Journal Pre-proof

A Prospective Study of Machine Learning-Assisted Radiotherapy Planning for Patients Receiving 54 Gy to the Brain

Derek S. Tsang MD MSc , Grace Tsui BSc , Anna T. Santiago MSc MPH , Harald Keller PhD , Thomas Purdie PhD FAAPM , Chris Mcintosh PhD , Glenn Bauman MD , Nancy La Macchia , Amy Parent , Hitesh Dama BSc , Sameera Ahmed MSc , Normand Laperriere MD , Barbara-Ann Millar MBChB , Valerie Liu BSc , David C. Hodgson MD MPH FASTRO

PII: S0360-3016(24)00334-1
DOI: https://doi.org/10.1016/j.ijrobp.2024.02.022
Reference: ROB 28713

To appear in: *International Journal of Radiation Oncology, Biology, Physics*

Received date: 21 August 2023
Revised date: 11 January 2024
Accepted date: 10 February 2024

Please cite this article as: Derek S. Tsang MD MSc , Grace Tsui BSc , Anna T. Santiago MSc MPH , Harald Keller PhD , Thomas Purdie PhD FAAPM , Chris Mcintosh PhD , Glenn Bauman MD , Nancy La Macchia , Amy Parent , Hitesh Dama BSc , Sameera Ahmed MSc , Normand Laperriere MD , Barbara-Ann Millar MBChB , Valerie Liu BSc , David C. Hodgson MD MPH FASTRO , A Prospective Study of Machine Learning-Assisted Radiotherapy Planning for Patients Receiving 54 Gy to the Brain, *International Journal of Radiation Oncology, Biology, Physics* (2024), doi: https://doi.org/10.1016/j.ijrobp.2024.02.022

A Prospective Study of Machine Learning-Assisted Radiotherapy Planning for Patients Receiving 54 Gy to the Brain

Derek S. Tsang MD MSc[a], Grace Tsui BSc[a], Anna T. Santiago MSc MPH[b], Harald Keller PhD[a], Thomas Purdie PhD FAAPM[a,c], Chris Mcintosh PhD[a,c], Glenn Bauman MD[d], Nancy La Macchia[a], Amy Parent[a], Hitesh Dama BSc[a], Sameera Ahmed MSc[a], Normand Laperriere MD[a], Barbara-Ann Millar MBChB[a], Valerie Liu BSc[a], David C. Hodgson MD MPH FASTRO[a]

[a] Radiation Medicine Program, Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada
[b] Department of Biostatistics, University Health Network, Toronto, Ontario, Canada
[c] Techna Institute, University Health Network, Toronto, Ontario, Canada
[d] London Regional Cancer Program, London, Ontario, Canada

Running title: ML planning for brain tumors

Conflicts of interest: DST received travel funding from Mevion Medical Systems and Elekta AB, and is a consultant with Need, all unrelated to this work. TP reports grants from CIHR and NSERC, unrelated to this work. TP and CM receive royalties from RaySearch Laboratories, unrelated to this work. CM holds a patent, "Method And System for Automated Quality Assurance And Automated Treatment Planning in Radiation Therapy". DCH serves as Medical Director at the Pediatric Oncology Group of Ontario, unrelated to this work.

Funding: This research was supported by the Mount Sinai Hospital-University Health Network Academic Medical Organization Innovation Fund.

Data sharing statement: Data and project code are stored in an institutional repository and are available upon request to the corresponding author.

Corresponding author:
Derek S. Tsang
610 University Ave, OPG 7-316
Toronto ON M5G 2M9
(tel) 416-946-2121
(fax) 416-946-2111
(e-mail) derek.tsang@uhn.ca

Abstract

Purpose: The capacity for machine learning (ML) to facilitate radiotherapy (RT) planning for primary brain tumors has not been described. We evaluated ML-assisted RT planning with regards to clinical acceptability, dosimetric outcomes and planning efficiency for adults and children with primary brain tumours.

Methods and Materials: In this prospective study, children and adults receiving 54 Gy fractionated radiotherapy for a primary brain tumor were enrolled. For each patient, one ML-assisted RT plan was created and compared with one or two plans created using standard

("manual") planning procedures. Plans were evaluated by the treating oncologist, who was blinded to the method of plan creation. The primary endpoint was the proportion of ML plans that were clinically acceptable for treatment. Secondary endpoints included the frequency with which ML plans were selected as preferable for treatment, and dosimetric differences between ML and manual plans.

Results: A total of 116 manual plans and 61 ML plans were evaluated across 61 patients. Ninety-four percent of ML plans and 93% of manual plans were judged to be clinically acceptable ($p = 1.0$). Overall, the quality of ML plans were similar to manual plans. ML plans comprised 34.5% of all plans evaluated, and were selected for treatment in 36.1% of cases ($p = 0.82$). Similar tumor target coverage was achieved between both planning methods. Normal brain (brain minus PTV) received an average of 1 Gy less mean dose with ML plans (as compared to manual plans, $p < 0.001$). ML plans required an average of 45.8 minutes less time to create, as compared with manual plans ($p < 0.001$).

Conclusion: ML-assisted automated planning creates high-quality plans for patients with brain tumours, including children. Plans created with ML assistance delivered slightly less dose to normal brain tissues, and can be designed in less time.

Keywords
Brain neoplasms; Machine learning; Radiotherapy planning

Introduction

Radiotherapy (RT) is a curative-intent treatment used for many children and adults with primary brain tumors. A high-quality RT plan delivers a focused dose of radiation to the tumor and surrounding at-risk regions, while minimizing radiation to normal brain tissues. This is particularly crucial for young patients, in whom excess RT to normal brain can lead to significant long-term side effects including cognitive changes [1, 2], hearing loss [3, 4] and endocrinopathy [5].

The standard manual method of RT planning is a time-consuming iterative process of testing beam/arc angles and optimization parameters by trained radiation therapists (also known as dosimetrists or planners). This process can be associated with substantial variation in plan quality, depending on the experience of the dosimetrist and oncology staff. If a less-than-optimal

plan is created, then a given patient may receive a higher-than-needed dose to specific organs-at-risk (OARs).[6, 7] Machine learning (ML) methods allow for automated processes to design, develop and deliver fully optimized RT plans that can improve the quality of RT plans, as well as reduce the time needed to create such plans.[8] Our group previously trained a ML radiotherapy model for children and adults with brain tumors; this novel method created RT plans that reduced doses to normal brain structures, while requiring less time than conventional RT planning to generate.[9]

The objective of study was to prospectively evaluate ML plans for children and adult patients receiving conventionally fractionated RT (54 Gy) for a primary brain tumor through blinded comparison with plans generated through a standard, manual planning process. We hypothesized that ≥90% of ML plans would be deemed clinically acceptable by expert radiation oncologist evaluation.

Methods and Materials

Patients

From January 2022 through May 2023, patients planned to receive conformal (focal) 54 Gy radiotherapy for an intracranial tumour were eligible to be prospectively enrolled and underwent concurrent manual and ML planning at a single institution. Ineligible patients were individuals with extracranial tumors or target volumes extending caudal to foramen magnum, or had emergency RT planning. There was no restriction on patient age or prior courses of RT. The target sample size was estimated using Simon's two-stage design, with the criterion for success being ≥90% of ML plans judged as clinically acceptable, while a clinical acceptability rate of

<80% would be unsatisfactory. With a power of 80% and type 1 error rate (one-sided) of 0.05, a minimum of 57 evaluable patients needed to be evaluated to distinguish these rates.

Study procedures

Enrolled patients first underwent standard-of-care simulation with non-contrast CT in a thermoplastic mask and MR simulation with gadolinium contrast. Images were imported into the clinical treatment planning system (TPS; RayStation 8B or 10B, RaySearch Laboratories, Stockholm, Sweden) for oncologist contours. Up to three plans were created based on the 54 Gy prescription: 1-2 manual and 1 machine-learning. Manual plans were created by trained dosimetrists in RayStation 8B or 10B; it is standard practice at our institution to create at least 1 or more plans for radiation oncologist review and choice, typically with differing beam geometry and/or optimization parameters (i.e. to focus on target coverage vs. OAR sparing). In a small number of patients, all clinical objectives could be met, and only one manual plan was created. Machine learning plans were created in a separate instance of RayStation 8B (which avoided biasing the dosimetrist creating manual RT plans) based on an existing, previously created ML model that was trained on high-quality 54 Gy plans in children and adults.[9] Briefly, this ML model was trained using atlas regression forests to associate image features with expected radiation dose within an atlas of 95 high-quality 54 Gy intracranial RT plans from two academic institutions. The most common tumor type in the training set was glioma (65% of all cases); among all cases, there was representation of midline (n = 45), left lateralized left (n = 18) and right lateralized (n = 32) tumors. When a novel case was presented for ML planning, a conditional random field model was used to generate a predicted dose distribution; that is, what an idealized dose distribution should look like (high doses to PTV, low doses to OARs).

Predicted dose plans were then converted into clinically deliverable volumetric modulated arc therapy (VMAT) plans using an inverse-planning optimization algorithm in RayStation 8B that mimicked predicted dose and minimized differences between the predicted and final dose, while ensuring technical beam delivery constraints were met to create a deliverable VMAT plan.

After the creation of the manual and ML plans, the ML plan was copied back into the clinical instance of RayStation (version 8B or 10B) and all plans were de-identified with removal of beam geometry and optimization parameters, so that it was not possible to determine whether a given plan was created manually or using ML. The treating radiation oncologist reviewed the blinded plans, identified which plans were clinically acceptable, selected a preferred plan, and completed a form which collected subjective, qualitative information about all three plans (Supplementary Material 1). After a plan was selected for clinical treatment, the dosimetrist then unblinded the plans, submitted the chosen plan for quality assurance (QA) by a medical physicist and a second dosimetrist. The plan also underwent patient-specific QA using a phantom (ArcCHECK, Sun Nuclear, Melbourne, FL) as per institutional standard operating procedures for VMAT plans and submitted to the treating radiation oncologist for final approval, prior to commencement of the first treatment fraction.

Data analysis

The primary endpoint was the proportion of ML plans that were judged to be clinically acceptable, defined as safe and suitable for treatment; this was a binary endpoint. Secondary endpoints included the proportion of cases in which the ML plan was used as the basis for treatment, time required to create ML and manual plans (defined as time between initiation of

planning by the dosimetrist to when RT plan(s) are ready to present to the radiation oncologist for review, as recorded by the dosimetrist), and evaluation of doses to OARs between planning methods (both quantitatively and by subjective oncologist evaluation). The observed probability of ML plan acceptability was compared with expected probability using a two-sided z-test. Times to plan manual vs. ML cases were compared using a two-sided paired t-test. A two-sided Student's t-test was used to compare times to plan patients enrolled in the 1^{st} vs. 2^{nd} half of the study. Dosimetric parameters for manual plans were averaged if two manual plans were available. Paired t-tests were used to compare dose data between manual and ML treatment plans. Multiple testing adjustment of paired t-tests were done using the Benjamini-Hochberg method to generate q-values (adjusted p-values) with a false discovery rate (Q) of 0.05. Statistical analysis was performed using R version 4.2.2 on Windows.

The project was reviewed and approved by the hospital research ethics board (REB) and the hospital Quality Improvement Review Committee (QIRC) as a quality improvement (QI) initiative, with waiver of patient or family consent. Details regarding staff participating in the project, as well as the QI study protocol and CONSORT-AI checklist are provided in Supplementary Material 2.

Results

A total of 61 patients were accrued. Baseline characteristics of patients are presented in Table 1. Five and nine of 61 patients were pediatric, defined as age <18 and age <21, respectively. An example of three RT plans created for a patient is shown in Figure 1.

Plan acceptability

A total of 116 manual plans and 61 ML plans were presented for physician evaluation. Six patients had one manual and one ML plan created; the remaining 55 patients had two manual and one ML plan created. A total of 61 of 177 (34.5%) of plans evaluated by oncologists were ML plans. Overall, 93% of ML plans (54 of 58) and 94% of manual plans (103 of 110) were judged to be clinically acceptable ($p = 1.0$, Table 2). The proportion of cases in which the ML plan was selected for treatment was not significantly different from the overall proportion of ML plans evaluated (i.e. 34.5% of plans evaluated were ML plans, and 36.1% of plans selected for treatment were ML plans; $p = 0.82$; Table 3). All ML plans selected for treatment were able to pass patient-specific QA without additional modification. Subjective oncologist evaluations of manual vs ML plans are presented in Table 2. Fifty-eight oncologist evaluations were available for analysis; three patients' plans did not have detailed subjective ratings recorded. Plans generated using the ML method were rated highly for OAR avoidance, with similar target coverage as compared to manual plans. Qualitative comments were provided by the reviewing oncologist for 17 cases and are shown in Supplementary Material 3.

In the subgroup of nine pediatric patients age <21 at the time of RT start, a total of 9 ML plans and 17 manual plans were presented for oncologist selection. ML plans were chosen for treatment in 3 of 9 patients (33.3%), as compared to an expected probability of ML selection of 34.6%.

Quantitative dosimetry

Quantitative evaluation of RT dosimetry is presented in Table 4. There was no statistically significant difference in GTV, CTV or PTV coverage; ML plans delivered slightly higher maximum dose to the PTV by 40 cGy ($p < 0.001$). Machine learning plans were clinically equivalent for target coverage, but were better able to avoid normal brain tissues (with a 101 cGy reduction in mean dose to the "brain minus PTV" structure, $p < 0.001$). Doses to the parotids and spinal cord were significantly lower with ML planning, while doses to the lenses were slightly higher with ML planning. An exploratory analysis separately comparing ML plans with the 1st or 2nd manual plan generated is presented in Supplementary Material 5; statistically significant differences to PTV maximum dose, normal brain, spinal cord and lenses were retained.

The time to create RT plans was compared between manual vs. ML methods is shown in Figure 2. The mean time savings is 45.8 minutes ($p < 0.001$), with less time required for ML plans. Machine learning cases planned in the first half of the study (first 30 patients) required a mean of 69.3 minutes to plan, while ML cases planned in the second half of the study (subsequent 31 patients) required a mean of 43.4 minutes to plan ($p < 0.001$; Supplementary Material 4). Prior to unblinding, oncologists were also asked if they could identify which of the RT plans they were presented with were designed using the ML method; 28 of 58 guesses were correct (51.7%), whereas if guessing was done at random, one would expect a correct rate of ML plan identification of 34.5% ($p = 0.084$).

Discussion

To our knowledge, this is the first prospective study to evaluate the performance of ML-assisted RT planning for patients with brain tumors, and the first to include pediatric patients. In over

90% of cases, ML was able to create high-quality RT plans that, in blinded comparisons, were of equivalent quality to plans created by specialized RT planners. ML plans produced small reductions in mean dose to the normal brain by 101 cGy, were slightly hotter within the PTV (40 cGy) and with higher conformity of the 50% (2700 cGy) isodose line, at a cost of slightly increased lens maximum doses. Creation of ML plans also required less time than manual planning; a potential training effect was observed over time among dosimetrists operating the TPS, with patients enrolled in the second half of the study requiring less time to generate an ML plan as compared to patients enrolled in the first half of the study.

There is potential in this ML platform to consistently create RT plans that are acceptable for clinical treatment (by subjective oncologist evaluation), while improving planning efficiency and potentially reducing turn-around time to plan design and approval. Qualitative evaluation by treating oncologists demonstrates consistent ability of ML plans to achieve tumor target coverage, which is of primary importance to ensure efficacy of ML planning. Other metrics relating to safety of ML plans (OAR sparing, high-dose conformity and dose fall-off) were competitive (though not superior) to manual planning. With this knowledge, machine learning RT planning can facilitate improved access to high quality RT care and provide important education opportunities, particularly in regions where availability of skilled radiation therapists is limited.[10] There are some instances where ML planning is used as an effective quality assurance tool to ensure that manual plans are of satisfactory quality.[11]

A strength of our study is the inclusion of patients with brain targets across different neuroanatomical regions. In some studies that use artificial intelligence or ML to assist RT

planning, treated volumes and targets were anatomically homogeneous, as with breast or prostate treatments.[12-14] However, our trained model [9] included patients with tumors from all regions of the brain, which allowed it to successfully create deliverable RT plans for any location of intracranial brain tumor.

This is one of the first known studies to prospectively evaluate ML planning for pediatric and adult brain tumors. Many other studies have evaluated artificial intelligence or ML systems to assist planning breast [15], thoracic [15], genitourinary [8], head-and-neck [16, 17], cervical [18] tumors. What is unique about our study is that we prospectively evaluated acceptability of ML plans, and proceeded to deliver the ML plans in clinical treatment of children and adults with brain tumors with RT. The project was able to implement ML planning in a real-world clinical setting with dosimetric improvement in plans delivered to patients, as well as potential resource savings with respect to dosimetrist time.

A limitation of our study is the possibility of bias for or against ML planning. Dosimetrists were requested not to use the ML plan to improve their manual plan; the ability of the ML plan to affect manual planning was reduced by use of a different treatment planning system database to design ML plans (as compared with manual plans). There was also a weak trend (not statistically significant) in the ability of the oncologist to guess which one of the presented plans was an ML plan. Nonetheless, given our finding that the ML plan was preferred for treatment no more or less frequently than manual plans, we do not believe that evaluator bias affected our findings. Oncologist ratings of ML plan acceptability were subjective, though quantitative metrics did demonstrate slightly superior normal brain tissue sparing with ML plans; whether this mean dose

reduction of -101 cGy to the brain is clinically significant (with respect to tumor control or acute/late toxicity) could not be determined by our study. Future work should incorporate analysis of factors that are associated with ML plan acceptability and oncologist preferences during dosimetric review.

Although ML planning results in time savings once implemented, this does not consider the resources and time required to train and develop the model. In addition, although our model can be used between institutions using the same treatment planning system, it is not directly scalable to other dose prescriptions, since it was tuned and trained to perform best for 54 Gy prescriptions. For example, the ML model is unable to selectively underdose a 60 Gy PTV that overlaps with optic structures (to meet optic tolerances of 54 Gy, for example), while ensuring dosimetric coverage of the remainder of the target volume to 60 Gy. Patients with tumor targets that extended into the orbit, inferior to the foramen magnum, or outside the bony skull base were excluded because the ML model could not create acceptable plans for these individuals, since patients with extracranial targets were not included in the initial training set for the ML model. Future work is planned to update the model to use newer AI algorithms (U-Net, a deep learning algorithm) [19, 20] to further improve RT planning and to train additional models for commonly used primary CNS dose prescriptions for glioblastoma, such as 40 Gy in 15 fractions and 60 Gy in 30 fractions.[21, 22] In addition, future ML models should ideally be portable between treatment planning systems to maximize generalizability.

Conclusion

In this prospective study, we demonstrate successful implementation and clinical delivery of radiotherapy plans designed using an automated, ML-enabled workflow. Plans created using ML were selected for clinical treatment as frequently as manual plans by treating radiation oncologists, and were deemed acceptable for treatment in >90% of patients. ML plans led to statistically significantly lower mean doses to the brain, with increased conformality of the 50% isodose line to the PTV, as compared to manual plans. ML planning also required less time to create as compared with manual plans. Therefore, ML-assisted RT planning may be considered alongside standard manual planning as a treatment option for patients receiving brain RT to a prescription of 54 Gy, but should be further evaluated to ensure reproducibility and generalizability in a multi-institutional setting.

References

1. Tsang, D.S., et al., *Intellectual changes after radiation for children with brain tumors: which brain structures are most important?* Neuro Oncol, 2021. **23**(3): p. 487-497.
2. Acharya, S., et al., *Association between hippocampal dose and memory in survivors of childhood or adolescent low-grade glioma: a 10-year neurocognitive longitudinal study.* Neuro Oncol, 2019.
3. Halperin, E.C., et al., *Pediatric Radiation Oncology*. 2012: Wolters Kluwer Health.
4. Keilty, D., et al., *Hearing Loss After Radiation and Chemotherapy for CNS and Head-and-Neck Tumors in Children.* J Clin Oncol, 2021: p. JCO2100899.
5. Vatner, R.E., et al., *Endocrine Deficiency As a Function of Radiation Dose to the Hypothalamus and Pituitary in Pediatric and Young Adult Patients With Brain Tumors.* J Clin Oncol, 2018. **36**(28): p. 2854-2862.
6. Nelms, B.E., et al., *Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems.* Pract Radiat Oncol, 2012. **2**(4): p. 296-305.
7. Moore, K.L., et al., *Quantifying Unnecessary Normal Tissue Complication Risks due to Suboptimal Planning: A Secondary Study of RTOG 0126.* Int J Radiat Oncol Biol Phys, 2015. **92**(2): p. 228-35.
8. McIntosh, C., et al., *Clinical integration of machine learning for curative-intent radiation treatment of patients with prostate cancer.* Nat Med, 2021. **27**(6): p. 999-1005.
9. Tsang, D.S., et al., *A pilot study of machine-learning based automated planning for primary brain tumours.* Radiat Oncol, 2022. **17**(1): p. 3.
10. McGinnis, G.J., et al., *Barriers and Facilitators of Implementing Automated Radiotherapy Planning: A Multisite Survey of Low- and Middle-Income Country Radiation Oncology Providers.* JCO Global Oncology, 2022(8).
11. Cao, W., et al., *Knowledge-based planning for the radiation therapy treatment plan quality assurance for patients with head and neck cancer.* J Appl Clin Med Phys, 2022. **23**(6): p. e13614.

12. McIntosh, C. and T.G. Purdie, *Voxel-based dose prediction with multi-patient atlas selection for automated radiotherapy treatment planning.* Phys Med Biol, 2017. **62**(2): p. 415-431.
13. McIntosh, C., et al., *Fully automated treatment planning for head and neck radiotherapy using a voxel-based dose prediction and dose mimicking method.* Phys Med Biol, 2017. **62**(15): p. 5926-5944.
14. McIntosh, C. and T.G. Purdie, *Contextual Atlas Regression Forests: Multiple-Atlas-Based Automated Dose Prediction in Radiation Therapy.* IEEE Trans Med Imaging, 2016. **35**(4): p. 1000-12.
15. Visak, J., et al., *An Automated knowledge-based planning routine for stereotactic body radiotherapy of peripheral lung tumors via DCA-based volumetric modulated arc therapy.* J Appl Clin Med Phys, 2021. **22**(1): p. 109-116.
16. Bai, P., et al., *A knowledge-based intensity-modulated radiation therapy treatment planning technique for locally advanced nasopharyngeal carcinoma radiotherapy.* Radiat Oncol, 2020. **15**(1): p. 188.
17. Jaworski, E.M., et al., *Development and Clinical Implementation of an Automated Virtual Integrative Planner for Radiation Therapy of Head and Neck Cancer.* Adv Radiat Oncol, 2023. **8**(2): p. 101029.
18. Wang, H., et al., *Tree-based exploration of the optimization objectives for automatic cervical cancer IMRT treatment planning.* Br J Radiol, 2021. **94**(1123): p. 20210214.
19. Chandran, L.P., et al., *A 3D U-Net based two stage deep learning framework for predicting dose distributions in radiation treatment planning.* International Journal of Imaging Systems and Technology, 2023.
20. Nguyen, D., et al., *3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture.* Physics in Medicine & Biology, 2019. **64**(6).
21. Perry, J.R., et al., *Short-Course Radiation plus Temozolomide in Elderly Patients with Glioblastoma.* N Engl J Med, 2017. **376**(11): p. 1027-1037.
22. Stupp, R., et al., *Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial.* Lancet Oncol, 2009. **10**(5): p. 459-66.

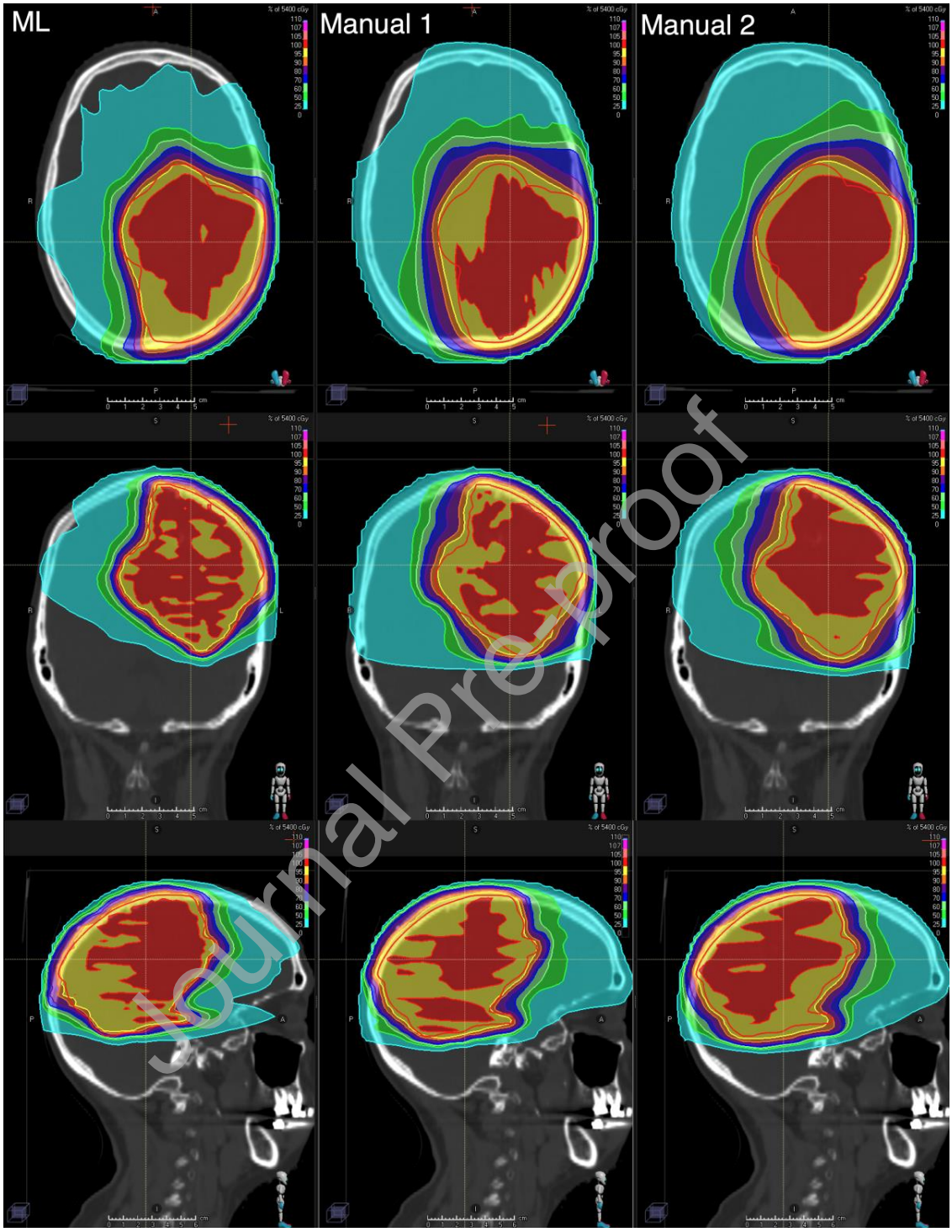

Figure 1. Screenshots of the ML (left column), manual (middle and right columns) plans created for a patient in the present study. Axial (top row), coronal (middle row) and sagittal (bottom row)

views are presented. The colorwash represents RT dose, while the PTV is shown in a thin red line. The moderate isodose lines (in shades of green) and low isodose line (in teal) are slightly more conformal to the target in the ML plan, as compared to the manual plans. The ML plan was chosen for clinical treatment.

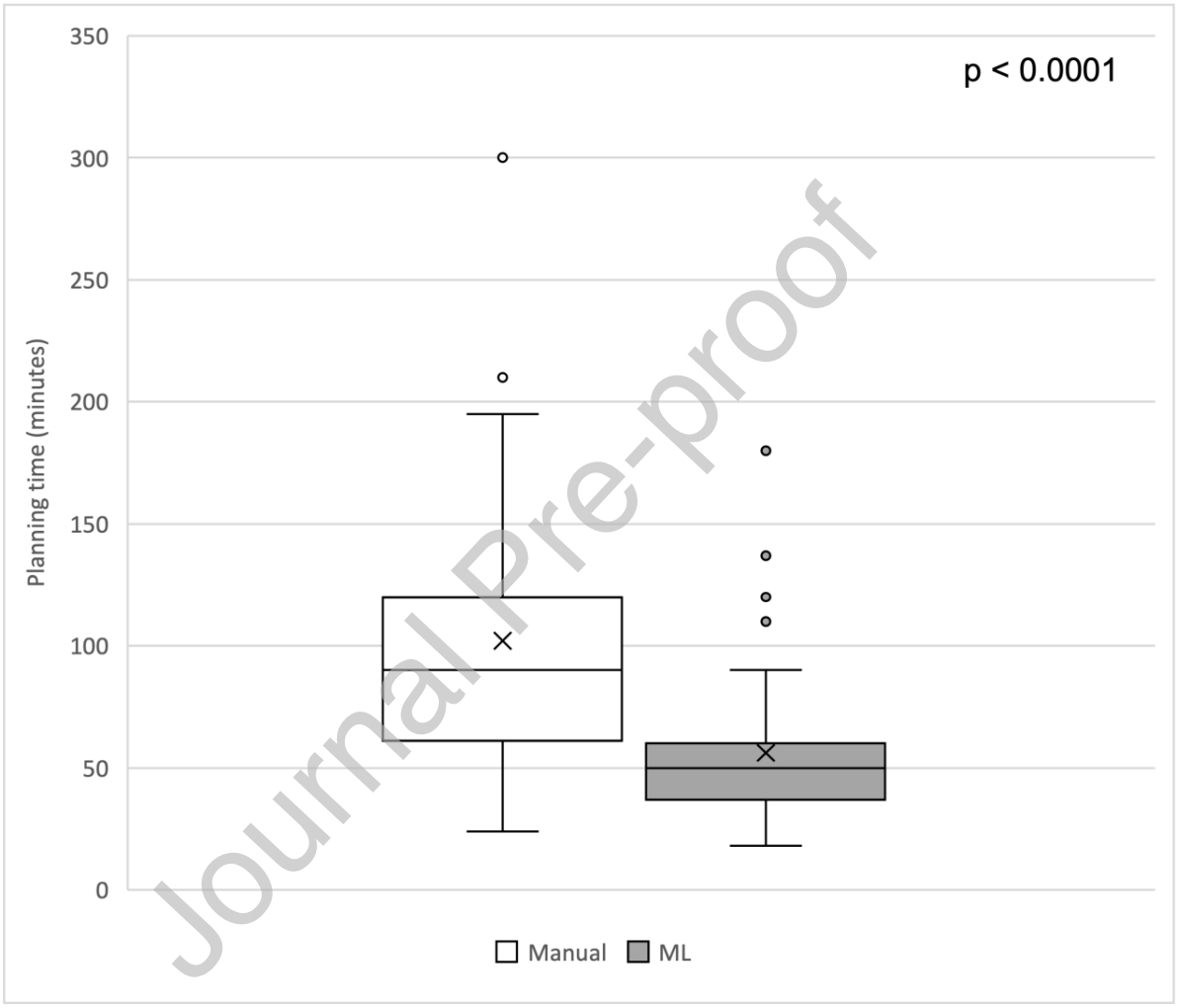

Figure 2. Boxplot comparison of time required (minutes) to create manual vs. machine learning (ML) plans. The upper/lower box represents the 1st and 3rd quartiles, the horizontal line represents the median, X represents the mean value, whiskers represent 1.5 times the interquartile range, and open circles represent outliers.

Table 1. Baseline characteristics

Characteristics	**N = 61**
Age at RT, years, median (range)	52 (3-81)
Pediatric, age <18 (%)	5 (8%)
Pediatric, age <21 (%)	9 (15%)
Female (%)	32 (52%)
Diagnosis	
Meningioma	24 (39%)
Glioma	23 (38%)
Craniopharyngioma	8 (13%)
Ependymoma	3 (5%)
Neurocytoma	2 (3%)
Atypical teratoid rhabdoid tumor, localized	1 (2%)
Target location	
Supratentorial	54 (89%)
Infratentorial	7 (11%)
Intracranial location	
Cavernous sinus	13 (21%)
Frontal	13 (21%)
Sellar, suprasellar, optic pathway	13 (21%)
Parietal	6 (10%)
Temporal	4 (7%)
Brainstem	3 (5%)
Cerebellar	3 (5%)
Occipital	3 (5%)
Cerebellopontine angle	2 (3%)
Cerebrum, multifocal	1 (2%)
Target laterality	
Left	17 (28%)
Right	17 (28%)
Midline or Bilateral	27 (44%)
GTV, cc, median (range)	16.2 (0.4-329.3)
PTV, cc, median (range)	61.3 (3.9-879.4)

Table 2. Oncologist ratings of blinded RT plans.

	Manual	**ML**	**p-value**	
Rated as acceptable for treatment (n = 168 plans)	103/110 (94%)	54/58 (93%)	1.0	
Subjective evaluation criterion (n = 58 patients)	**Target coverage**	**OAR sparing**	**High dose conformity**	**Dose fall-off**
ML plan superior to one or both	9 (16%)	24 (41%)	16 (28%)	22 (38%)

manual plans				
ML plan equivalent to both manual plans	42 (72%)	12 (21%)	24 (41%)	16 (28%)
Manual plan superior	7 (12%)	22 (39%)	18 (31%)	20 (34%)

Table 3. Characteristics of ML and manual plans presented for oncologist evaluation. In almost all patients, one ML plan was compared to two manual plans. The p-value for observed (vs expected) preference for ML plans was p = 0.82.

	Manual	**ML**
Number of plans presented for evaluation	116 (65.5%)	61 (34.5%)
Selected by oncologist for treatment	39	22
% preferred by oncologist	63.9%	36.1%

Table 4. Comparison of doses to organs-at-risk between manual and ML plans. Positive difference values indicate higher value for ML (vs. manual), whereas negative difference values indicate lower value for ML (vs. manual). Mean values are presented; standard deviations (SD) are shown in parentheses.

Structure	**Metric**	**Manual (SD)**	**ML (SD)**	**Mean Difference**	**p**	**q**
GTV	V95 (%)	100 (0.1)	100 (0.0)	0.00	0.67	0.72
CTV	V95 (%)	99.9 (0.2)	100 (0.1)	+0.05	0.09	0.15
PTV	V95 (%)	98.4 (1.4)	98.0 (1.3)	-0.42	0.07	0.13
PTV	Max (cGy)	5566 (71)	5606 (62)	+40.2	<0.001	<0.001
PTV	Conformity index (5130 cGy)	0.82 (0.11)	0.83 (0.11)	+0.010	0.38	0.45
PTV	Conformity index (4860 cGy)	0.66 (0.12)	0.67 (0.13)	+0.018	0.048	0.10
PTV	Conformity index (2700 cGy)	0.26 (0.10)	0.27 (0.11)	+0.012	0.005	0.014
Whole brain	Mean (cGy)	1537 (1057)	1463 (960)	-73.5	<0.001	0.001
Brain minus PTV	Mean (cGy)	1221 (738)	1120 (593)	-101.4	<0.001	<0.001
Brain minus PTV	V27Gy (%)	16.0 (14.2)	13.6 (10.9)	-2.4	<0.001	<0.001
Cochlea left	Mean (cGy)	1622 (1655)	1685 (1687)	+62.3	0.39	0.45
Cochlea right	Mean (cGy)	1644 (1640)	1720 (1695)	+76.0	0.24	0.33
Hippocampus left	Mean (cGy)	2078 (1522)	2084 (1502)	+6.2	0.90	0.93
Hippocampus right	Mean (cGy)	2111 (1286)	2115 (1264)	+3.3	0.96	0.96
Pituitary	Mean (cGy)	3431 (1916)	3279 (2020)	-151.8	0.09	0.15
Hypothalamus	Mean (cGy)	2910 (1933)	2845 (1997)	-64.6	0.26	0.35
Parotid left	Mean (cGy)	174 (278)	143 (277)	-30.6	0.007	0.016
Parotid right	Mean (cGy)	161 (244)	131 (216)	-29.2	0.007	0.016
Body	Max (cGy)	5566 (70.5)	5606 (62.2)	+40.2	<0.001	<0.001

Spinal cord	Max (cGy)	788 (1430)	623 (1455)	-164.7	<0.001	<0.001
Lens left	Max (cGy)	466 (161)	585 (237)	+118.9	<0.001	<0.001
Lens right	Max (cGy)	476 (158)	583 (252)	+106.8	<0.001	<0.001
Optic chiasm	Max (cGy)	4340 (1700)	4314 (1752)	-26.6	0.44	0.49
Optic nerve left	Max (cGy)	3695 (1921)	3599 (1900)	-96.3	0.10	0.15
Optic nerve right	Max (cGy)	3776 (1845)	3670 (1843)	-106.0	0.06	0.13
Brainstem	Max (cGy)	4745 (1297)	4792 (1258)	46.1	0.29	0.36