidPlan 经国外的临床前验证表明,其在头颈^[12]、肺^[13]、食管^[14]、肝^[15]、前列腺^[16]等部位的初步应用可以达到甚至超越原人工计划质量。还有研究将 RapidPlan 用于个体计划的自动化评估,也取得了不错的效果^[17]。

在 RapidPlan 自动优化算法的所有核心组件中,剂量体积直方图预测模型是决定最终计划成败的首要因素。该模型通常需要由用户自行利用既往同类型、高质量计划建立,并且经过统计学和剂量学双重验证。按照说明书要求,验证过程需要根据统计学提示对每个离群值对应的原计划进行逐一排查处理,而每改变一次模型组分都需要重新进行训练拟合,并由此产生新的离群值^[18]。整个流程十分繁琐耗时。

由于 RapidPlan 在全球的应用尚处于起步阶段,模型验证中迭代过程的终点指征及其对不同放疗部位的剂量学影响仍处于摸索和经验积累阶段^[19]。本文以直肠癌术前同步推量放疗计划为例,探讨容积调强模型训练中对不同类型离群值的处理,并比较验证前后模型优化测试计划的结果。

1 材料与方法

1.1 病例选择及初始模型创建

从2014年5月~2015年6月在北京大学肿瘤医院接受术前同步推量放疗的直肠癌患者中选取80例计划,均在 Eclipse 11 版系统上利用1~2个整弧(取决于计划的难易程度),10 MV 光子, PRO_11031 优化算法和 AAA_11031 剂量体积算法,分22次使95%肿瘤计划靶区(PGTV)及计划靶区(PTV)同步达到50.6 Gy和41.8 Gy的处方剂量,并控制 $V_{107\%}$ 热点区域。在 PTV 与 PGTV 之间设置5 mm的剂量跌落区:PTV-(PGTV+5 mm),并将其作为辅助结构赋上限值进行剂量跌落优化。危及器官保护目标包括尽可能降低股骨头和膀胱的平均剂量(D_{mean})。利用上述计划在 V13.5 版 Eclipse 系统的 RapidPlan 模块中创建剂量体积直方图(DVH)预测模型。

1.2 模型的统计学确认及修改

除按照生产商提供的指南观察各种统计图

外^[20],为发现并分类型处理训练完成后模型中可能存在的几何、剂量离群值以及拟合强影响点计划,本文为改进的 Z 值(Modified Z-score, mZ)、学生化残差值(Studentized Residual Values, SR)和库克距离(Cook's Distance, CD)分别设置阈值为3.5、3.0、10.0(曾经建议值为4.0),并对超过阈值的计划进行分批逐次逐一处理,以重新训练后模型参数值的变化趋势判断上述处置是否得当。按照离群值可能产生的原因及严重程度,本文设计了3个渐进步骤:

(1)初步验证。记录每次训练后所有超限参数对应的原计划 ID,检查并剔除可能混入的非同类型计划,修改模型建立过程中可能存在的器官结构误配、处方剂量不一致或不是最高剂量处方等问题。之后重新训练模型并采用相同方法再次排查,直到确认所有离群值均不由上述原因导致。

(2)几何离群值的处理。对于重新训练的模型,检查拟合曲线和数据点分布,对其中的离群值、 $mZ > 3.5$ 以及由结构勾画导致 $CD > 10.0$ 的计划,分以下4种情况进行处理:①剔除靶区形状或体积明显异于其他患者的病例,否则保留所有仅 PTV-(PGTV+5 mm)参数超限的计划以维持靶区的几何多样性;②对于正常解剖结构的个体差异造成的离群值,如果不能为差异较大者找到足够数量的相似病例加入模型,则从库中予以剔除,但为了保护模型中危及器官的几何多样性及代表性,本文选择保留其中大部分计划;③对于因勾画者主观原因造成的危及器官几何离群值,通过返回修改原计划以提高模型的几何一致性;④对于控制 U 型靶区包绕部分正常组织热区的剂量学辅助结构(如部分小肠、结肠等),因其主观随意性和计划者间差异较大,故除明显异常病例外不作处理。对再次训练拟合的模型重复上述过程,进一步处理新出现的几何离群值。

(3)剂量离群值的处理。完成上述步骤后重新训练拟合模型,审查残差分布图和 DVH 分布图,检查 $SR > 3.0$ 以及由于剂量原因导致 $CD > 10.0$ 的计划,除因与靶区交叠明显过多或危及器官本身体积过小等客观原因导致的高剂量离群值外,均被列为可疑欠优计划并由高年资物理师进行重新优化。如计划质量因此提高则以其替代原计划并重新训练模型。对于危及器官剂量异常偏低的计划则需排除勾画失当等非客观原因导致的离群值。

1.3 验证前后模型的剂量学比

另选取与模型类型相同的20例临床计划作为测试病例,分别以验证前后的模型对其进行 RapidPlan 优化并分别与原计划进行比较,同时比较两种模型间的剂

量学效果。所有计划均按处方剂量归一(95%体积满足100%处方剂量)以求在相似靶区剂量覆盖的前提下比较以下参数:靶区均匀性指数(HI): $HI=(D_{2\%}-D_{98\%})/D_{50\%}$;靶区适形指数(CI): $CI=V_{100\%}/V_{target}$;PGTV中超过107%处方剂量的热点体积 $V_{107\%}$;PTV-(PGTV+5mm)、股骨头及膀胱的 D_{mean} (D_{mean_p} 、 $D_{mean_股}$ 和 $D_{mean_膀}$)。

1.4 统计学方法

利用SPSS 21.0软件中提供的Shapiro-Wilk方法检验数据分布的正态性,对于符合及不符合正态的数据分别采用配对 t 检验和Wilcoxon秩和检验以判断差异是否显著(双侧, $P<0.05$)。将表单格式的DVH数据从Eclipse系统中导入自编的MATLAB程序,计算20例患者的平均DVH数据的计划类型分别为:原计划、验证前模型以及验证后模型优化的RapidPlan计划,并以SigmaPlot V 10.0软件(Systat, San Jose, CA)进行绘图比较。

2 结果

2.1 离群值处置及统计学验证结果

模型的统计学验证过程直接剔除各类离群值对应的训练计划20例。修改原计划12例,其中7例因再次产生新的较大离群值而被排除,5例得以最终保留,最终新模型由53例计划构成。表1分析了构成新旧模型的计划中最大离群值参数的分布统计(不含剂量辅助结构)。

表1 验证前后构成各模型的计划离群参数最大值分布比较
Tab.1 Comparison of maximum outlier indicators before and after verification

Item	Before verification ($n=80$)	After verification ($n=53$)
Mean	4.51	3.98
95% CI	3.01-6.00	3.01-4.95
SD	6.73	3.51
Normality	$P<0.001$	$P<0.001$
P value	0.002	

SD: Standard deviation; 95%CI: 95% confidence intervals

由于DVH预测模型以危及器官为主要优化对象,虽然PTV-(PGTV+5 mm)在模型中被视为非靶区结构进行高量控制,但受靶区复杂程度的影响,形状和体积的一致性远差于其他危及器官,因此出现了大量的离群值。为了消除该影响,当计划的超限离群参数最大值由PTV-(PGTV+5 mm)产生时,则统计产生次大值的危及器官参数,结果如表2所示。

表2 排除由PTV-(PGTV+5 mm)导致的超限离群参数最大值之后模型验证前后各计划离群参数最大值的分布比较

Tab.2 After excluding the maximum outlier indicator caused by PTV-(PGTV+5 mm), comparing maximum outlier indicators before and after model verification

Item	Before verification ($n=80$)	After verification ($n=53$)
Mean	3.21	3.13
95% CI	2.41-4.00	2.57-3.68
SD	3.56	2.01
Normality	$P<0.001$	$P<0.001$
P value	<0.001	

PGTV: Planning gross target volume; PTV: Planning target volume

2.2 剂量学比较

表3对20例人工优化的临床计划与利用验证前后模型优化的RapidPlan计划之间进行了两两比较。除此之外,在3例(占比15%)临床计划的PGTV中发现了 $V_{107\%}>0$ 的剂量热点,分别为0.14%、0.01%和0.26%,而使用任何一种模型优化的RapidPlan均未发现超过107%处方剂量的热区。图1对比绘制了20例患者的平均DVH图,直观展示了人工计划与两套RapidPlan计划之间的剂量学差异。

3 讨论

除异类型计划、结构误配、处方剂量定义等客观原因外,模型验证过程中对于离群值计划的处理存在较大主观性。 mZ 、SR和CD值等提示的仅是可能的离群值,需要进一步结合其在分布图中的位置并返回原计划进行分析判断。

验证后模型的最大离群参数值显著低于验证前($P=0.002$),且分布范围更窄(表1)。在排除统计由PTV-(PGTV+5 mm)产生的最大离群值后(表2),模型最大离群参数值进一步大幅降低,且变化幅度进一步缩窄。仅从统计参数评价,验证后的模型在计划构成一致性等方面优于原模型。

分析表3测试组的原临床计划与RapidPlan计划的剂量学差异发现,无论模型是否经过统计学验证,RapidPlan在PGTV和PTV剂量HI以及膀胱 D_{mean} 控制方面都显著优于传统人工优化计划($P<0.05$)。相比原计划,RapidPlan计划的PGTV HI降幅为19.0%(验证前和验证后相同),PTV HI降幅分别为5.0%(验证前)和2.7%(验证后),且均未见超过107%处方剂量热区;而膀胱 D_{mean} 的降幅可分别达到13.2%(验证前)和12.8%(验证后)。另一方面,股骨头的 D_{mean}

表3 模型验证前后 RapidPlan 计划与原临床计划的相互比较

Tab.3 Dosimetric comparison between original plan and RapidPlan optimized by the models before and after verification

Item		HI		CI		$D_{\text{mean}}/\text{cGy}$		
		PGTV	PTV	PGTV	PTV	PTV-(PGTV+5 mm)	Femoral head	Urinary bladder
Clinical	Mean	0.058	0.260	0.986	1.026	4 397	1 576	2 886
	SD	0.009	0.009	0.026	0.023	39	169	223
	95%CI	0.053-0.616	0.255-0.264	0.973-0.998	1.015-1.037	4 379-4 415	1 497-1 655	2 782-2 990
Before	Mean	0.047	0.247	0.973	1.027	4 402	1 498	2 505
	SD	0.006	0.004	0.020	0.024	33	64	100
	95%CI	0.044-0.049	0.245-0.249	0.964-0.983	1.016-1.038	4 387-4 417	1 468-1 528	2 458-2 552
After	Mean	0.047	0.253	0.974	1.022	4 405	1 503	2516
	SD	0.006	0.007	0.111	0.017	31	30	64
	95%CI	0.044-0.050	0.250-0.256	0.922-1.026	1.014-1.029	4 390-4 419	1 489-1 517	2 486-2 546
P value	P_1	<0.001*	<0.001*	0.191	0.911*	0.580*	0.093	<0.001
	P_2	<0.001*	0.036*	0.737	0.162*	0.382*	0.135	<0.001
	P_3	0.802*	0.005*	0.117	0.114*	0.524*	0.654	0.178

*was tested by paired t test, and the other was tested by Wilcoxon test. P_1 : Clinical vs. Before; P_2 : Clinical vs. After; P_3 : Before vs. After.

HI: Homogeneity index; CI: Conformity index

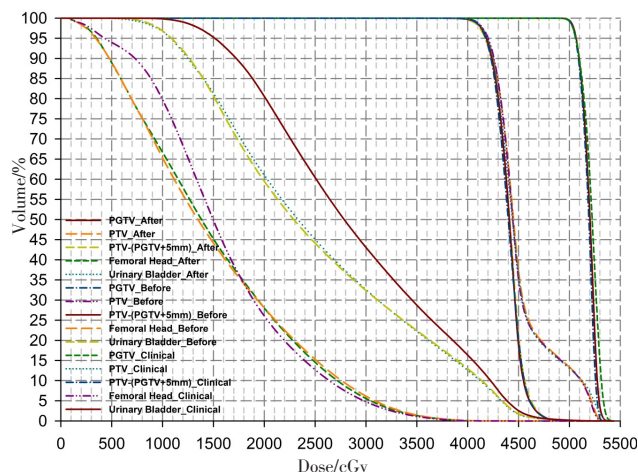


图1 以临床计划、验证前、后模型优化的 RapidPlan 计划进行分组而绘制成的 20 例患者平均 DVH 图

Fig.1 Average DVH of 20 patients grouped by clinical plan, RapidPlan optimized by model before and after verification

降低 5.0% (验证前) 和 4.6% (验证后), 但无统计学意义。除此之外 RapidPlan 计划与原人工计划的剂量学差异可以忽略, 包括 PGTV CI 相差 1.3% (验证前) 和 1.2% (验证后), PTV CI 相差 0.1% (验证前) 和 0.4% (验证后), PTV-(PGTV+5 mm) D_{mean} 相差 0.1% (验证前) 和 0.2% (验证后), 且均无统计学意义。导致 RapidPlan 表现较优的算法层面原因是: DVH 预测模型会依据靶区、危及器官与射野的相对位置关系等

对优化结构做量化细分和评估, 并利用训练计划提取数据拟合而成的模型对每一部分分别作出剂量预测。该算法有效避免了人工优化过程主观性导致的欠优计划。因此, 选择优质且具有广泛代表性的计划用于模型训练对于 RapidPlan 的最终表现至关重要。

对比表 3 两套模型分别优化的 RapidPlan 计划可以发现, 除 PTV HI 一项差异具有统计学意义外 (确认后的模型高 2.4%, $P=0.005$), 其余差异均不显著 ($P>0.050$), 且差异幅度可以忽略, 包括: PGTV HI 相差 0.0%, PGTV CI 相差 0.1%, PTV CI 相差 0.5%, PTV-(PGTV+5 mm) D_{mean} 相差 0.1%, 股骨头 D_{mean} 相差 0.3%, 膀胱 D_{mean} 相差 0.4%。

图 1 所示的平均 DVH 图进一步印证了对表 3 数据的解释。模型优化前后 RapidPlan 计划各结构 DVH 曲线近乎重合 (对应表 3 无统计学意义的小幅差异), 而与此相对的原临床计划在 PGTV 和 PTV 高剂量区跌落速度 (对应表 3 HI 的大幅显著差异)、股骨头和膀胱 D_{mean} 控制等方面均有较大幅度的劣势。

利用既往经验知识和个体化结构剂量关系评估, RapidPlan 可以有效发现人工优化的欠优计划, 并在靶区剂量均匀性和危及器官保护等方面取得进步。然而, 作为模型训练的必要过程之一, 统计学指标的改善并不一定能够提高最终的剂量学表现。这

与 Delaney 等^[19]在头颈部模型验证实验中观察到的结论一致。导致该现象的原因之一是少数离群值的作用可能被大量的训练病例所“稀释”, 另外也可能与统计学验证过程中删除的病例特征有关: 该操作可能导致与某些测试计划具有相似靶区/危及器官相对位置关系的计划被删除, 从而使模型的几何多样性和代表性变差。从现有文献资料和研究结果来看, 通过分析统计图表和参数进而排除明显异类计划和结构误配, 以及通过修改勾画和重新优化后的更优计划替代原离群值计划等手段可能更加可靠, 更容易以相对较低的时间代价取得相似的剂量学效果。

鉴于统计验证过程的复杂性、繁琐性和不确定性, 未来需要在模型计划入排标准、迭代过程终止指征、高效剂量学验证方法等方面做进一步的研究探讨。

4 结 论

无论 DVH 预测模型是否经过统计学验证, 基于知识的 RapidPlan 均能较原人工计划获得更优的剂量学结果, 包括提高靶区剂量均匀性以及降低危及器官受量。依据统计参数盲目删除模型中的离群值计划并不一定能够显著提高测试计划的剂量学表现, 临床上需要在模型的几何代表性与拟合优度两方面寻找平衡。

【参考文献】

- [1] SHARPE M B, MOORE K L, ORTON C G. Within the next ten years treatment planning will become fully automated without the need for human intervention[J]. Med Phys, 2014, 41(12): 120601.
- [2] APPENZOLLER L M, MICHALSKI J M, THORSTAD W L, et al. Predicting dose-volume histograms for organs-at-risk in IMRT planning[J]. Med Phys, 2012, 39(12): 7446-7461.
- [3] ZAREPISHEH M, LONG T, LI N, et al. A DVH-guided IMRT optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning[J]. Med Phys, 2014, 41(6): 061711.
- [4] BOLAN C. Expediting the treatment planning process[J]. Appl Radiat Oncol, 2013: 19-23.
- [5] MOORE K L, BRAME R S, LOW D A, et al. Experience-based quality control of clinical intensity-modulated radiotherapy planning[J]. Int J Radiat Oncol Biol Phys, 2011, 81(2): 545-551.
- [6] YANG Y, FORD E C, WU B, et al. An overlap-volume-histogram based method for rectal dose prediction and automated treatment planning in the external beam prostate radiotherapy following hydrogel injection[J]. Med Phys, 2013, 40(1): 011709.
- [7] CHANYAVANICH V, DAS S K, LEE W R, et al. Knowledge-based IMRT treatment planning for prostate cancer[J]. Med Phys, 2011, 38(5): 2515-2522.
- [8] ZHU X, GE Y, LI T, et al. A planning quality evaluation tool for prostate adaptive IMRT based on machine learning[J]. Med Phys, 2011, 38(2): 719-726.
- [9] WU B, PANG D, SIMARI P, et al. Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study[J]. Med Phys, 2013, 40(2): 021714.
- [10] LIAN J, YUAN L, GE Y, et al. Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: an intertechnique and interinstitutional study[J]. Med Phys, 2013, 40(12): 121704.
- [11] WU B, MCNUTT T, ZAHURAK M, et al. Fully automated simultaneous integrated boosted-intensity modulated radiation therapy treatment planning is feasible for head-and-neck cancer: a prospective clinical study[J]. Int J Radiat Oncol Biol Phys, 2012, 84(5): e647-e653.
- [12] TOL J P, DELANEY A R, DAHELE M, et al. Evaluation of a knowledge-based planning solution for head and neck cancer[J]. Int J Radiat Oncol Biol Phys, 2015, 91(3): 612-620.
- [13] FOGIATA A, BELOSI F, CLIVIO A, et al. On the pre-clinical validation of a commercial model-based optimisation engine: application to volumetric modulated arc therapy for patients with lung or prostate cancer[J]. Radiother Oncol, 2014, 113(3): 385-391.
- [14] FOGIATA A, NICOLINI G, CLIVIO A, et al. A broad scope knowledge based model for optimization of VMAT in esophageal cancer: validation and assessment of plan quality among different treatment centers[J]. Radiat Oncol, 2015, 10: 220.
- [15] FOGIATA A, WANG P M, BELOSI F, et al. Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer [J]. Radiat Oncol, 2014, 9: 236.
- [16] NWANKWO O, MEKDASH H, SIHONO D S, et al. Knowledge-based radiation therapy (KBRT) treatment planning versus planning by experts: validation of a KBRT algorithm for prostate cancer treatment planning[J]. Radiat Oncol, 2015, 10: 111.
- [17] TOL J P, DAHELE M, DELANEY A R, et al. Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans? [J]. Radiat Oncol, 2015, 10: 234.
- [18] Varian Medical Systems. Treatment planning 13.5 new features RapidPlan[Z]. EC13.5-WBK-01-B. 2014: 80-108.
- [19] DELANEY A, TOL J P, DAHELE M, et al. Effect of dosimetric outliers on the performance of a commercial knowledge-based planning solution[J]. Int J Radiat Oncol Biol Phys, 2016, 94(3): 469-477.
- [20] Varian Medical Systems. Application quick tip: verification of a custom RapidPlan model[Z]. EC13.5-QTS-02-A. 2015: 1-12.

(编辑: 陈丽霞)