A sub-analysis of multi-center planning radiosurgery for intracranial metastases through automation (MC-PRIMA) comparing UK and international centers

Chan, Mark K.H.

Published in:
Medical Engineering & Physics

DOI:
10.1016/j.medengphy.2023.103996

Publication date:
2023

Medical Engineering and Physics 117 (2023) 103996

Contents lists available at ScienceDirect

Medical Engineering and Physics

journal homepage: www.elsevier.com/locate/medengphy

Communication

A sub-analysis of multi-center planning radiosurgery for intracranial metastases through automation (MC-PRIMA) comparing UK and international centers

Mark K.H. Chan [a,*]

[a] University Medical Center Groningen and University of Groningen, Groningen, The Netherlands

ARTICLE INFO

Keywords:
Plan crowdsourcing
Autoplanning
Stereotactic radiosurgery
Multiple brain metastases

ABSTRACT

Objectives: A sub-analysis of the MC-PRIMA study was performed to compare the plan quality of stereotactic radiosurgery (SRS) to multiple brain metastases (MBM) between UK and other international centres.
Methods and materials: Six centres from the UK and nineteen from other international centres autoplanned using Multiple Brain Mets™ (AutoMBM; Brainlab, Munich, Germany) software for a five MBM study case from a prior planning competition that was originally organized by the Trans-Tasmania Radiation Oncology Group (TROG). Twenty-three dosimetric metrics and the resulting composite plan score per the TROG planning competition were compared between the UK and other international centres. Planning experience and planning time from each planner were recorded and statistically compared.
Results: Planning experiences between two groups are equal. Except for mean dose to the hippocampus, all other 22 dosimetric metrics were comparable between two groups. The inter-planner variations in these 23 dosimetric metrics and the composite plan score were also statistically equivalent. Planning time is slightly longer in the UK group (mean = 86.8 min) with a mean difference of 50.3 min.
Conclusions: AutoMBM effectively achieves standardization of the plan quality of SRS to MBM within UK and further against the other international centres. Significant planning efficiency gain by AutoMBM both among the UK and other international centres may help to increase the capacity of SRS service by alleviating the clinical and technical loadings.

1. Introduction

In United Kingdom, Stereotactic radiosurgery / radiotherapy (SRS / SRT) is centralized at a number of specialist centres. A prerequisite for centres in England to be commissioned for SRS by the National Health Service in England (NHSE) was to pass the quality assurance (QA) that was established by the national trials QA group (RTTQA).

The RTTQA group published the planning benchmark results for two cases of 3 and 7 multiple brain metastases (MBM) [1]. This report showed wide variation in plan quality, exemplified by the large spread of values of Paddick conformity index (PCI) [2], dose gradient index (GI) [3] and half prescription isodose (R50%) [4]. More importantly, the variation of R50%, for example, was much greater on C-arm linac-based plans than GammaKnife (Elekta AB, Stockholm, Sweden) and CyberKnife (Accuray Inc., CA, USA) plans. This partly reflects the wide variety of planning and delivery systems within this broad category, but also may well be dependent on planning philosophy (e.g. normalization) and on the planner's skill influencing the plan quality for SRS on linacs.

Recently, a multicentre study (MC-PRIMA) has been performed to investigate the potential of autoplanning SRS to MBM on linac-based platforms using single isocentre [5]. Six of these participating centres were from UK and took part in the previous NHSE/RTTQA program. This study benchmarked a single case of five MBM autoplanned on the Elements Multiple Brain Mets (AutoMBM) SRS™ treatment planning solution (TPS) against 160 other plans on six other TPSs. Although AutoMBM significantly reduced the inter-institutional / -planner variability in various dosimetric metrics and overall plan quality, observable variability in the planning performance still existed. It was hypothesized that the observed variability might be connected to different planning philosophy behind autoplanning, specifically, differences in the definition of templates called Clinical Protocols and Setup Protocols, each catering to specific clinical objectives and beam irradiation

* Corresponding author at: Department of Radiation Oncology, University Medical Center Groningen, P.O. Box 30001, 9700 Groningen, The Netherlands.
E-mail address: k.h.chan@umcg.nl.

https://doi.org/10.1016/j.medengphy.2023.103996
Received 6 January 2023; Received in revised form 23 April 2023; Accepted 17 May 2023
Available online 18 May 2023

arrangement, respectively, in the plan evaluation and re-optimization.

In UK, multiple user meetings was organized by the vendor ofAutoMBM to promote experience sharing and therefore a more UK-wide standardized and optimized autoplanning practice could be expected. This work performed a sub-analysis of MC-PRIMA[5] to study if there existed differences in the autoplanning approach and practices using AutoMBM between the UK and the other international centres and the resulting impacts on the planning performance. Such comparative analysis is expected to provide further insights into factors that may contribute to improvements using AutoMBM.

2. Methods and materials

The study case was originally from the Local-HER-0 trial protocol [6] of Trans-Tasmania Radiation Oncology Group (TROG) and was used in a prior international planning competition. It comprised five metastases of 0.52, 0.39, 0.07, 2.82 and 0.12 cm^3. Further details of this study case are found in Ref. [7].

Twenty-five international institutions from seven regions (North / South America $n = 2$ /2; Europe $n = 16$; Asia $n = 2$; Middle East $n = 1$; Africa $n = 1$; Australasian $n = 1$), including 6 from the UK retrospectively autoplanned this case on AutoMBM. Briefly, AutoMBM is a template-based autoplanning solution dedicated to non-coplanar linac-based SRS using MLC and single isocenter [8].

Twenty-three dosimetric metrics (Table 2) and resulting composite plan score (maximum score = 150) were calculated for each center according to the same scoring matrix that was devised by the Local-HER-0 trial protocol as in the TROG planning competition. The scoring algorithm [9] was detailed for each of 23 dosimetric metrics in the cloud-based plan challenge platform called ProKnow[1] and in MC-PRIMA [5]. All participants were provided with detailed information of the scoring matrix before generating the AutoMBM plan. All scoring metrics were extracted from the AutoMBM software and along with the composite plan score were compared between the UK and other international centres by non-parametric Mann-Whitney U tests for difference in their values and squared rank tests for equality of their variations using Matlab v.R2018a (Mathwork Inc. MA, USA). Statistical difference was considered significant at p-value< 0.01 for comparisons of plan quality metrics and < 0.05 in other planning evaluations (e.g., number of table angles, gantry arc length, planning time, etc.).

Additionally, the general and SRS planning experiences of individual planner were surveyed. Each planner was also asked to log the actual planning time spent after the data, including the time of changing and creating new Clinical and Setup protocols, the optimizer to generate the plan dose distribution, and the planner to evaluate the dose distribution. Both the planning experience and planning time were compared between the UK and other centres. Spearman's correlations of planning experience and planning time with the composite plan score were performed separately for the UK and other centres.

3. Results

Table 1 summarizes the characteristics of equipment used in the autoplanning (AP) on AutoMBM. The UK and international centres used comparable number of couch kicks (mean = 5; range: 4–6 and 5–7, respectively). Means of gantry arc length were 148° (range: 120 ° – 160 °) for the UK centres and other centres 153° (range: 110 ° – 160 °) without statistical difference ($p > 0.05$).

Results of various dosimetric parameters in the scoring matrix per the TROG planning competition are given in Table 2. Fig. 1 plotted the GI versus the PCI values of all GTVs separately for the UK and other centres. Ranges of the PCI values resulting from the UK and other centres are 0.71–0.79 and 0.59–0.81, for GI 3.70–5.61 and 3.82–6.43, and for the spread of R50% values 3.19–3.90 and 3.11–5.57, respectively.

Fig. 2 shows the normalized scores of different scoring metrics and the normalized total composite plan scores.

Averaged SRS planning experiences of the UK and other centres are 5.2 and 7.8 years ($p > 0.05$). The mean planning time was 86.8 and 40.0 min for the UK and other centres, respectively ($p < 0.05$).

Neither the UK nor other international centres showed significant correlations between planning time and the overall plan quality in terms of the composite plan score (both $p > 0.05$; Spearman's correlation coefficients rho = 0.543 and −0.290, respectively), and between planning time and SRS planning experience (both $p > 0.05$). For the six UK centres, the planners estimated a mean treatment planning time of 688 min (range: 53–2250 min) using other non-AP solutions prior to using AutoMBM (range: 30–180 min). For other international centres, 40 and 298 min on average were estimated using AutoMBM and other non-AP solutions, respectively.

Table 1
Equipment used for benchmark case submissions, with numbers of platforms shown.

	UK	Others
Varian Linac	5	15
Multileafcollimator (MLC)		
• 2.5 mm (inner 8 cm) / 5 mm (outer 14 cm); max field size 22×30 cm^2	4	11
• 5.0 mm (inner 20 cm) / 10 mm (outer 20 cm); max field size 40×40 cm^2	1	2
Elekta Linac	1	4
Multileafcollimator (MLC)		
• 5 mm; max field size 40×40 cm^2	1	4
Nominal energy		
6MV	4	11
6MV Flattening filter free	2	8

Table 2
Means ± one standard deviations (SD) of various dosimetric parameters calculated for UK centers and other international centers according to the scoring matrix of TROG planning competition.

	UK ($n = 6$)	Others ($n = 19$)
Gross tumor volume (GTV)		
GTV1 *V20Gy* (%)	99.00 (0.00)	99.00 (0.00)
GTV2 *V20Gy* (%)	99.00 (0.00)	99.00 (0.00)
GTV3 *V20Gy* (%)	99.00 (0.00)	99.00 (0.00)
GTV4 *V20Gy* (%)	99.00 (0.00)	99.00 (0.00)
GTV5 *V20Gy* (%)	99.00 (0.00)	99.00 (0.00)
PCI of all GTVs(20 Gy)	0.75 (0.03)	0.70 (0.06)
R50% of all GTVs	4.72 (0.65)	5.15 (0.77)
GI of all GTVs	3.69 (0.28)	3.87 (0.60)
Normal organs		
Normal brain *V12Gy* (cc)	10.41 (2.02)	10.92 (1.99)
Normal brain *V10Gy* (cc)	14.94 (2.66)	15.24 (3.58)
Right hippocampus *Dmean* (Gy)	**2.92** (0.61)	2.28 (0.48)
Left hippocampus *Dmean* (Gy)	1.92 (0.15)	1.73 (0.32)
Right eye *Dmax* (Gy)	3.70 (1.01)	3.79 (1.32)
Left eye *Dmax* (Gy)	3.51 (1.20)	4.28 (1.25)
Right lens *Dmax* (Gy)	0.00 (0.00)	0.00 (0.00)
Left lens *Dmax* (Gy)	0.00 (0.00)	0.00 (0.01)
Right optic nerve *V8Gy* (cc)	1.88 (0.36)	1.71 (0.52)
Left optic nerve *V8Gy* (cc)	1.54 (0.38)	1.60 (0.45)
Chiasm *Dmax*(Gy)	4.92 (0.87)	4.24 (1.12)
Brainstem *D0.3cc*(Gy)	5.95 (0.88)	5.74 (0.89)
Composite plan score	126.95 (5.89)	127.75 (6.62)
Monitor units	8944 (7006)	9544 (4403)

Abbreviations: V*x*(Gy) = absolute or relative volume receiving *x* Gy; PCI = Paddick conformity index; R50% = half prescription isodose; GI = gradient index; D*x*cc = dose to *x* cc of the volume of interest. D*mean* = mean dose; D*max* = maximum dose.

[1] https://proknowsystems.com/

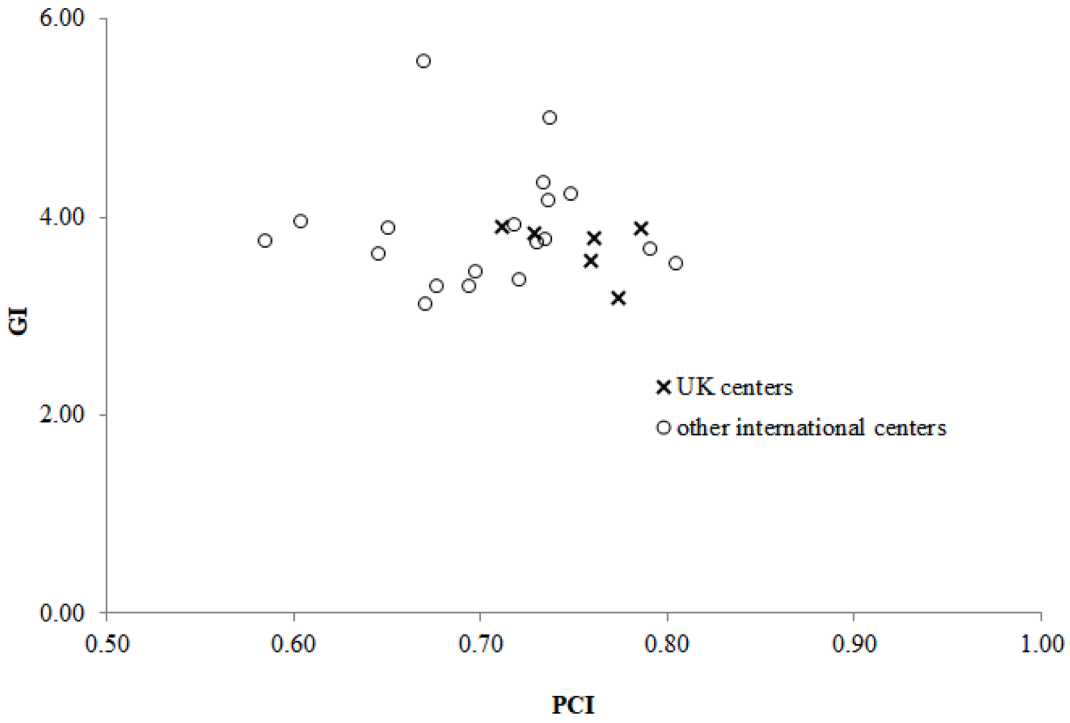

Fig. 1. GI against PCI of all GTVs ($n = 5$) was plotted for the UK ($n = 6$) and the other international centers ($n = 19$).

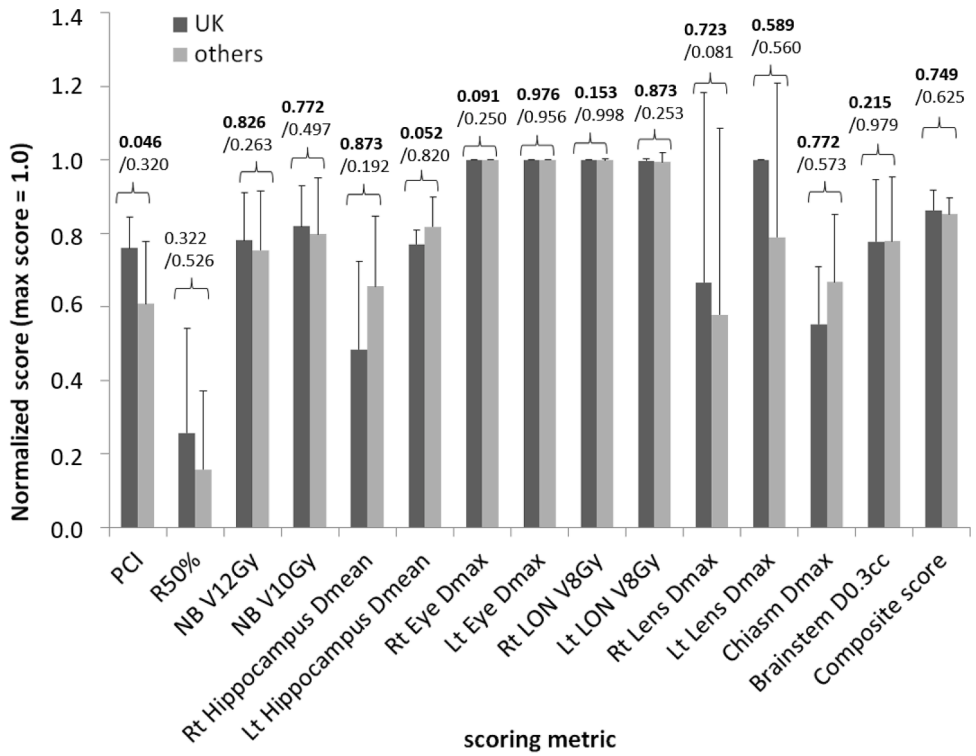

Fig. 2. Normalized scores for different scoring metrics achieved by AutoMBM treatment planning software. *p* values are given for Mann-Whitney U test for difference (bold) and for two-sample squared ranks test for equality of variance of the scoring metrics between UK and other international centers.

4. Discussions

This sub-analysis of the previous MC-PRIMA study [5] found that except for one (mean dose to right hippocampus) all dosimetric parameters are statistically equivalent between the UK and other international centres that autoplanned on the same study case by AutoMBM using similar linac treatment platforms. The inter-planner variability (one standard deviations; 1 S.D.) of all 23 dosimetric parameters was comparable between the UK and other centres. The lowest two PCI values of 0.59 and 0.60 in the international cohort were associated with the Elekta Versa HD platform with MLC width of 5 mm and the Varian Novalis treatment platform with MLC width of 2.5 mm, respectively.

However, these relatively poor PCI values were presumably unrelated to the machine platforms applied to AutoMBM nor the number of the treatment table positions and gantry arc lengths because other similar platforms and irradiation geometry were seen to achieve better results. For example, the other two centres that autoplanned with MLC width of 10 mm achieved PCI values of 0.65 and 0.73 based on the same number of treatment table positions as in the plans with the lowest two PCI values. The small variation of plan quality in the respective groups was partly attributed to the different radiation beam characteristics produced by a range of linac machines / models, the variation of user-predefined Clinical Goal templates as input to the AutoMBM planning software, and ultimately the interactive fine-tuning of planning objectives such as monitor unit modulation, prescription isodose level per lesions, etc.

The above results were unlikely influenced by the different planning experience (2.6 years) either. There was, however, a significance difference in time to plan this case. Because the specification of the hardware of the AutoMBM software was standardized by the vendor, the relatively large time difference ($p < 0.05$) in the autoplanning process may arise from repetition of optimizations by changing different preconfigured irradiation geometry and the plan re-evaluation. This subanalysis suggested that the UK centres can limit the manual intervention using AutoMBM by restricting the attempts to change the templates of irradiation geometry and by streamlining the plan evaluation process. A database of the Clinical Goal templates and the final fine-tuning plan parameters may foster the knowledge sharing to achieve further standardization and optimization of autoplanning practice with AutoMBM.

The small dispersions of PCI and GI values from the UK centres in MC-PRIMA (Fig.1) were in contrast to the wide spread corresponding values 0.39–0.91 (PCI), 2.94–8.22 (GI) reported in Fig.2 of the RTTQA study [1]. Their means were improved in MC-PRIMA compared to means of the three metastases case in the RTTQA study (PCI: 0.68; GI: 4.17), even though this related to a five-lesions case which should be more difficult to plan with a single isocentre. One may argue that direct comparison between MC-PRIMA and RTTQA studies cannot be strictly made because of the difference between the study cases in terms of the number, size and shape of the lesions. Direct comparison between MC-PRIMA and RTTQA studies is also challenging for other reasons. Firstly, all the platforms and techniques have evolved in the time since the NHSE/RTTQA study. For example, only a few plans used single isocentre VMAT on C-arm linacs, whereas this is now widespread and forms all the plans in the MC-PRIMA study. Secondly, in the 2016 study no guidance was given on what plan metrics were achievable, and large improvements were possible when feedback was given with results from other centres. However, it is likely that variation was influenced by planning philosophy and skill as well. One other major difference between the two studies is that a single TPS (AutoMBM) with dedicated AP functionality was used in MC-PRIMA whereas multiple TPS (iPlan, Pinnnacle, Eclipse, and Monaco) without dedicated AP solution were used in the RTTQA study. Reduced variation was found in the RTTQA study within plans produced for GK and CK platforms. These results lend further support that dedicated AP solution to SRS for C-arm linacs, such as AutoMBM can reduce the inter-planner variability and improve the l plan quality, which both has a strong dependence on the TPS in use [5].

This sub-analysis inherited the limitation of the primary MC-PRIMA baseing the plan evaluation on a single study case. Ideally, any treatment planning study should involve as many institutions as possible and cover a large number of clinical cases to allow faithful multivariate analysis of the planning results [9,10]. However, this is not always possible in practice as in most treatment planning quality assurance (QA)of clinical trials [11,12], multi-center study of planning benchmark [7,13–15], and even the Commission through Evaluation(CtE) program of NHSE for stereotactic radiosurgery a single benchmark case was often used [16]. When a single case is used for planning benchmark in multi-center study, it would be more important to select one case that is as clinically as representative by the experts, as in the Local-HER-0 trial [6,7] on which the study case of primary MC PRIMA study and the present sub-analysis was based on.

For the six UK centres that participated both in the MC-PRIMA and RTTQA studies, there was a drastic improvement of the planning efficiency by autoplanning. For this studied case of five metastases, 610 min (or 10.2 h) would have been saved if it had been planned using AutoMBM. This translates into an efficiency gain of over 5 working days if four MBM patients are scheduled in a single month. This study suggests that the further introduction of dedicated AP solutions to MBM can widen the geographic access to SRS by allowing consistently high quality plans to be produced independent of the experience of the center.

5. Conclusions

Dedicated SRS autoplanning solution to MBM effectively achieves plan standardization within UK and further against other international centres. Significant planning efficiency gain potentially increases the capacity of SRS service by alleviating the clinical and technical loadings.

Advances in knowledge

AutoMBM potentially enables standardization of SRS plan quality of multiple brain metastases among the UK and other international centres despite the observed difference in the autoplanning practice.

Funding

None.

Declarations

The following additional information is required for submission. Please note that this form runs over two pages and failure to respond to these questions/statements will mean your submission will be returned to you. If you have nothing to declare in any of these categories then this should be stated.

Ethical approval

Work on human beings that is submitted to *Medical Engineering & Physics* should comply with the principles laid down in the Declaration of Helsinki; Recommendations guiding physicians in biomedical research involving human subjects. Adopted by the 18th World Medical Assembly, Helsinki, Finland, June 1964, amended by the 29th World Medical Assembly, Tokyo, Japan, October 1975, the 35th World Medical Assembly, Venice, Italy, October 1983, and the 41st World Medical Assembly, Hong Kong, September 1989. You should include information as to whether the work has been approved by the appropriate ethical committees related to the institution(s) in which it was performed and that subjects gave informed consent to the work.

Declaration of Competing Interest

Mark K.H. Chan was a full-time employee at Imperial College Healthcare Trust, London, UK and University Hospital Essen, Essen, Germany during the partial preparation of this work.

Acknowledgements

The authors would like to thank all participants who contributed to the plan submissions, including Thierry Gevaert at Department of Radiation Oncology, UniversitairZiekenhuis Brussel, Vrije Universiteit Brussel, Belgium; SylviaGarny at University Hospital, Ludwig-Maximilians University of Munich, Germany; Manfred Wiencierz at University Hospital Essen, Germany; Noriyuki Kadoya at Department of

Radiation Oncology, Tohoku University Graduate School of Medicine, Japan; John Dorr at AdventHealth, USA; Carlos Otero Martinez at ComplexoHospitalario Universitario de Santiago, Spain; Ronnie Leung at Auckland Radiation Oncolgy, New Zealand; Samer Alheet at Department of Radiation Oncology, King Hussein Cancer Center, Jordan; Abdelkader Toutaoui at ChahidsMahmoudi Hospital, Algeria; Ruben Farias at Department of Radiation Oncology, Mevaterapia Medical Center, Argentina; Matthew Wong at GenesisCare, Cabrini Hospital, Australia; Daniel Venencia at Instituto Zunino-Fundacion Marie Curie, Argentina; Alexander Malesevic at Brainlab, Germany. Christina Skourou at Department of Physics, St. Luke's Radiation Oncology Network, Ireland, Marco Valenti at Unit of Medical Physics - OspedaliRiuniti Umberto I, G.M. Lancisi, G. Salesi Ancona Italy, Ignacio ModolellFarré at Servei de Física Mèdica i Protecció Radiològica, InstitutCatalàd'Oncologia, L'Hospitalet del Llobregat, Spain; Maximilian Grohmann at Department of Radiotherapy and Radiation Oncology, University Medical Center Hamburg, Germany; Haisong Liu at Department of Radiation Oncology, Thomas Jefferson University, USA

References

[1] Eaton DJ, Lee J, Paddick I. Stereotactic radiosurgery for multiple brain metastases: results of multicenter benchmark planning studies. Pract Radiat Oncol 2018;8(4):e212–20.

[2] Paddick I. A simple scoring ratio to index the conformity of radiosurgical treatment plans. J Neurosurg 2000;93(Suppl 3):219–22.

[3] Paddick I, Lippitz B. A simple dose gradient measurement tool to complement the conformity index. J Neurosurg 2006;105(Supplement):194–201.

[4] Timmerman R, et al. Stereotactic body radiation therapy for inoperable early stage lung cancer. JAMA, 2010;303(11):1070–6.

[5] Chan MKH, et al. Multi-center planning study of radiosurgery for intracranial metastases through automation (MC-PRIMA) by crowdsourcing prior web-based plan challenge study. Phys Med 2022;95:73–82.

[6] https://clinicaltrials.gov/ct2/show/NCT02898727 [Accessed date: 21-06 2021].

[7] Hardcastle N, et al. OC-0521 SRS plan quality with variation in modality: results of an international planning competition. Radiother Oncol 2019;133:S272–3.

[8] Gevaert T, et al. Evaluation of a dedicated brain metastases treatment planning optimization for radiosurgery: a new treatment paradigm? Radiat Oncol 2016;11(1):13.

[9] Hansen CR, et al. Radiotherapy treatment plannINg study guidelines (RATING): a framework for setting up and reporting on scientific treatment planning studies. Radiother Oncol 2020;153:67–78.

[10] Moustakis C, et al. Planning benchmark study for stereotactic body radiation therapy of liver metastases: results of the DEGRO/DGMP working group on stereotactic radiation therapy and radiosurgery. Int J Radiat Oncol Biol Phys 2022;113(1):214–27.

[11] Habraken SJM, et al. The TRENDY multi-center randomized trial on hepatocellular carcinoma – trial QA including automated treatment planning and benchmark-case results. Radiother Oncol 2017;125(3):507–13.

[12] Williams MJ, et al. Multicentre quality assurance of intensity-modulated radiation therapy plans: a precursor to clinical trials. Australas Radiol 2007;51(5):472–9.

[13] Blanck O, et al. Inverse treatment planning for spinal robotic radiosurgery: an international multi-institutional benchmark trial. JACMP 2016;17(3):313–30.

[14] Hardcastle N, et al. The challenge of planning vertebral body SBRT: optimizing target volume coverage. Med Dos 2020;45(3):302–7.

[15] Nelms BE, et al. Variation in external beam treatment plan quality: an inter-institutional study of planners and planning systems. Pract Radiat Oncol 2012;2(4):296–305.

[16] Lee J, et al. Multi-center evaluation of dose conformity in stereotactic body radiotherapy. Phys Imaging Radiat Oncol 2019;11:41–6.