BJR

https://doi.org/10.1259/bjr.20201354

Received: 19 November 2020 | Revised: 07 January 2021 | Accepted: 11 January 2021

Cite this article as:
Clemente S, Falco MD, Cagni E, Talamonti C, Boccia M, Gino E, et al. The influence of small field output factors simulated uncertainties on the calculated dose in VMAT plans for brain metastases: a multicentre study. *Br J Radiol* 2021; **94**: 20201354.

FULL PAPER

The influence of small field output factors simulated uncertainties on the calculated dose in VMAT plans for brain metastases: a multicentre study

[1]STEFANIA CLEMENTE, MSc, [2]MARIA DANIELA FALCO, PhD, [3]ELISABETTA CAGNI, PhD, [4]CINZIA TALAMONTI, PhD, [5]MAFALDA BOCCIA, MSc, [6]EVA GINO, MSc, [7]ELENA LORENZINI, MSc, [8]FEDERICA ROSICA, MSc, [9]SERENELLA RUSSO, MSc, [10]ALESSANDRO ALPARONE, MSc, [11]DANIELE ZEFIRO, MSc and [12]CHRISTIAN FIANDRA, PhD

[1]Unit of Medical Physics and Radioprotection, Federico II University Hospital, Napoli, Italy
[2]Department of Radiation Oncology, "G. D'Annunzio" University, "SS. Annunziata" Hospital, Chieti, Italy
[3]Medical Physics Unit, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy
[4]Medical Physics Unit, University Of Florence, Azienda Ospedaliero Universitaria Careggi, Florence, Italy
[5]Polydiagnostic Center Check - Up S.r.l, Salerno, Italy
[6]Medical PhysicDepartment, A.O. Ordine Mauriziano, Turin, Italy
[7]U.O.C Fisica Sanitaria Area Nord, Azienda USL Nord Ovest Toscana, Massa Carrara, Italy
[8]U.O.C. Fisica Sanitaria, ASL di Teramo, Italy
[9]Medical Physics Unit, AUSL Toscana Centro, Florence, Italy
[10]R&D Department, Tecnologie Avanzate, Turin, Italy
[11]MedicaPhysics Unit, ASL5 Sistema Sanitario Regione Liguria, La Spezia, Italy
[12]Department of Oncology, University of Turin, Turin, Italy

Address correspondence to: Dr Christian Fiandra
E-mail: *christian.fiandra@unito.it*

Objectives: This multicentric study was carried out to investigate the impact of small field output factors (OFs) inaccuracies on the calculated dose in volumetric arctherapy (VMAT) radiosurgery brain plans.

Methods: Nine centres, realised the same five VMAT plans with common planning rules and their specific clinical equipment Linac/treatment planning system commissioned with their OFs measured values (OFbaseline). In order to simulate OFs errors, two new OFs sets were generated for each centre by changing only the OFs values of the smallest field sizes (from 3.2 × 3.2 cm^2 to 1 × 1cm^2) with well-defined amounts (positive and negative). Consequently, two virtual machines for each centre were recommissioned using the new OFs and the percentage dose differences ΔD (%) between the baseline plans and the same plans recalculated using the incremented (OFup) and decremented (OFdown) values were evaluated. The ΔD (%) were analysed in terms of planning target volume (PTV) coverage and organs at risk (OARs) sparing at selected dose/volume points.

Results: The plans recalculated with OFdown sets resulted in higher variation of doses than baseline within 1.6 and 3.4% to PTVs and OARs respectively; while the plans with OFup sets resulted in lower variation within 1.3% to both PTVs and OARs. Our analysis highlights that OFs variations affect calculated dose depending on the algorithm and on the delivery mode (field jaw/MLC-defined). The Monte Carlo (MC) algorithm resulted significantly more sensitive to OFs variations than all of the other algorithms.

Conclusion: The aim of our study was to evaluate how small fields OFs inaccuracies can affect the dose calculation in VMAT brain radiosurgery treatments plans. It was observed that simulated OFs errors, return dosimetric calculation accuracies within the 3% between concurrent plans analysed in terms of percentage dose differences at selected dose/volume points of the PTV coverage and OARs sparing.

Advances in knowledge: First multicentre study involving different Planning/Linacs about undetectable errors in commissioning output factor for small fields.

INTRODUCTION

In the last ten years volumetric modulated arc therapy (VMAT) has gained popularity and intra cranial stereotactic radiosurgery VMAT treatments are routinely performed in both leading academic medical centres and in smaller, non-academic centres.[1] Stereotactic Radiosurgery (SRS) procedures use small radiation fields leading to a possible increase of dosimetric uncertainties.[2]

Accurate characterization of small fields is challenging due to the lack of lateral electronic equilibrium, the partial occlusion of the finite radiation source, the perturbation of the charged particle fluence caused by the detector, the high dose gradients and peaked dose distributions. Therefore, a good practice for small field dosimetry is to select suitable detectors and to take into account their perturbation correction factors.[3] As regards the latter, Alfonso et al introduced a specific formalism.[4] Interestingly, a new code of practice was published[5] which offered comprehensive data on all of the detectors available on the market at that time and their correction factors. According to this code of practice, after applying the appropriate correction factor, more than one detector must be used in order to verify the consistency of the measurements. Moreover, for an accurate experimental determination of small field output factors (OFs),[6,7] it is essential to verify the calculated small field OFs in order to identify potential dose discrepancies.[8]

Although the topic of small field OF measurements has been widely investigated in the literature,[9–16] few studies have assessed how inaccuracies in the measurement of small field OFs affect treatment planning system (TPS) dose calculation accuracy in clinical cases.[7,8,17,18]

The purpose of this work is to evaluate, in the context of large multi centre study, how undetectable errors in the measurement of small fields OFs values can affect the accuracy of dose calculations in VMAT brain radiosurgery treatments. To this aim, two new OFs sets were generated for each TPS/centre by varying the small field OFs values into known amounts (positive and negative). Consequently, differences between the original plans (that used the $OF_{baseline}$ set) and the same plans recalculated using two new virtual machines and configured with the incremented (OF_{up}) and decremented (OF_{down}) OFs values, were analysed.

METHODS AND MATERIALS

Nine centres, with teams of expert medical physicists capable of performing SRS plans using VMAT techniques, participated in the study. For this study, we decided to use VMAT technique only because is potentially more sensitive to the variation of small fields output factors respect to easier techniques, in terms of fields complexity, such as IMAT (conformal arc therapy) or the use of stereotaxic cones.[19] Details of the TPSs, linacs, MLCs, detectors and OFs set-up measurements are reported in Table 1.

All of the centres, measured OFs according to the IAEA TRS-398 code of practice and to type-specific TPS/linac reference conditions. Using the different next-generation detectors, small fields measurements were taken and analysed in terms of ratio of detector readings according to the formalism proposed by Alfonso.[4] Correction factors were not considered, since each TPS was commissioned before the factors were available in the literature. A good agreement between calculated and measured OFs was reported during the beams modelling by each centre, in accordance with international methodological recommendations.[20–22]

Five patients were enrolled in the study, four with single brain metastases and one with multiple metastases, in order to diversify the cases in size and position. The relative CT images and structures were converted into Dicom-RT format and shared between all of the participating centres. Of note, data uncertainties in RT structures should be taken into account as different software were used in the study. A strongly dependence was reported from the initial phantom's voxel resolution (0.1 × 0.1 × 1 mm) and the boundary locations algorithm as well.[23] Here, the maximum variation in the volumes of lesions among centres was within 5% showing that the used computational phantom's resolution was suitable to be shared for a correct reconstruction by all TPSs. Every centre developed a plan for each patient (defined here 'baseline' plan) using their clinical equipment Linac/TPS (therefore their $OF_{baseline}$ values) and common planning rules: dose prescription to the tumour and organs at risks (OARs) constraints all satisfied (Table 2), VMAT planning beam geometry (five and six non-coplanar arcs), dose calculation grid (1 mm), isocentre position and beam energy (6 MV). Patients

Table 1. Patients, lesions number and size, prescription dose, VMAT techniques, PTV and OARs dose–volume constraints for plans optimization

# Patient	# Lesions	Size of lesions	Prescription dose	VMAT techniques	Planning rules			
					PTV coverage	OARs	Constraints	
1	1	d = 8 mm, V = 0.3 cm^3	36 Gy in three fractions	Five non-coplanar arcs	$V_{100\%}$ = 95%	Brainstem	V_{18Gy}<0.5cc	$V_{23.1Gy}$ ≤0.035cc
2	1	d = 1 cm, V = 0.4 cm^3	27 Gy in three fractions	Five non-coplanar arcs		Brain	-	$V_{28.8Gy}$ <3–7cc
3	1	d = 2 cm, V = 1.5 cm^3	27 Gy in three fractions	Five non-coplanar arcs		Cord	V_{18Gy}<0.35cc	$V_{21.9Gy}$ ≤0.035cc
4	1	d = 3 cm, V = 9 cm^3	35 Gy in five fractions	Five non-coplanar arcs		Opt. Pathway	$V_{15.3Gy}$ <0.2cc	$V_{17.4Gy}$ ≤0.035cc
5	3	d_1 = 2.2 cm, V_1 = 3.3 cm^3 d_2 = 2.5 cm, V_2 = 7.7 cm^3 d_3 = 1.7 cm, V_3 = 2.8 cm^3	27 Gy in three fractions	Six non-coplanar arcs (single isocentre)				

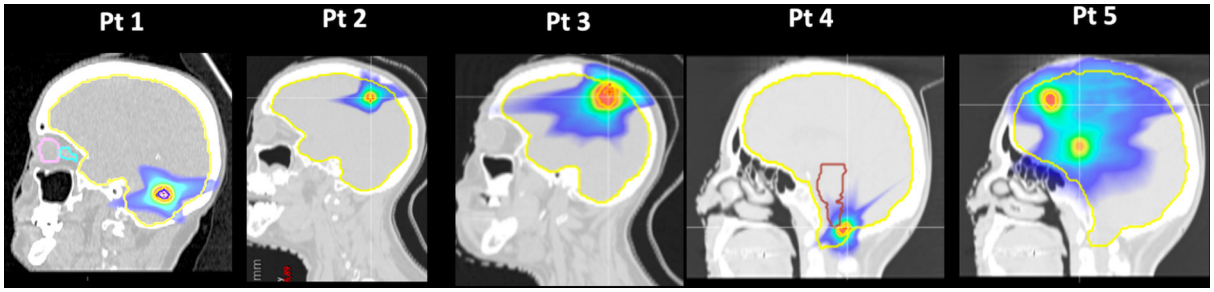
#: number d: diameter; V: volume

Table 2. TPSs, Linacs, MLCs and OFs measurements details of the enrolled centres

		TPSs		LINACs/MLCs					OFs MEASUREMENTS DETAILS			
Centres	Type	Version	Algorithm	Type	MLC	MLC leaves number/width (mm)	Maximum collimator field size (cm^2)	JTT	Set- up	Fields size definition	Detector used for small fields	Minimum field size measured & commissioned (cm^2)
1	Monaco (ElektaAB, Stockholm, Sweden)	5.11.02	MC	Synergy (Elekta)	Agility (Elekta)	160/5	40 × 40	YES	SSD = 90 cm Depth = 10 cm	MLC (transversal) Jaws (longitudinal)	CC01 (IBA)	1 × 1
2	Pinnacle (Philips Radiation Oncology Systems, Fitchburg, WI)	9.10	CCC	Synergy (Elekta)	Beam Modulator (Elekta)	80/4	16 × 21	NO	SSD = 90 cm Depth = 10 cm	MLC (jaw fixed to 16 × 21 cm)	Gaf Chromic EBT3 film	0.8 × 0.8
3	Pinnacle (Philips Radiation Oncology Systems, Fitchburg, WI)	9.10	CCC	Trilogy (Varian)	HD 120 (Varian)	120/5–2.5	40 × 22	NO	SSD = 90 cm Depth = 10 cm	MLC +Jaws	CC01 (IBA)	1 × 1
4	Eclipse (Varian Medical Systems, Palo Alto, CA)	13.7	AAA & Acuros	True Beam (Varian)	HD 120 (Varian)	120/5–2.5	40 × 22	YES	SSD = 100 cm Depth = 5 cm	MLC	Diode E (PTW)	1 × 1
5	Pinnacle (Philips Radiation Oncology Systems, Fitchburg, WI)	16	CCC	Synergy (Elekta)	Agility (Elekta)	160/5	40 × 40	YES	SSD = 90 cm Depth = 10 cm	MLC +Jaws	Microdiamond (PTW)	1 × 1
6	Ray Station (RaySearch Laboratories, Stockholm)	7.0	CCC	True Beam (Varian)	HD 120 (Varian)	120/5–2.5	40 × 22	YES	SSD = 90 cm Depth = 10 cm	Jaw	Microdiamond (PTW)	1 × 1
7	Ray Station (RaySearch Laboratories, Stockholm)	7.0	CCC	Axesse (Elekta)	Beam Modulator (Elekta)	80/4	16 × 21	NO	SSD = 90 cm Depth = 10 cm	MLC (jaw fixed to 16 × 21 cm)	TLD100, Gaf Chromic EBT3 film, micro Lion (PTW)	1 × 1
8	Ray Station (RaySearch Laboratories, Stockholm)	7.0	CCC	Synergy (Elekta)	Agility (Elekta)	80/5	40 × 40	YES	SSD = 90 cm Depth = 10 cm	MLC (transversal) Jaws (longitudinal)	EDGE Detector (Sun Nuclear)	1 × 1
9	Pinnacle (Philips Radiation Oncology Systems, Fitchburg, WI)	9.10	CCC	True Beam (Varian)	HD 120 (Varian)	120/5–2.5	40 × 22	YES	SSD = 100 cm Depth = 10 cm	MLC	Razor (IBA)	1 × 1

MLC, multileaf collimator; OF, output factor; TPS, treatment planning system.

Figure 1. Patients enrolled in the study.



and planning characteristics are reported in Figure 1 and Table 1. The immobilization equipment was carefully contoured and each Houfield unit (HU) value was converted to mass density via linear interpolation of the mass density-to-HU calibration curve clinically commissioned in each centre.

For each centre, two new sets of OFs values were generated by changing only the values of the smallest OFs field sizes with well-defined amounts (up to ±3% of the OFs $_{baseline}$). The amounts of variation were established fitting the Sauer model[24] to the incremented and decrements new OFs data.

First, the $OF_{baseline}$ set for field sizes ranging from $1 \times 1\,cm^2$ to $40 \times 40\,cm^2$, normalised to $10 \times 10\,cm^2$, was fitted using the Sauer equation reported also by Cagni et al.[9] :

$$OF\left(FS\right) = P_{\infty}\frac{FS^n}{l^n+FS^n} + S_{\infty}\left[1-\exp\left(-bFS\right)\right] \qquad .1$$

where FS is the field size and P_{∞}, S_{∞}, l, b and n are fit coefficients. More specifically, P_{∞} represents the maximum primary dose component and S_{∞} represents the maximum scatter component. The function was forced to be equal to one at the reference field (OF = 1; FS_{Ref} = 10 cm) as a boundary condition in the fit. Sets of FIT parameters were calculated using data obtained from all of the centres.

Then, for each centre, the $OF_{baseline}$ values for field sizes ranging from $3.2 \times 3.2\,cm^2$ to $1 \times 1\,cm^2$ were modified up to a maximum of ±3% (the new OF_{up} and OF_{down} sets), that is until data were still well fitted by the Sauer model (both visually and numerically (root-mean-square deviation (R^2)), in order to simulate hardly detectable errors (Table 3).

A total of 9 With Flattening Filter (WFF) 6 MV beams plus 3 Flattening Filter (FF) Free beams (available only in three centres) were collected and lastly, two new virtual treatment machines were recommissioned for each TPS/centre using the new incremented (OF_{up}) and decremented (OF_{down}) values. Maintaining the fixed number of monitor units (μ), each baseline plan, for each centre, was recalculated using the new virtual machines. The percentage dose differences, ΔD (%) defined as $\left(\frac{Dup\ or\ Ddown-Dbaseline}{Dbaseline}\right) \times 100$ between recalculated plans and the relative baseline plans were analysed to assess planning target volume (PTV) coverage and OARs sparing at selected dose/volume points: $D_{95\%}$, $D_{80\%}$, $D_{50\%}$, $D_{2\%}$and $D_{7cc\ Brain}$, $D_{mean\ Brain}$, $D_{0,1cc\ Brainstem}$, respectively. The Student's *t*-test was used to compare the results with a statistical significance level of $p \leq 0.05$ value after testing for normality with the Shapiro-Francia test using MedCalc v. 15.8. The same analysis was carried out for plans optimised also with 6 MV FFF energy.

The data were then analysed according to the following factors: patients (size and number of targets), TPSs (algorithms), type of MLCs, and delivery mode (Static Jaw (SJ) *vs* Jaw Tracking (JT)). A three-way analysis of variance (3w-ANOVA) was performed to determine the *F*-test statistics; *p* values of each variance component were computed and a post-hoc Scheffé test (PHSt) was carried out in order to identify intragroup differences or each factor identified as a significant predictor of OFs differences.

RESULTS

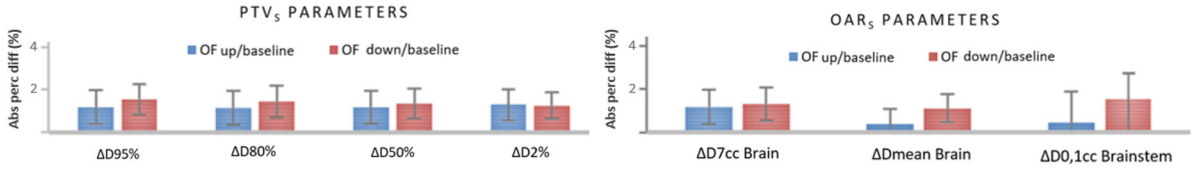
On applying Equation 1, the $OF_{baseline}$ fits showed a coefficient of determination $R^2 = 0.987$. The deviation of the $OF_{up/down}$ from the $OF_{baseline}$ fitted curves proved to be ≤10% with $R^2 = 0.982$ for both the OF_{up} and OF_{down} fits.

The absolute percentage dose differences between the recalculated plans and the relative baseline plans at selected dose/volume targets and OAR points, as the mean of the data for all centres and patients, are shown in Figure 2a and b. No statistically significant differences were found for all of the dose–volume points analysed. However, compared to baseline plans, and using 6 MV energy, the treatment plans with OF_{down} values resulted in higher differences ranging between −1.6% and −3.4% to PTVs and OARs respectively, while plans with OF_{up} values

Table 3. Variations of OFs $_{baseline}$ values, from the field $3.2 \times 3.2\,cm^2$ to $1 \times 1\,cm^2$ up to a maximum of + 3% (OF_{up} set) and a minimum of −3% (OF_{down} set) for the field size $1 \times 1\,cm^2$

Field size (cm^2)	Increment of the OFs baseline value (%)	Decrement of the OFs baseline value (%)
3.2 × 3.2	+0.8	−0.8
3.0 × 3.0	+1.0	−1.0
2.4 × 2.4	+1.6	−1.6
2.0 × 2.0	+2.0	−2.0
1.6 × 1.6	+2.4	−2.4
1.0 × 1.0	+3.0	−3.0

Figure 2. Average absolute percentage differences between recalculated plans (using the OF_{up} set and the OF_{down} set) and the relative baseline plans (with the $OF_{baseline}$ set) at: 2a) selected dose/volume points: $D_{95\%}$, $D_{80\%}$, $D_{50\%}$, $D_{2\%}$ for the PTV and 2b) $D_{7cc\ Brain}$, $D_{mean\ Brain}$, $D_{0,1cc\ Brainstem}$ for the OARs. OAR, organ at risk; OF,output factor; PTV, planning target volume



resulted in lower differences within +1.3% to both PTVs and OARs. When comparing 6 MV FFF and 6 MV beams, slightly higher percentage dose differences were observed for both OF_{up} and OF_{down} plans of approximately +0.1% and −1% respectively, with no statistically significant differences between them.

Focusing on the outcome variable ΔD_{80}(%) to PTV, the analysis performed according to the size and number of targets showed statistically significant differences between patient number five and patients' number 1 and 2 ($p < 0.001$) (Figure 3a). Moreover, while no statistically significant differences were observed for the three different MLC models (Figure 3b), they were found for the Monaco TPS (Elekta AB, Stockholm, Sweden) (Figure 3c and Table 4) compared to the other TPSs ($p < 0.001$). The Scheffé test also showed statistically significant differences between RayStation (RaySearch Laboratories, Stockholm) and Pinnacle (Philips Radiation Oncology Systems, Fitchburg, WI) which are both based on the Collapsed Cone Convolution (CCC) algorithm ($p < 0.001$). No differences were found between the Acuros (AXB) and the Analytical Anisotropic Algorithm (AAA) in the Eclipse TPS (Varian Medical Systems, Palo Alto, CA) therefore the AAA was selected to represent the data. Finally, the analysis on all the plans, showed median ΔD_{80} (%) values of −0.80% [range (−4.12; 0.71)] and −1.38% [range (−5.06; 0.20)] for the SJ *vs* the JT delivery mode respectively, with a trend towards a greater variation observed for the JT delivery mode, although not statistically significant ($p = 0.14$).

DISCUSSION

In volumetric modulated arc therapy (VMAT) the accuracy of dose calculation for treating intra cranial tumours depends on various dosimetric parameters, such as MLC modelling, the calculation algorithm and the field size dependence of small field output factors.[9,16] This is the first large multicentre study to evaluate the impact of each parameter on calculated doses, through purposely generated inaccuracies of OFs values.

Azimi et al evaluated the effect of modifying small field OF values (for the fields $2 \times 2\ cm^2$ and $3 \times 3\ cm^2$) commissioned in a single TPS (Pinnacle) by a considerably amounts,±5%, ±10% and ±20%, from the measured values.[6] It was observed that the extent of the discrepancy between measured and calculated intensity modulated radiotherapy (IMRT) plans depends on the jaw tracking ability of the accelerator. Kairn et al also investigated the effects of a set of OFs, acquired using an inaccurate detector, on planned doses.[7] They concluded that errors in excess of 10% may occur in radiotherapy treatments using fields smaller than 15 mm, if accurate data measurements are not acquired. Interestingly, they suggested that the overall clinical effects of inaccurate small field OFs may depend on the shape and volume of the targets (inaccurate OF measurement effects may be overcome when beam segments of various sizes are combined together, but not when treatment plans involve the use of small beam segments).

This study focused on spherical single or multiple targets with volumes ranging from the smallest measuring $0.3\ cm^3$ (Patient 1) to the largest of $9\ cm^3$ (Patient 4). The $OF_{baseline}$ values for each centre were modified up to a maximum of ±3% for the $1 \times 1\ cm^2$ field so that the two new OFs sets were still well described by the equation proposed by Sauer with the aim of simulating errors that cannot be easily identified even by a skilled medical physicist. Plans with OF_{down} sets resulted in higher variation of doses within 1.6 and 3.4% to PTVs and OARs respectively, while plans with OFs_{up} sets resulted in lower variation of doses within 1.3% to both PTVs and OARs. Our study suggests that slight errors in small field OFs achieve dose calculation accuracies within 3%

Figure 3. Average percentage dose differences ΔD_{80} (%) between recalculated plans (using the OF_{down} sets) and the relative baseline plans (with the $OF_{baseline}$ set) at dose/volume points D_{80} for the PTV according to: 3a) size and number of targets volume (patients); 3b) MLC models and 3c) algorithms/TPS. OF, outputfactor; PTV, planning target volume; TPS, treatment planning system

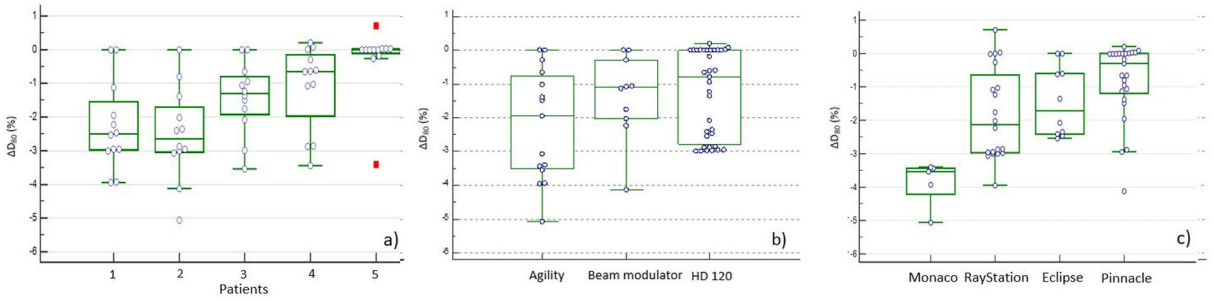


Table 4. Average percentage dose differences ΔD (%) between recalculated plans (using OF_{down} sets) and the relative baseline plans at selected dose/volume points: $D_{95\%}$, $D_{80\%}$ (showed also in Figure 3c), $D_{50\%}$, $D_{2\%}$ for the PTV and $D_{7cc\ Brain}$, $D_{mean\ Brain}$, $D_{0,1cc\ Brainstem}$ for the OARs referred to the four algorithms/TPS

Algorithms	PTV				Brain		Brainstem
	$\Delta D_{95\%}$	$\Delta D_{80\%}$	$\Delta D_{50\%}$	$\Delta D_{2\%}$	ΔD_{7cc} (%)	ΔD_{mean} (%)	$\Delta D_{0.1cc}$ (%)
MC	$-4{,}4 \pm 0{,}5^{a,c,d}$	$-3{,}8 \pm 0{,}7^{a,c,d}$	$-3{,}4 \pm 0{,}8^{a,c,d}$	$-2{,}6 \pm 0{,}9^{a,c,d}$	$-0{,}5 \pm 2{,}5$	$2{,}0 \pm 2{,}9^{a,c,d}$	-2.4 ± 1.9
CCC	$-1{,}7 \pm 1{,}4^{b,c}$	$-1{,}7 \pm 1{,}4^{b,c}$	$-1{,}6 \pm 1{,}4^{b,c}$	$-1{,}5 \pm 1{,}5^{b,c}$	$-2{,}4 \pm 1{,}2^{c}$	$-1{,}6 \pm 1{,}5^{b}$	-1.8 ± 1.4^{c}
AAA	$-1{,}4 \pm 1{,}1^{b}$	$-1{,}4 \pm 1{,}1^{b}$	$-1{,}4 \pm 1{,}1^{b}$	$-1{,}3 \pm 1{,}2^{b}$	$-1{,}8 \pm 0{,}8$	$-1{,}5 \pm 1{,}0^{b}$	-1.4 ± 1.1
*CCC**	$-0{,}8 \pm 1{,}3^{b,d}$	$-0{,}7 \pm 1{,}1^{b,d}$	$-0{,}7 \pm 1.0^{b,d}$	$-0{,}6 \pm 1{,}0^{b,d}$	$-0{,}9 \pm 1{,}4^{d}$	$-0{,}6 \pm 1{,}2^{b}$	-0.6 ± 1.2^{d}

TPS Pinnacle TPS RayStation; AAA = Analytical Anisotropic Algorithm, TPS Eclipse; CCC*=Collapsed Cone Convolution, TPS Monaco; CCC = Collapsed Cone Convolution,MC = Monte Carlo;OAR, organ at risk; PTV, planning target volume.
Statistical significance: [a] Test vs AAA; [b] Test vs MC; [c] Test vs CCC* (Pinnacle); [d] Test vs CCC (Ray Station).
Statistically significant differences ($p < 0.05$) for each algorithm towards the alternative ones are also reported.

with discrepancy inversely proportional (but not significant) to the size of the lesions considered (Figure 3a). Our results are in line with Azimi et al,[6] who found that decreasing/increasing the OFs by 5% resulted in approximately 0.8% higher and 2.9% lower calculated doses respectively.

In order to achieve high accuracy, other authors suggested fixing the jaws above the minimum size and generating shielding and modulation by the MLC, thus disabling the JT technique.[25–27] According to Azimi et al and Sendani et al,[6,17] also our analysis showed a trend towards less dependence from OFs values without Jaw tracking ($p = 0.14$). We can find a further indication of this by looking at the data represented in Figure 3a: in Patient 5, where all the TPS/MLC had to work with the more open jaws, given the presence of a number of metastases equal to three (single isocenter), the percentage variations of $\Delta D_{80(\%)}$ are certainly less than in all other patients with a single lesion.

It is interesting to note that in this scenario Monte Carlo dose calculation algorithm is essential, in fact the Figure 3c shows that Monaco is the most sensitive TPS to OF variations ($p < 0.05$)[28] ;

This was an important confirmation but it doesn't surprise so much because the MC approach is generally considered the gold-standard for determining dose distributions in any medium and it has been used to determine the accuracy of different dose calculation techniques by many authors.[29–32] Also, in the specific context of small fields using an anthropomorphic head phantom, Behinaein et al[33] found the average dose differences between measurements and MC, AAA and AXB equal to 0.2%, 3.2 and 2.7% respectively for small field sizes up to $0.5 \times 0.5\,cm^2$.

Unlike Mzenda et al,[34] who used both RayStation and Pinnacle TPS to develop comparable high-quality treatment plans, in our study small but statistically significant differences were found between the two algorithms in terms of target coverage and OARs sparing (Figure 3c and Table 4) with a trend of RayStation data towards the MC. Although the dose calculation algorithm is the same for both TPSs a different customization of the clinical machine in the physics module could be of importance.

Moreover, by adding a flattening filter weight to the RayStation TPS, Chen et al,[35] improved dose calculation accuracy for MLC defined fields to within 2%. In our study, no significant differences were observed in dose accuracy, irrespective of whether the FF component was used or not, however only three of the centres use FF Free beams which may be considered as one of the limitations of the analysis.

A great deal of time and effort was spent collecting the data from nine centers which enabled us to carry out an analysis that allow to highlight various individual factors and this is the true strength of this study. The authors are aware that the final results include various inseparable aspects (individual equipment, single or multiple targets, MLC type, algorithm, FF *vs* FFF) that may be subject for further investigation in the future.

However, framing the results of this work in the context of their clinical significance, we could say that errors up to ±3% in the measurement of OFs small fields have a small impact on the final dose calculation of VMAT radiosurgery treatment plans; in addition, our data emphasised that a fine algorithm is needed to highlight small uncertainties.

Therefore, our work could define a 'comfort zone' within a measurement errors on small fields does not translate into a relevant clinical error. Anyway, the goal of pursuing accuracy in measurements thought the use of suitable detectors and the search for their correction factors, should be the standard for a good clinical practice.

CONCLUSIONS

The commissioning process and beam modelling for dosimetric accuracy continues to be a major challenge for medical physicists.[30] The aim of our study was to evaluate how OFs undetectable inaccuracies in small fields can affect dose calculation in VMAT brain radiosurgery treatment plans. It was observed that simulated measurement OFs errors (up to ±3%) on small fields does not translate into a relevant clinical error. These errors are purely calculation based and do not take into account any machine related daily fluctuations or deviations from the commissioning data that might have occurred over time.

EQUAL CONTRIBUTION

Stefania Clemente and Maria Daniela Falco contributed equally

CONFLICT OF INTEREST

None of the authors has any conflict of interest with the published data and the software used in the present work. Alessandro Alparone is an employee of the company "Tecnologie Avanzate TA Srl" that distributes the software RayStation in Italy; this company supports the group in terms of collection of data. Data of all involved commercial platforms were analyzed objectively and independently from the vendor.

FUNDING

The authors received no financial support for the research, authorship, and publication of this article.

REFERENCES

1. Dimitriadis A, Kirkby KJ, Nisbet A, Clark CH. Current status of cranial stereotactic radiosurgery in the UK. *Br J Radiol* 2016; **89**: 20150452. doi: https://doi.org/10.1259/bjr.20150452
2. Schmitt D, Blanck O, Gauer T, et al. Technological quality requirements for stereotactic radiotherapy: expert review group consensus from the dGMP Working group for physics and technology in stereotactic radiotherapy. *StrahlentherOnkol* 2020; **196**: 421–43.
3. Das IJ, Ding GX, Ahnesjö A. Small fields: nonequilibrium radiation dosimetry. *Med Phys* 2008; **35**: 206–15. doi: https://doi.org/10.1118/1.2815356
4. Alfonso R, Andreo P, Capote R, Huq MS, Kilby W, Kjäll P, et al. A new formalism for reference dosimetry of small and nonstandard fields. *Med Phys* 2008; **35**: 5179–86. doi: https://doi.org/10.1118/1.3005481
5. International atomic energy agency, dosimetry of small static fields used in external beam radiotherapy, technical reports series No. 483. *IAEA* 2017;.
6. Azimi R, Alaei P, Higgins P. The effect of small field output factor measurements on IMRT dosimetry. *Med Phys* 2012; **39**: 4691–4. doi: https://doi.org/10.1118/1.4736527
7. Kairn T, Charles P, Crowe SB, Trapp JV. Effects of inaccurate small field dose measurements on calculated treatment doses. *Australas Phys Eng Sci Med* 2016; **39**: 747–53. doi: https://doi.org/10.1007/s13246-016-0461-y
8. Lechner W, Wesolowska P, Azangwe G, Arib M, Alves VGL, Suming L, et al. A multinational audit of small field output factors calculated by treatment planning systems used in radiotherapy. *Physics and Imaging in Radiation Oncology* 2018; **5**: 58–63Volume. doi: https://doi.org/10.1016/j.phro.2018.02.005
9. Cagni E, Russo S, Reggiori G, Bresciani S, Fedele D, Iori M, et al. Technical note: multicenter study of TrueBeam FFF beams with a new stereotactic diode: can a common small field signal ratio curve be defined? *Med Phys* 2016; **43**: 5570–6. doi: https://doi.org/10.1118/1.4961744
10. Russo S, Reggiori G, Cagni E, Clemente S, Esposito M, Falco MD, et al. Small field output factors evaluation with a microDiamond detector over 30 Italian centers. *Phys Med* 2016; **32**: 10121644–50. doi: https://doi.org/10.1016/j.ejmp.2016.10.017
11. Mancosu P, Pasquino M, Reggiori G, Masi L, Russo S, Stasi M. Dosimetric characterization of small fields using a plastic scintillator detector: a large multicenter study. *Phys Med* 2017; **41**: 33–8. doi: https://doi.org/10.1016/j.ejmp.2017.03.024
12. Clemente S, Masi L, Fiandra C, Cagni E, Villaggi E, Esposito M, et al. A multi-center output factor intercomparison to uncover systematic inaccuracies in small field dosimetry. *Phys Imaging Radiat Oncol* 2018; **5**: 93–6. doi: https://doi.org/10.1016/j.phro.2018.03.007
13. Talamonti C, Russo S, Pimpinella M, Falco MD, Cagni E, Pallotta S, et al. Community approach for reducing small field measurement errors: experience over 24 centres. *Radiotherapy and Oncology* 2019; **132**: 218–22. doi: https://doi.org/10.1016/j.radonc.2018.10.012
14. AkinoY MH, Tanaka Y, et al. Inter-institutional variability of small field-dosimetry beams among HD120™ multileaf collimators: a multi-institutional analysis *Phys. Med. Biol* 2018; **63**: 205–18.
15. Casar B, Gershkevitsh E, Mendez I, Jurković S, Huq MS. A novel method for the determination of field output factors and output correction factors for small static fields for six diodes and a microdiamond detector in megavoltage photon beams. *Med Phys* 2019; **46**: 944–63. doi: https://doi.org/10.1002/mp.13318
16. Girardi A, Fiandra C, Giglioli FR, Gallio E, Ali OH, Ragona R. Small field correction factors determination for several active detectors using a Monte Carlo method in the Elekta Axesse linac equipped with circular cones. *Phys Med Biol* 2019; **64**: 11NT0111NT01. doi: https://doi.org/10.1088/1361-6560/ab1f26
17. Sendani NG, Karimian A, Mahdavi SR, Jabbari I, Alaei P. Effect of beam configuration with inaccurate or incomplete small field output factors on the accuracy of treatment planning dose calculation. *Med Phys* 2019; **46**: 5273–83. doi: https://doi.org/10.1002/mp.13796
18. Lechner W, Primeßnig A, Nenoff L, Wesolowska P, Izewska J, Georg D. The influence of errors in small field dosimetry on the dosimetric accuracy of treatment plans. *Acta Oncol* 2020; **59**: 511–7. doi: https://doi.org/10.1080/0284186X.2019.1685127
19. Ohira S, Ueda Y, Akino Y, Hashimoto M, Masaoka A, Hirata T, et al. HyperArc VMAT planning for single and multiple brain metastases stereotactic radiosurgery: a new treatment planning approach. *Radiat Oncol*, (2018; **13**: 1313. doi: https://doi.org/10.1186/s13014-017-0948-z
20. International atomic energy Agency. commissioning and quality assurance of computerized planning systems for radiation treatment of cancer. *Technical Reports Series N. 430, IAEA, Wien* 2004;.
21. International atomic energy Agency. commissioning of radiotherapy treatment planning systems: testing for typical external beam treatment techniques. *IAEA-TECDOC No. 1583, IAEA, Wien* 2008;.
22. Fraass B, Doppke K, Hunt M, Kutcher G, Starkschall G, Stern R, et al. American association of physicists in medicine radiation therapy Committee task group 53: quality assurance for clinical radiotherapy treatment planning. *Med Phys* 1998; **25**: 1773–829. doi: https://doi.org/10.1118/1.598373
23. Taschereau R, Chatziioannou AF. Compressed voxels for high-resolution phantom simulations in gate. *Mol Imaging*

Biol 2008; ; **10**: 40–7PMIDPMCIDJan-Feb. doi: https://doi.org/10.1007/s11307-007-0110-7

24. Sauer O, Wilbert J. Measurement of output factors for small photon beams *Med. Phys* 2007; **34**: 1983–8.
25. Kerns JR, Followill DS, Lowenstein J, Molineu A, Alvarez P, Taylor PA, et al. Agreement between institutional measurements and treatment planning system calculations for basic Dosimetric parameters as measured by the imaging and radiation oncology Core-Houston. *Int J Radiat Oncol Biol Phys* 2016; **95**: 1527–34. doi: https://doi.org/10.1016/j.ijrobp.2016.03.035
26. Swinnen ACC, Öllers MC, Roijen E, Nijsten SM, Verhaegen F. Influence of the jaw tracking technique on the dose calculation accuracy of small field VMAT plans. *J Appl Clin Med Phys* 2017; **18**: 186–95. doi: https://doi.org/10.1002/acm2.12029
27. Fogliata A, Clivio A, Vanetti E, Nicolini G, Belosi MF, Cozzi L. Dosimetric evaluation of photon dose calculation under jaw and MLC shielding. *Med Phys* 2013; **40**: 1017062000. doi: https://doi.org/10.1118/1.4820443
28. Han T, Mikell JK, Salehpour M, Mourtada F. Dosimetric comparison of Acuros xB deterministic radiation transport method with Monte Carlo and model-based convolution methods in heterogeneous media. *Med Phys* 2011; **38**: 2651–64. doi: https://doi.org/10.1118/1.3582690
29. Brualla L, Rodriguez M, Lallena AM. Monte Carlo systems used for treatment planning and dose verification. *Strahlenther Onkol* 2017; **193**: 243–59. doi: https://doi.org/10.1007/s00066-016-1075-8
30. Andreo P. Monte Carlo simulations in radiotherapy dosimetry. *Radiat Oncol* 2018; **13**: 121. doi: https://doi.org/10.1186/s13014-018-1065-3
31. Hasani M, Mohammadi K, Ghorbani M, Gholami S, Knaup C. A Monte Carlo evaluation of dose distribution of commercial treatment planning systems in heterogeneous media. *J Cancer Res Ther* 2019; **15**(Supplement): S127–34. doi: https://doi.org/10.4103/jcrt.JCRT_1210_16
32. Alagar AGB, Mani GK, Karunakaran K. Percentage depth dose calculation accuracy of model based algorithms in high energy photon small fields through heterogeneous media and comparison with plastic scintillator dosimetry. *J Appl Clin Med Phys* 2016; **17**: 132–42. doi: https://doi.org/10.1120/jacmp.v17i1.5773
33. Behinaein S, Osei E, Darko J, Charland P, Bassi D, et al. Evaluating small field dosimetry with the Acuros xB (AXB) and analytical anisotropic algorithm (AAA) dose calculation algorithms in the eclipse treatment planning system. *J Radiother Pract* 2019; **18**: 353–64. doi: https://doi.org/10.1017/S1460396919000104
34. Mzenda B, Mugabe KV, Sims R, Godwin G, Loria D, et al. Modeling and dosimetric performance evaluation of the RayStation treatment planning system. *J Appl Clin Med Phys* 2014; **15**: 29–46. doi: https://doi.org/10.1120/jacmp.v15i5.4787
35. Chen J, Morin O, Weethee B, et al. Optimizing beam models for dosimetric accuracy over a wide range of treatments *Phys. Med* 2019; **58**: 47–53.