Results

Between November 2015 and July 2023, 508 eligible patients were recruited from 40 centres in UK, Germany, France, USA, Australia. 313 (62%) underwent TORS, and 195 (38%) underwent TLM. There was no significant difference in length of hospital stay after surgery between patients undergoing TORS and TLM, when centre was taken in consideration (HR=0.89, 95%CI 0.69-1.16, p=0.401). NGT insertion rates were significantly higher after TORS than TLM (85/189 – 45.0% vs 10/126 - 7.9%, respectively, OR=4.41, 95%CI=1.01-19.3, p=0.049) but there was no difference in duration (median 5 (95% CI=0.5-12) days TLM, 6 (95% CI=4-6) days TORS; HR=1.05, 95%CI=0.52-2.12, p=0.897). Mean scores significantly favoured TLM (relative to TORS) in all MDADI domains and the H&N35 swallowing item at 4 weeks post-surgery (see table 1); between group difference (95% CI): MDADI composite -4.89 (8.27,-1.50), p=0.005; MDADI physical -6.37 (-10.15, -2.59), p=0.001; MDADI global -10.02 (-16.50, -3.54), p=0.002; H&N35 swallowing 7.24 (2.17, 12.30), p=0.005. There was a trend (p<0.1) for difference in EORTC H&N 35 pain score (4.58, 95%CI(-0.90, 9.96), p=0.095) and water swallow capacity (mL/second) (-1.51, 95%CI(-3.11, 0.10), p=0.067) favouring TLM. There was no significant difference between the following scores: EORTC C30 global, constipation, and summary; H&N35 opening mouth, pain killers, and weight loss.

Conclusion

PATHOS presents a unique opportunity to compare two different transoral surgical techniques. In this study population, TORS was associated with significantly higher rates of NGT use, worse H&N35 swallowing scores, and worse MDADI scores at 4 weeks post-surgery compared to TLM. There was also a trend (p<0.1) favouring TLM in H&N35 pain score and water swallow capacity. This is the largest comparative study of functional outcomes following TORS vs TLM. The recruiting institutions' practices are likely to impact on length of stay and NGT use and has been accounted for in the analysis. This represents a non-randomised, unpowered sub-study for multiple secondary endpoints across which multiplicity was unadjusted. As such the results should be seen as hypothesis generating rather than confirmatory. Furthermore, the study has focused solely on the post-operative recovery period following surgery. It cannot comment on the impact of surgical philosophy (TORS vs TLM) on margins and how this may relate to the PATHOS randomised groups.

Head and neck surgical oncologists may wish to reconsider the role that laser surgery, both as an energy source and a philosophy, has in the emerging field of robotic surgery.

Funded by Cancer Research UK (Grant no: A25317), co-sponsored by Cardiff University and Velindre University NHS Trust.

122

Objective evaluation of plan quality in the PATHOS clinical trial using automated treatment planning

Salvatore Berenato[1], Mererid Evans[2,3], Richard Webster[2], Nachi Palaniappan[2], Lisette Nixon[4], Emma Higgins[2], Rush Patel[5], Anthony E Millin[1], Christian Hurt[6], Christie Heiberg[4], Joanna Canham[4], Terence M Jones[7], Elizabeth Miles[5], Matthew Beasley[8], Philip A Wheeler[1]

[1]Velindre Cancer Centre, Radiotherapy Physics Department, Cardiff, United Kingdom. [2]Velindre Cancer Centre, Medical Directorate, Cardiff, United Kingdom. [3]Cardiff University, Division of Cancer and Genetics, School of Medicine, Cardiff, United Kingdom. [4]Cardiff University, Centre for Trials Research, Cardiff, United Kingdom. [5]Mont Vernon Cancer Centre, National Radiotherapy Trials Quality Assurance (RTTQA) Group, Northwood, United Kingdom. [6]University of Southampton, Southampton Clinical Trials

Unit, Southampton, United Kingdom. [7]University of Liverpool, Liverpool Head and Neck Centre, Liverpool, United Kingdom. [8]United Hospitals Bristol, Bristol Cancer Institute, Bristol, United Kingdom

Topic

Imaging, radiomics and artificial intelligence

Keywords

Quality Assurance, Automation, Treatment Planning

Purpose/Objective

Radiotherapy (RT) plan quality is critical in ensuring treatment efficacy. Poor quality RT can increase the risks of treatment failure, overall mortality and detrimentally impacting a patient's quality of life [1–4]. This is especially important within RT clinical trials, where standardisation of treatment plan quality is paramount. However, widespread objective quantitative assessment of plan quality within trials is not performed routinely, leading to uncertainty on the magnitude of quality variations. Automated planning enables the possibility to efficiency and objectively assess the quality of individual clinical plans (CP) through comparison with an automatically generated standardised 'baseline' plan. Utilising this innovative auditing methodology within a trial enables full quantitative characterisation of: (i) overall plan quality, (ii) potential outliers and (iii) variation solely due to planning practice. The aim of this study was to use fully automated planning to objectively assess plan quality within the Cancer Research UK funded (A25317) multi-centre international phase III trial PATHOS.

Material/Methods

337 patients enrolled in the PATHOS clinical trial before 1st July 2021 were included in this study. 55 cases were excluded due to incomplete data and 16 for calibrating the automated solution, leaving 264 patients for analysis. 219 (83%) and 45 (17%) cases were treated with unilateral (Unilat) and bilateral (Bilat) volumes respectively. Planning was performed in alignment with the PATHOS protocol, with prescriptions of Bilat66Gy, Bilat60Gy, Unilat66Gy or Unilat60Gy in 30 fractions and Unilat50Gy in 25 fractions. Automated treatment plans (AP) were generated in RayStation using a locally developed 'Protocol Based Automatic Iterative Optimization' automated planning solution [5]. CP were quantitatively compared to AP across all the PATHOS trial metrics (including: Parotid Dmean; SpinalCord/BrainStem PRV D1cc; and PTV D98%, D2% and D50%) together with conformality (CI) and homogeneity (HI) indices. Analysis was performed with data categorised in terms of prescription and also tumour laterality. Statistical significance was assessed via a two-sided Wilcoxon matched-paired signed-rank test.

Results

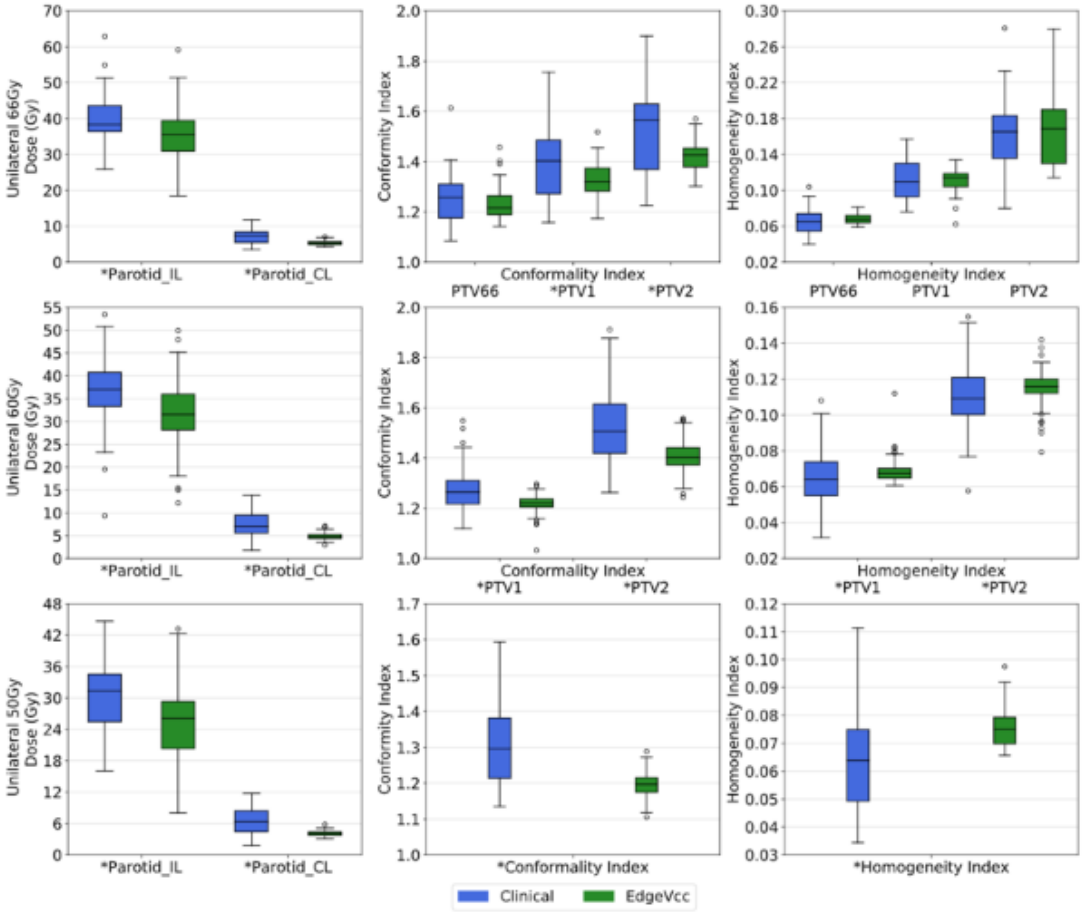

*Figure 1: The boxplots show the dosimetric results for the clinical plan (in blue) and the automated plan (in green) for unilateral cases. Significant differences (p-values ≤0.05) are highlighted with a *symbol.*

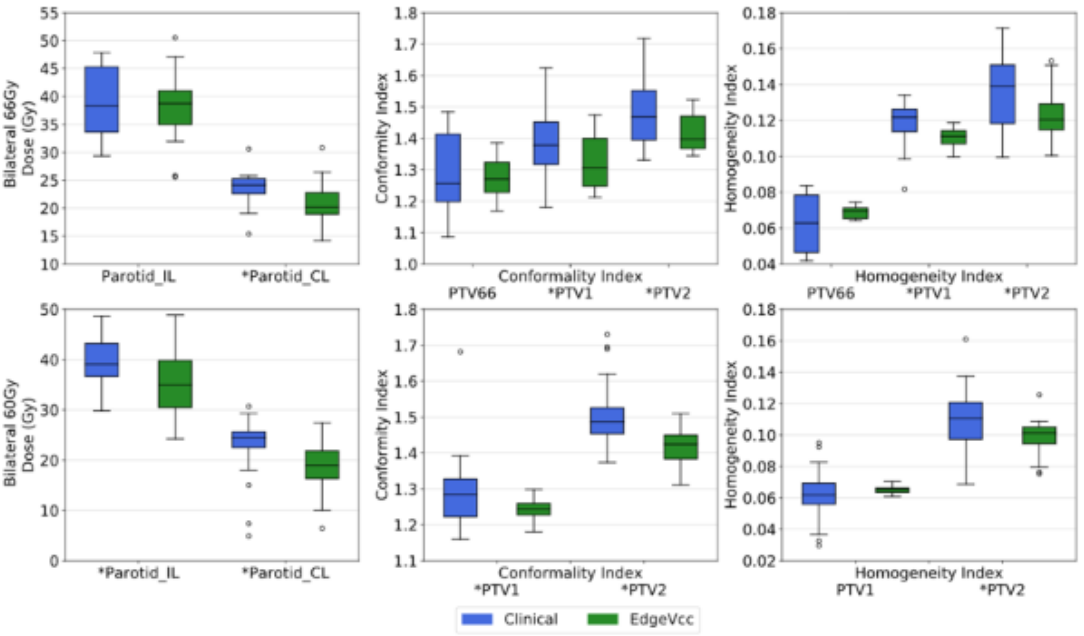

*Figure 2: Dosimetric analysis of the bilateral cases for the clinical plans (in blue) and the automated plans (in green). Significant differences (p-values ≤0.05) are highlighted with a *symbol.*

Fig. 1 and Fig. 2 present a summary of the dosimetric results, categorised in terms of prescription. When comparing CP to the AP baseline (CP-AP), statistically significant ($p \leq 0.05$) differences, Δ, in median values were observed across most key metrics. For HI, small changes across all prescriptions were detected for the primary PTV with the largest Δ equalling (-0.012, $p<0.001$) for Unilat50Gy prescriptions. This indicated CP were marginally more homogeneous that the AP baseline. For CI, significant differences were observed across primary PTVs for three prescriptions (Unilat50Gy, Unilat60Gy and Bilat60Gy) and all secondary PTVs. Median differences were substantial, with a max Δ of +0.110 ($p<0.001$, Unilat66:PTV54), which represented a 10% increase in the volume treated to 54Gy for CP. When categorised in terms of tumour laterality, differences in contralateral Parotid (Parotid_CL) Dmean were small for Unilat (Δ=+2.2Gy, $p<0.001$) and moderate for Bilat cases (Δ=+3.5Gy, $p<0.001$). For ipsilateral Parotids (Parotid_IL), differences were substantial for Unilat cases (Δ=+4.8Gy, $p<0.001$) but nominally equivalent to Parotid_CL for Bilat (Δ=+3.1Gy, $p<0.001$).

At an individual patient level, AP baseline plans highlighted potential quality improvements that could have been realised for CP. For 50% of all patients, AP led to a reduction in Parotid_IL and Parotid_CL Dmean of between 4.4Gy-14.7Gy and 2.5Gy-8.9Gy respectively. In terms of conformality, for 50% of all patients AP reduced CI by between 0.06-0.35 and 0.08-0.28 for PTV60 and PTV54 respectively.

In terms of overall variation with the trial, Fig. 1 and Fig. 2 demonstrate that a high proportion of the variation observed in the majority dose metrics was a direct result of plan quality. For example, a standardised AP planning method would have reduced the inter-quartile range (IQR) for Parotid_CL Dmean from 5.4Gy to 1.4Gy, for HI (PTV54) from 0.031 to 0.015 and for CI (PTV54) from 0.194 to 0.071. Parotid_IL Dmean was a key exception, with similar IQRs for both AP and CP.

Conclusion

Clinics participating in PATHOS undergo a comprehensive quality assurance process prior to patient recruitment, with additional 'on trial' qualitative reviews performed on small subset of patients. Furthermore, all patient plans must, where practicable, meet trial dose metric tolerances. Results of this study demonstrate that despite these procedures, which are common to many high-quality trials, meaningful variations in plan quality remain. Automated planning was found to be an effective tool in objectively assessing plan quality within a large trial. Implementation on a prospective basis could be a powerful QA tool to reduce this observed variation.

References

[1] Fairchild A, Straube W, Laurie F, Followill D. Does quality of radiation therapy predict outcomes of multicenter cooperative group trials? A literature review. Int J Radiat Oncol Biol Phys 2013; 87:246–60. https://doi.org/10.1016/j.ijrobp.2013.03.036.

[2] Ohri N, Shen X, Dicker AP, Doyle LA, Harrison AS, Showalter TN. Radiotherapy protocol deviations and clinical outcomes: A meta-analysis of cooperative group clinical trials. J Natl Cancer Inst 2013; 105:387–93. https://doi.org/10.1093/jnci/djt001.

[3] Weber DC, Tomsej M, Melidis C, Hurkmans CW. QA makes a clinical trial stronger: Evidence-based medicine in radiation therapy. Radiother Oncol 2012;105:4–8. https://doi.org/10.1016/j.radonc.2012.08.008.

[4] Peters LJ, O'Sullivan B, Giralt J, Fitzgerald TJ, Trotti A, Bernier 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 Clin Oncol 2010;28:2996–3001. https://doi.org/10.1200/JCO.2009.27.4498.

[5] Wheeler PA, Chu M, Holmes R, Smyth M, Maggs R, Spezi E, et al. Utilisation of Pareto navigation techniques to calibrate a fully automated radiotherapy treatment planning solution. Phys Imaging Radiat Oncol 2019;10:41–8. https://doi.org/10.1016/j.phro.2019.04.005.

133

Proton therapy significantly reduces acute and late toxicity in nasopharyngeal cancer

Johannes A. Langendijk[1], Tineke W.H. Meijer[1], Johanna G.M. van den Hoek[1], Stefan Both[1], Edwin Oldehinkel[1], Hans H.G. Verbeek[1], Gyorgy Halmos[2], Sjoukje F. Oosting[3], Roel H.M. Steenbakkers[1]

[1]University Medical Center Groningen, Radiation Oncology, Groningen, Netherlands. [2]University Medical Center Groningen, Otolarygology, Groningen, Netherlands. [3]University Medical Center Groningen, Medical Oncology, Groningen, Netherlands

Topic

Innovative treatments

Keywords

Nasopharyngeal carcinoma; Proton therapy; Toxicity

Purpose/Objective

The aim of the study was to test the hypothesis that Intensity Modulated Proton Therapy (IMPT) reduces acute and late radiation toxicity in nasopharyngeal cancer (NPC) patients compared to photon-based radiation techniques including IMRT and VMAT.

Material/Methods

The study population of this prospective cohort study was composed of 131 NPC patients treated with curative radiotherapy (RT) or chemoradiotherapy. Between July 2007 and December 2017, all patients were treated with IMRT or VMAT. Since January 2018, 97 out of 99 patients (98%) qualified for IMPT according to model-based selection. All patients were included in a prospective data registration program in which acute and late toxicity was prospectively scored weekly during RT and at fixed time points after RT (6 weeks, 6, 12, 18 and 24 months). To determine the overall effect on acute and late toxicity, the Weighted Overall Toxicity Burden (WOTB) was calculated, defined as the sum of all toxicities weighted by toxicity grading. In addition, the WOTB Area Under the Curve (WOTB-AUC) was calculated representing the WOTB from the start of treatment until 24 months after completion of treatment.

Results

The two groups were well balanced regarding gender, age, race, T-stage, N-stage, AJCC-stage, and EBV-status. However, there was a significant difference regarding the chemotherapy regimens used between the two groups. In the photon cohort, 33% of patients were treated with conventional RT, 7% with concurrent chemoradiation, 2% with induction chemotherapy + concurrent chemoradiation and 57% with concurrent chemoradiation + adjuvant chemotherapy, while this was 25%, 37%, 36% and 2% in the proton cohort, respectively.