



Original paper

## Clinical evaluation of a transmission detector system and comparison with a homogeneous 3D phantom dosimeter

Francesca Romana Giglioli<sup>a,\*</sup>, Elena Gallio<sup>a</sup>, Pierfrancesco Franco<sup>b</sup>, Serena Badellino<sup>c</sup>, Umberto Ricardi<sup>b</sup>, Christian Fiandra<sup>b</sup>

<sup>a</sup> Medical Physics Unit, A.O.U. Città della Salute e della Scienza, Corso Bramante 88/90, 10126 Turin, Italy

<sup>b</sup> Department of Oncology, Radiation Oncology, University of Turin, Via Genova 3, 10126 Turin, Italy

<sup>c</sup> Radiotherapy Unit, A.O.U. Città della Salute e della Scienza, Corso Bramante 88/90, 10126 Turin, Italy

## ARTICLE INFO

## Keywords:

IMRT  
QA  
Dosimetry  
Gamma metric  
Pre-treatment verification

## ABSTRACT

**Purpose:** To evaluate the performances of the Dolphin system for pre-treatment verification (IBA Dosimetry, Schwarzenbruck, Germany) based on transmission measurements, employing a clinical perspective.

**Methods and materials:** Fifty treatment plans were verified by Dolphin and Delta<sup>4</sup> detectors (Scandidos, Uppsala, Sweden) during the same session and subsequently by the Delta<sup>4</sup> itself. The attenuation factor of the transmission detector (required for on-line dosimetry) was evaluated by comparing Delta<sup>4</sup> measurements with and without Dolphin. *Gamma* evaluation was performed to compare the plan dose with the one delivered in case of Delta<sup>4</sup> and Dolphin (gamma analysis within the structures). Dose–volume based parameters for PTV and OARs doses were considered for Compass calculation and Dolphin reconstruction and clinical decisions were made by two expert physicians in order to assess the “pass”, “fail” or “evaluate” grade of the treatment plans. A statistical analysis was performed to investigate the eventual correlation between Delta<sup>4</sup> gamma analysis and Dolphin clinical evaluation.

**Results:** A value of  $10.7\% \pm 0.7\%$  was found for detector attenuation. No patients were classified as “fail” by the two instruments as well as by physicians. No correlation was found between the Delta<sup>4</sup> *gamma* metric and physician classification; conversely, a significant correlation was observed for Dolphin between the numbers of points with  $\gamma \geq 1$  (gamma failure rate, as evaluated by the Dolphin) in the PTV area and clinical decision.

**Conclusion:** The Dolphin system demonstrated to be an accurate detector for pre-treatment purposes and could be used as a clinical decision making tool for plan acceptance.

### 1. Introduction

Over the last few years, numerous technological advancements took place in modern radiation therapy delivery, gradually shifting from static intensity modulated radiation therapy (IMRT) to volumetric modulated arc therapy (VMAT). Due to the complexity of beam modulation, patient specific pre-treatment measurement is recommended for these two techniques by different international documents [1–3]. The most common method for pre-treatment verification relies on the transfer of the treatment plan to a standard phantom geometry within the TPS, in order to recalculate the dose distribution, comparing it with measured dose distribution with a *gamma* analysis [4]. *Gamma* passing rate analysis is the most widely used QA metric for establishing dosimetric calculation accuracy and treatment deliverability. Certain reports suggest various methods to define a set of action levels to support

decision on the acceptance of pre-treatment verification results (AAPM Task Group 119 [5] and ICRU 83 [6]). The recent recommendations of AAPM Task Group No. 218 [7] describe tolerance limits and methodologies for IMRT measurement-based verification QA. The recommendations are intended to improve the IMRT QA process and to establish universal and comparable IMRT QA criteria among different institutions, differently from the past when the acceptance level was based on a specific population. The report provides evaluation metrics, measurement methods, methodology for absolute dose verification and process based on tolerance and action limits. The *Gamma* analysis is not patient-specific ‘per se’, since it does not provide information on the anatomical location or dose level where the failure occurs. Moreover the dose differences, whenever present, may be clinically irrelevant, even if IMRT- QA performance goals are unmet [8,9]. To provide a reliable response concerning patient-specific QA, the same DVH-based

\* Corresponding author. Address: Medical Physics Unit, A.O.U. Città della Salute e della Scienza di Torino, Corso Bramante 88/90, 10126 Turin, Italy.  
E-mail address: [fgiglioli@cittadellasalute.to.it](mailto:fgiglioli@cittadellasalute.to.it) (F.R. Giglioli).

<https://doi.org/10.1016/j.ejmp.2019.01.016>

Received 3 July 2018; Received in revised form 23 January 2019; Accepted 26 January 2019

Available online 04 February 2019

1120-1797/ © 2019 Published by Elsevier Ltd on behalf of Associazione Italiana di Fisica Medica.

metrics employed to approve the plan at first instance can be used. For this purpose, three-dimensional dose distributions within patient anatomy may be estimated from measured dose, using several commercially available software applications such as Compass system (IBA Dosimetry, Schwarzenbruck, Germany), Mobius3D software (Mobius Medical Systems, Houston, TX), Dosimetry Check software (Math Resolutions, LLC, Columbia, MD), Octavius 4D (PTW, Freiburg, Germany), arcCHECK 3DVH (Sun Nuclear Corporation, Melbourne, Florida) and Delta<sup>4dvh</sup> Anatomy (Scandidos, Uppsala, Sweden) [10–12]. These approaches, even if different, are considered more useful compared to conventional pre-treatment QA because the traditional *gamma* analysis is performed inside the anatomic structures of the patient and DVHs comparisons (planned and measured) can be carried out with clinical considerations.

The aim of this study was to evaluate the performance of a verification platform based on transmission measurements (Compass and Dolphin, IBA Dosimetry, Schwarzenbruck, Germany) [13,14] from a clinical point of view and to compare the results with standard *gamma* metric obtained by a homogeneous 3D phantom dosimeter (Delta<sup>4</sup>, Scandidos, Uppsala, Sweden). In collaboration with 2 radiation oncologists, an evaluation criterion (tolerance and action levels) was set, based on DVHs values for PTV, parallels or serials organs at risk, simulating the plan evaluation process to finally provide a clinical perspective on the QA outcome. An investigation of the possible correlations between the clinical decisions derived from Dolphin DVHs analysis and *gamma* metric of the Delta<sup>4</sup> measurements was carried out.

## 2. Materials and methods

### 2.1. Dolphin detector

The Dolphin detector is a transmission detector made of a two-dimensional ionization chamber array for pre-treatment verification and on-line treatment monitoring. It has 1514 detectors (1513 air-vented parallel ionization chambers and 1 diode) arranged in a grid. The ion chambers are 2 mm high and have a diameter of 3.2 mm with an effective volume of 16.1 mm<sup>3</sup>. They cover an active area of 24 × 24 cm<sup>2</sup> (40 × 40 cm<sup>2</sup> at isocenter distance). The center-to-center distance between the chambers is 5 mm in the inner 15 × 15 cm<sup>2</sup> field and 10 mm in the outer area. During measurements, the Dolphin detector is mounted on the gantry in order to ensure that it is constantly perpendicular to the central beam axis for all gantry angles. After dismounting and remounting the detector, positioning errors for a certain detector setup are measured; a correction based on the measured errors is applied to the acquired data also accounting for the daily output of the Linac. The nominal Dolphin attenuation is in the range of 6–8% for a 10 × 10 cm<sup>2</sup> field size and SSD = 100 cm for 6 MV photon beams [14]. However, the manufacturer recommends determining the precise attenuation factor for on-line dosimetry.

The Compass is a software with a dose engine based on collapsed cone convolution superposition to calculate the dose using the fluence of the imported RT plan (calculated dose) and/or to reconstruct the dose inside the patient based on Dolphin measured fluence (reconstructed dose). It is therefore possible to make a 3D comparison between the TPS dose and Compass calculated and Compass reconstructed doses. Compass software can be employed as a secondary TPS [13–16] and the congruence between TPS involved in this study and Compass calculated dose has been verified at the stage of Dolphin acceptance.

### 2.2. Delta<sup>4</sup> detector

All dose distributions were also measured employing the Delta<sup>4</sup> detector; it is composed of 2 orthogonal planar diode arrays placed 45° oblique to the sagittal and coronal planes and arranged in a cylindrical polymethyl methacrylate (PMMA) phantom. A total of 1096p-type

**Table 1**

Plans included in the study.

Treatment sites	Number of plans	TPS	Linac
Lung SBRT	10	Monaco	Elekta Axesse
CNS	10	Monaco	Elekta Axesse
Lung	10	Monaco	Elekta Synergy
Anal/Rectum cancer	10	Pinnacle <sup>3</sup>	Elekta Synergy
Head & neck	10	Pinnacle <sup>3</sup>	Elekta Synergy

diodes cover the measurement area of 20 × 20 cm<sup>2</sup> in each measurement plane. The diodes are spaced at 5 mm intervals over the central 6 × 6 cm<sup>2</sup> area of the planes and at 1 cm outside the central region. In the literature, several studies have evaluated Delta<sup>4</sup> phantom for IMRT and VMAT plans verification [11,17–20]. A 4 fields plan was measured before every session and a correction factor for daily output and setup was applied to the measurements. Our Delta<sup>4</sup> is in its basic configuration without the Delta<sup>4dvh</sup> Anatomy option.

### 2.3. Plans, treatment planning systems and linear accelerators

A set of 50 patients were measured using both Dolphin and Delta<sup>4</sup> detectors. Different treatment sites were included and 10 cases were randomly selected from our database for each site (Table 1). The specific scenarios were lung stereotactic body radiation therapy (LSBRT), central nervous system (CNS), locally advanced lung (L), anal/rectal (AR) and head and neck (HN) cancer cases. All plans were generated with an inverse optimization engine using a volumetric modulated arc therapy technique (VMAT). LSBRT, CNS and L were planned with Monaco TPS (v. 5.0, Elekta AB, Stockholm, Sweden); AR and HN were planned with Pinnacle<sup>3</sup> (v. 9.10, Philips Medical Systems, Fitchburg, WI). All the plans were delivered with a 6 MV photon beam with 2 different linear accelerators: LSBRT and CNS with an Elekta Axesse with Beam Modulator multi leaf collimator (40 leaf pair; 0.4 cm leaf width at isocenter); L, AR and HR with an Elekta Synergy with MLCi<sub>2</sub> multi leaf collimator (40 leaves pair; 1 cm leaf width at isocenter). Both TPS and linacs were accredited with dosimetry audit using the IROC Houston IMRT phantoms thus we were confident enough on the accuracy of absolute delivered dose.

### 2.4. Dolphin attenuation

The Dolphin is also configured for patient on-line verification (even if not yet certified for this purpose); the system is mounted on the gantry during the treatment session. Its attenuation is measured with the Delta<sup>4</sup> detector. Different dimensions of the tumours were considered to account for the consistent variability of the field sizes. All 50 plans were measured using Delta<sup>4</sup> with and without Dolphin: the measurement with Delta<sup>4</sup> alone was used as a reference dose and the measurement performed with the Dolphin was rescaled to find a correction factor. This factor (in percentage) was evaluate “matching” the rescaled dose (Dolphin inserted) with the reference one in terms of median dose and passing rate *gamma* index. The mean correction factor for the 50 patients and the corresponding standard deviation were calculated.

### 2.5. Reproducibility

Several reports outline the characteristics and performances of the 2 systems [11,13–15,20,21]; the reproducibility of the 2 detectors used in our study (Dolphin and Delta<sup>4</sup>) was tested in a clinical context (lung case). The plan was measured 10 consecutive times with both Dolphin and Delta<sup>4</sup>. To measure the reproducibility, the standard deviation in % of number of points with local *gamma* index ≤ 1 was evaluated for Delta<sup>4</sup> and the standard deviation in % of number of points with local *gamma* index ≥ 1 in the PTV region (failure rate) for Dolphin (as

**Table 2**  
Clinical constraints. ( $D_{max}$  refers to 0.035 cc).

Clinical context	“Dose Parameters”	Dose Constraints
Lung SBRT	Ipsilateral Lung Mean Dose eq 2 Gy	12 Gy 30 (3 fx)/38 (5 fx) Gy
	Heart $D_{max}$	25 (3 fx)/35 (5 fx) Gy
	Esophagus $D_{max}$	30 (3 fx)/40 (5 fx) Gy
	Trachea $D_{max}$	22 (3 fx)/30 (5 fx) Gy
	Spinal Cord $D_{max}$	37 (3 fx)/43 (5 fx) Gy
	Rib $D_{max}$	45 (3 fx)/53 (5 fx) Gy
	Aorta $D_{max}$	
Lung	Lungs-PTV V20Gy	35%
	Ipsilateral Lung V20Gy	55%
	Heart Mean Dose	20 Gy
	Heart V30Gy	50%
	Esophagus Mean Dose	34 Gy
	Esophagus V40Gy	40%
	Spinal Cord $D_{max}$	45 Gy
CNS	Brain Mean Dose	20 Gy
	Brain V20Gy	40%
	Brain V40Gy	20%
	Brainstem $D_{max}$	60 Gy
	Left Lens $D_{max}$	5 Gy
	Right Lens $D_{max}$	5 Gy
	Left Cochlea Mean Dose	45 Gy
	Right Cochlea Mean Dose	45 Gy
	Left Optical Nerve $D_{max}$	54 Gy
	Right Optical Nerve $D_{max}$	54 Gy
Anal Cancer/Rectum	Bladder V30Gy	50%
	Bladder V40Gy	35%
	Bladder V50Gy	5%
	Bowel $D_{max}$	50 Gy
	Bowel V30Gy	200 cc
	Right/Left Thigh-bone V30Gy	50%
Head & Neck	Contralateral Parotid Mean Dose	26 Gy
	Oral Cavity Mean Dose	55 Gy
	Larynx Mean Dose	45 Gy
	Spinal Cord $D_{max}$	45 Gy
	Brainstem $D_{max}$	60 Gy

calculated by Compass).

### 2.6. Patient measurements

Dolphin system: the dose distribution was computed and reconstructed in Compass for each treatment and dose-volume indices were evaluated for the quantitative analysis. The presence of the couch was included both in the TPS and in the Compass calculation by contouring the support and considering its real electron densities. The dose constraints for each site used in our clinical practice were employed for the comparison (Table 2) and evaluated against the measured values. A clinical evaluation of the plan measured with Dolphin was performed: 2 dedicated radiation oncologists analysed the measured dose values for PTV and OARs and defined a clinical decision criterion: “fail” if any of the serial organs exceeded its constraint or if V95% of the PTV was less than 90% of the prescribed dose, “evaluate” if the PTV showed high over dosage within the PTV or if any of the parallel organs exceed the constraint of less than 10% of the constraint itself (CO) and “pass” if all the criteria were met. Prior to the clinical evaluation, the 2 physicians attended a training session to homogenize the decision making process.

Delta<sup>4</sup> detector: the measurements acquired with Delta<sup>4</sup> alone were evaluated. For each plan the *gamma* indexes 3%–3 mm, 3%–2 mm and 2%–2 mm were reported with threshold of 20% isodose and local evaluation as in our routine; the decision criteria used for the two systems, are shown in Table 3.

The metrics proposed by Du et al. [22] to quantify plan complexity were adopted to find possible correlation with the measurements results of the 2 systems: plan average beam area (PA), plan average beam irregularity (PI) and plan average beam modulation (PM)

$$PA = \frac{\sum_i (BA_i \cdot MU_i)}{MU_p} \tag{1}$$

with

$$BA_i = \frac{\sum_j (MU_{ij} \cdot AA_{ij})}{MU_i} \tag{2}$$

where  $BA_i$  is the area of beam  $i$ ,  $MU_i$  are monitor units of beam  $i$ ,  $MU_p$  total monitor units in the plan,  $MU_{ij}$  monitor units of segment  $j$  of beam  $i$  and  $AA_{ij}$  is the area of segment  $j$  of the beam  $i$ . Plan average beam modulation is defined as:

$$PM = \frac{\sum_i (BM_i \cdot MU_i)}{MU_p} \tag{3}$$

with

$$BM_i = 1 - \frac{\sum_j (MU_{ij} \cdot AA_{ij})}{MU_i \cdot U(AA_{ij})} \tag{4}$$

$$AA_{ij} = \frac{AP_{ij}^2}{4\pi(AA_{ij})} \tag{5}$$

where  $U(AA_{ij})$  is the union area of all apertures of beam  $i$ . PI is computed using, instead of  $BA_i$  and  $AA_{ij}$ ,  $B_{ii}$  (beam irregularity) and  $AI_{ij}$  (segment aperture irregularity), calculated like (5) where  $AP_{ij}$  is the aperture perimeter of segment  $j$  of the beam  $i$ .

The t-Student (two-side) test and the analysis of variance (ANOVA) were performed in order to find correlations among different parameters (linac, TPS, monitor units, plan complexity index, pathology, clinical decision, *gamma* metric); p-values < 0.05 were considered as statistically significant.

## 3. Results

### 3.1. Dolphin attenuation

Fig. 1 shows the distribution of attenuation values at 6 MV for all patients; the mean correction factor calculated from the patients measured with and without Dolphin mounted onto the gantry was  $10.7\% \pm 0.71\%$  corresponding to 6.6% of variation.

### 3.2. Reproducibility

Reproducibility in terms of standard deviation in % of number of points with local gamma index  $\leq 1$  was 0.42% for Delta<sup>4</sup>; for Dolphin the standard deviation in % of number of points with local gamma index  $\geq 1$  in the PTV region was 0.22%.

### 3.3. Patient measurements

Dolphin system: Table 4 shows the results of the clinical evaluation with the Dolphin: among the 50 patients evaluated, 78% were classified “Pass”, 22% were considered for re-evaluation by repeating the measurement or checking step by step the flowchart and 0% were classified as “fail” (i.e. to be replanned).

Delta<sup>4</sup> detector: the results in terms of pass, evaluate and failure rate are reported in Table 5, as functions of different parameters of the *gamma* metric (dose difference in % and Distance to agreement). Among the 50 patients under study, 84% were classified Pass, 16% were considered for re-evaluation by checking the flowchart or repeating the measurement and 0% were classified as fail (i.e. to be replanned) due to *gamma* 2%–2 mm, otherwise 98%, 2% and 0% for *gamma* 3%–3 mm.

On the analysis of all parameters, a significant correlation was found between clinical evaluation and the number of points with  $\gamma \geq 1$  (as calculated in Compass) inside the PTV for the Dolphin system ( $p = 0.03$ ). On the contrary, with Delta<sup>4</sup>, no correlation was found

**Table 3**  
Decision criteria for Delta<sup>4</sup> and clinical evaluation criteria for Dolphin.

Delta <sup>4</sup>	3% – 3 mm	2% – 2 mm	Dolphin	Serial organs	Parallel organs	PTV
pass (tolerance limit)	$\gamma \geq 90\%$	$\gamma \geq 80\%$	pass	$D_{max} \leq CO$		$V95\% \geq 90\%$
evaluate (action limit)	$85\% \leq \gamma < 90\%$	$75\% \leq \gamma < 80\%$	evaluate		$CO < DVH_{value} \leq 1.1^* CO$	$107\% < D1 < 115\%$
fail	$\gamma < 85\%$	$\gamma < 75\%$	fail	$D_{max} > CO$		$V95\% < 90\%$

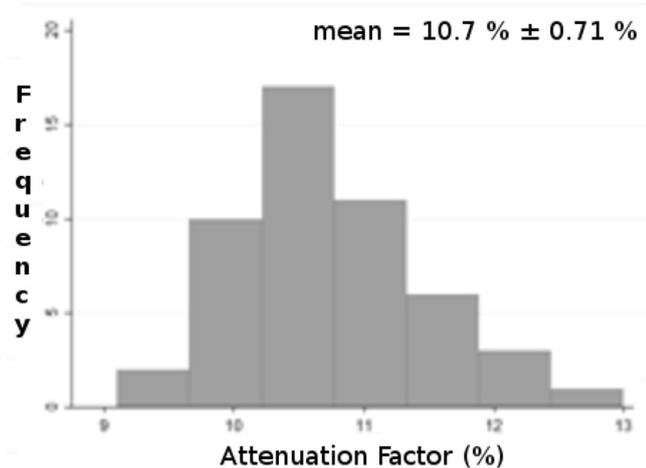


Fig. 1. Percentage attenuation distribution.

**Table 4**  
Clinical evaluation results in terms of “pass”, “evaluate” or “fail” for Dolphin.

Site	Pass	Evaluate	Fail	Fail or Evaluate motivation
H&N	5	5	0	Dose to parotids PTV overdose Oral cavity dose
Lung	10	0	0	
CNS	7	3	0	Lens doses PTV underdosage
Lung SBRT	8	2	0	Dose to trachea
Rectum	10	0	0	

between the same clinical evaluation and the *gamma* metric (Fig. 2).

Fig. 3 shows a significant correlation between average *gamma* in the PTV region (Dolphin) and the oncological site and a correlation between the *gamma* metric (Delta<sup>4</sup>) and TPS (Monaco or Pinnacle<sup>3</sup>). In the case of Dolphin, the Fisher-Hayter pairwise comparisons test showed that rectum&anus was significantly different from the other groups ( $p < 0.05$ ).

The values of the plan complexity metric are shown in Table 6; in case of Anus&rectum the mean values of PA and PI were higher compared to other clinical sites but no significant correlation was found.

#### 4. Discussion

Pre-clinical QA such as patient specific pre-treatment measurements are required for modern radiotherapy approaches, given the complexity

**Table 5**  
Number of patients classified as pass, evaluate or fail in function of parameters of *gamma* metric for Delta<sup>4</sup> detector.

Sites	Pass	Evaluate	Fail	Pass	Evaluate	Fail
	2%–2 mm	2%–2 mm	2%–2 mm	3%–3 mm	3%–3 mm	3%–3 mm
H&N	6	4	0	10	0	0
Lung	10	0	0	10	0	0
CNS	10	0	0	10	0	0
Lung SBRT	10	0	0	10	0	0
Rectum	6	4	0	9	1	0

of planning and treatment delivery, especially during the commissioning phase of a new machine, technique or software [10,11,13–15,20–24]. In our study, we focused on two different platforms: a traditional 3D homogeneous phantom to be positioned on the couch and a new transmission detector inserted on the gantry. The Delta<sup>4</sup> detector is a bi-planar 3D diode arrays for which the patient plan is computed on the 3D homogeneous phantom and the *gamma* analysis is performed to compare the two dose distributions (calculated and measured) without any relationship to the patient’s anatomy and structure. Conversely, the Dolphin transmission detector, combined with Compass, allows evaluating the absorbed dose on the patient CT, calculating the DVHs, the *gamma* index within the structures and allowing to clinically compare calculated and measured DVHs. This system can also be used as in vivo dosimetry even if, at present, there is no agreement with manufacturer companies in terms of patient safety. Many authors have studied the performances of the two devices employed in our comparative study, namely Delta<sup>4</sup> and Compass with transmission detector (Dolphin) [3–15,20]: all of them found the two instruments suitable for treatment verification. Of interest the study by Thoelking et al [13] that compared Dolphin with a 2D platform and analyzed 18 patients for only 3 clinical sites. Their analysis was performed defining the acceptance criteria based on average *gamma* values (pass  $\gamma \leq 0.4$  evaluate  $0.4 < \gamma > 0.5$  and fail  $\gamma \geq 0.5$ ). Our effort has been addressed to introduce a clinical perspective in strict collaboration with radiation oncologists, defining clinical criteria as normally done by physicians during plan approval. Nakaguchi et al. [15] as well evaluated the performances of Dolphin (sensitivity, detectability) using various methods but only tested the clinical behavior of the system for SBRT with no clinical considerations performed. The aim and the novelty of our study were to investigate the potential correlations between the two methods used for pre-treatment QA on a group of 50 patients with different tumors: *gamma* evaluation in a homogeneous phantom (Delta<sup>4</sup>) and clinical DVHs analysis (Dolphin). We intended to correlate the results in term of pass, evaluate and fail between applied *gamma* criteria (Delta<sup>4</sup>) and clinical decision (Dolphin).

To compare calculated and measured dose distributions in a homogeneous phantom (Delta<sup>4</sup>), common metrics, including local or global *gamma* analysis pass rates based on thresholds of 3%–3 mm or 2%–2 mm, are generally used. The sensitivity of these metrics in detecting delivery errors has been widely studied and discussed within the literature [8,9,23–25]. The AAPM TG 218 was published in April 2018 and proposed universal tolerance and action limit [7]; at the publication time we already had full experience in pre-treatment QA based on consolidated procedures and have not yet implemented the TG 218 recommendations in the department routine. As mentioned above we calculated *gamma* metric with local normalization, 20% threshold and

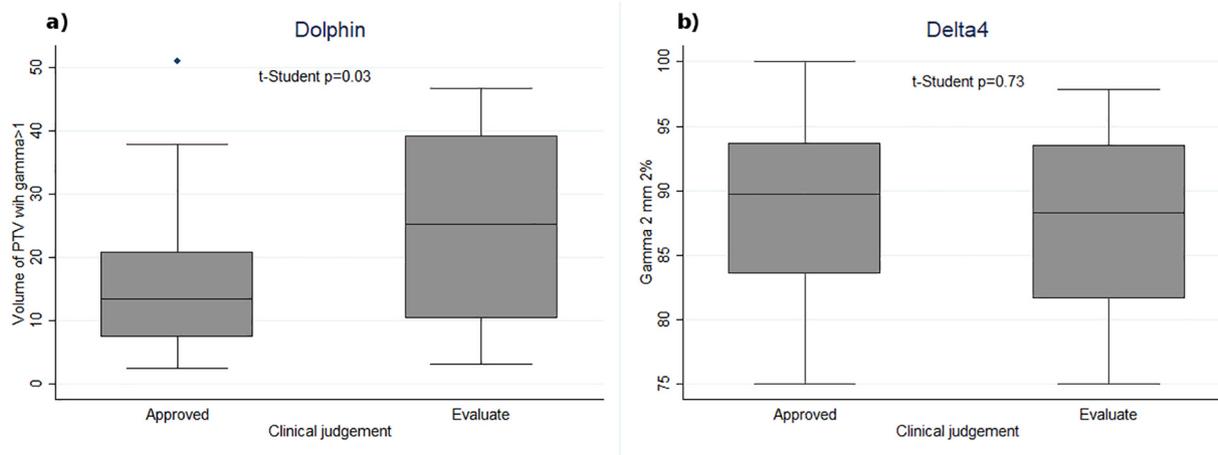


Fig. 2. Correlation between the clinical evaluation and the results obtained for the Dolphin (a) and Delta<sup>4</sup> (b).

2%-2mm also based on published data and the same parameters were used for the 2 instruments. It should be noted that the use of local normalization implies lower *gamma* values than the global normalization proposed by the report and hence a consistent effort will be needed to correlate the new standard data with the historical results. The DVH analyses were performed, according to the TG 218, considering the distribution of *gamma* values inside the structures of the patient without any consideration on clinical relevance. Nevertheless, the clinical evaluation of the measured DVHs was not accounted for in the aforementioned report, leaving this methodology prone to interpretation and non-standardised.

Moreover, as far as dose-volume-based analysis is concerned, the most important aspect was the interpretation of delivery errors within patient’s anatomy. This provides insights into the potential harm for the patient and helps us to improve the decisional process. Since Dolphin may be used as an on-line dosimeter; we evaluated a correction factor accounting for its attenuation. Thoelking et al. [13,14] measured it in a planar geometry for various field sizes, SSDs and depths in water phantom. They found an absorption varying from 10.3% to 7.9% depending on the field size, while our geometry has followed a clinical criterion involving various treatment complexities and geometries. The Delta<sup>4</sup> phantom was used to measure attenuation for all the 50 patients included; the results are not conflicting with those of Thoelking et al but, conversely, are complementary: the value of 10.7% ± 0.7% translates the planar data found by Thoelking to a volumetric 3D geometry. A noteworthy finding has been also the small result variability, which confirms the consistency of the measurement.

It has to be noticed that we are comparing 2 systems having the

Table 6

Results of plan complexity metrics.

TPS	Anatomical Site	PA [cm <sup>2</sup> ] [mean (min, max)]	PI [mean; min, max]	PM [mean; min, max]
Monaco	Lung	45.5 (28.3, 92.6)	5.4 (2.8, 10.0)	0.5 (0.1, 1.0)
	CNS	21.9 (5.9, 54.4)	4.3 (3.2, 5.7)	0.9 (0.3, 1.0)
	Lung SBRT	8.9 (4.0, 15.4)	4.8 (2.4, 8.5)	0.9 (0.6, 1.0)
Pinnacle <sup>3</sup>	H&N	49.3 (26.3, 75.5)	5.7 (2.9, 7.7)	0.8 (0.2, 1.0)
	Anal/Rectum	72.1 (46.0, 94.3)	7.1 (3.6, 9.1)	0.9 (0.8, 1.0)

same task but different metrics: a numerical evaluation (*gamma* value) in a homogeneous media (the Delta<sup>4</sup>) or a DVH analysis within the patient anatomy (the Dolphin) scored following a clinical decision. The clinical evaluation of the plans in terms of pass, fail or evaluate enabled us to find correlations among the various parameters related to the plans itself (complexity score, TPS, linac) and gamma indexes based on Delta<sup>4</sup> or Dolphin.

Stasi et al. [25] investigated the predictive power of the conventional gamma passing rate metric through the analysis of the correlation between planned and perturbed DVH. The perturbed DVH was calculated by using dedicated software, able to modify the dose distribution according to the dose discrepancies measured with a detector, to predict the delivered 3D dose distribution in the patient. They found

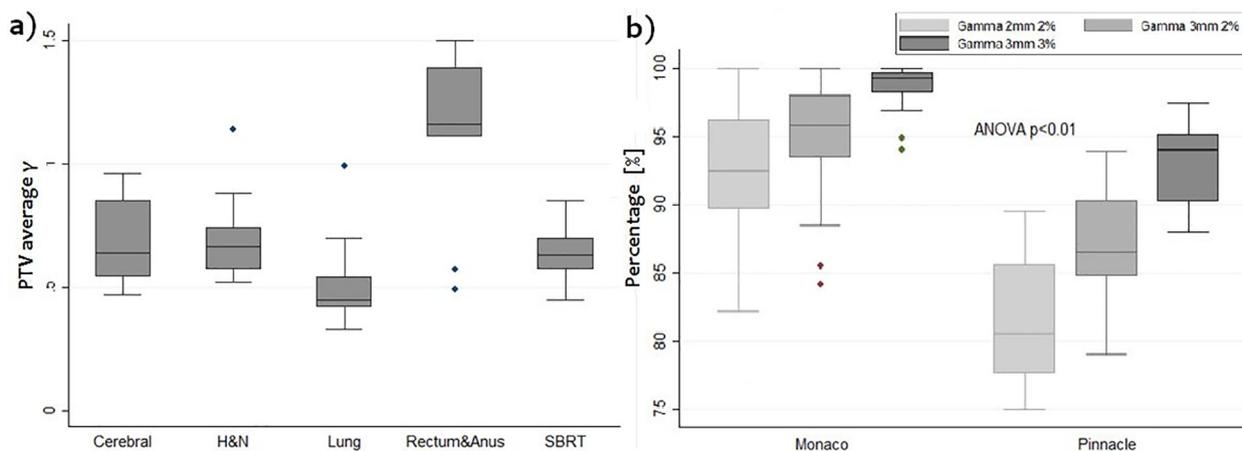


Fig. 3. Correlation between average gamma PTV (Dolphin) and site (a) and correlation between gamma metrics (Delta<sup>4</sup>) and TPS (Monaco and Pinnacle<sup>3</sup>) (b).

weak correlations between gamma pass rates and DVH values for PTV volumes and a lack of correlation for OAR volumes. Their conclusions show that further research on the clinical relevance of the  $\gamma$ -index analysis is required, calling for the need to integrate the IMRT QA analysis results with a methodology enabling clinicians to predict the impact of the discrepancies between planning and delivery in the patient dose distribution. Our results demonstrated a similar paradigm: Fig. 2 displays the correlation between the clinical evaluation and the results obtained for the Dolphin and Delta<sup>4</sup>, pointing out the lack of correlation between the Delta<sup>4</sup> gamma pass rate and the clinical evaluation obtained with our data. This confirms the conclusions drawn by Stasi et al. who suggested using this QA test as a constancy test and not as a method for detecting clinical relevant errors. Although numerous recent publications support this theory [26–29], the gamma analysis with a homogeneous phantom is still being used by most institutions; for this reason the users should have a clear knowledge on the performances of each dosimetric platform. Furthermore, it is important to highlight the significant correlation found between the clinical decision and the number of points with gamma  $\geq 1$  for PTV with the Dolphin system. This, probably, reveals that radiation oncologist's opinion on the reliability of the measure obtained with calculation was mostly focused on target volume coverage or over dosage. In our study, the indices of plan complexity for each patient were computed, since several studies reported on the correlation between plan complexity and  $\gamma$ -index analysis in patient specific QA. The maximum correlation between the modulation complexity score [28] and  $\gamma$ -pass rate was determined for 2%-2mm global  $\gamma$ -criteria; our data, however, didn't show any correlation between complexity plan parameters and gamma metric. An outcome of the pre-treatment program is, undoubtedly, the focus on the commissioning procedures and the evaluation of possible differences between the dose calculation accuracy by the TPS or dose delivery errors. In our study, a significant difference was found for Dolphin system between PTV average gamma and the clinical site involved; for rectal cancer cases (Fig. 3a), the measured data presented an average gamma value higher than the one observed in other settings. Although no significant correlations were found between the plan complexity parameters and the measurement results, a higher level of "stress" for the MLC can be assumed, because of the need to conform to very large concave volumes. However, this discrepancy doesn't translate into clinical evidence as 10 out of 10 of the rectal cancer cases were scored as "pass" and the plan complexity indices did not differ significantly. Finally, Delta<sup>4</sup> showed differences in gamma metric according to the TPS used (Fig. 3b), namely Monaco or Pinnacle<sup>3</sup>; both these systems are consolidated in our professional practice but, the presence of the Monte Carlo algorithm within Monaco could provide more accurate results in terms of measured dose.

## 5. Conclusions

The aim of this study was to compare the clinical performance of a verification platform based on transmission measurements (Compass and Dolphin) with the gamma metric obtained by a homogeneous 3D phantom dosimeter (Delta<sup>4</sup>). Gamma tests and DVH-based comparisons were carried out for a cohort of 50 patients affected with different tumors; in particular we set a decision criterion in strict collaboration with the radiation oncologist to provide the final evaluation of the QA outcome. Our data showed a correlation between clinical decision and PTV gamma metric in the PTV area for Dolphin system while no correlation was observed with Delta<sup>4</sup> measurements. Therefore the PTV number of points with gamma  $\geq 1$  (2%-2mm) seems to predict the consistency of the irradiated dose to the patient: the reliability of the irradiated plan could be evaluated by calculating this parameter. Future research should be carried out to improve the robustness of the clinical evaluation of the irradiated DVHs and effort should be undertaken to

make this process less subjective with the use of plan quality scores.

## Acknowledgment

We would like to thank G. Bartesaghi and S. La Civita for their helpful comments and suggestions. Thanks are also due to Jimmy Jimenez for continuous support.

## References

- [1] Frenzel T, Krüll A. The use of IMRT in Germany. *Strahlenther Onkol* 2015;191(11):821–6.
- [2] Guidelines for the verification of IMRT. ESTRO Booklet no. 9. 2009 Brussels, Belgium.
- [3] Moran JM, Dempsey M, Eisbruch A, Fraass BA, Galvin JM, Ibbott GS, et al. Safety considerations for IMRT: executive summary. *Med Phys* 2011;38(9):5067–72.
- [4] Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys* 1998;25(5):656–61.
- [5] TG-119 IMRT Commissioning Tests Instruction for Planning, Measurement, and Analysis Digital File Support .
- [6] International Commission on Radiation Units and Measurements ICRU Report 83: Prescribing, recording, and reporting photon-beam intensity modulated radiation therapy (IMRT). *J ICRU*, 10 (2010), p. 1–10.
- [7] Miften M, Olch A, Mihailidis D, Moran J, Pawlicki T, Molineu A, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM Task Group No. 218. *Med Phys* 2018;45(4):e53–83.
- [8] Negri A, Scaggion A, Rossato M, Roggio A, Simonato F, Bacco S, et al. Evaluation of a phantom related gamma index threshold for VMAT QA. *Phys Med* 2014;30:e70.
- [9] Nelms BE, Zhen H, Tomé WA. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys* 2011;38(2):1037.
- [10] Tyagi N, Yang K, Yan D. Comparing measurement-derived (3DVH) and machine log file-derived dose reconstruction methods for VMAT QA in patient geometries. *J Appl Clin Med Phys* 2014;15:54–66.
- [11] Freygelman V, Stambaugh G, Opp D, Zhang G, Moros EG, Nelms BE. Cross validation of two commercial methods for volumetric high-resolution dose reconstruction on a phantom for non-coplanar VMAT beams. *Radiother Oncol* 2014;110:558–61.
- [12] McDermott LN, Wendling M, Nijkamp J, Mans A, Sonke JJ, Mijnheer BJ, et al. 3D in vivo dose verification of entire hypo-fractionated IMRT treatments using an EPID and cone-beam CT. *Radiother Oncol* 2008;86:35–42.
- [13] Thoelking J, Fleckenstein J, Sekar Y, Boggula R, Lohr F, Wenz F, et al. Patient-specific online dose verification based on transmission detector measurements. *Radiother Oncol* 2016;119:351–6.
- [14] Thoelking J, Sekar Y, Fleckenstein J, Lohr F, Wenz F, Wertz H. Characterization of a new transmission detector for patient individualized online plan verification and its influence on 6MV x-ray beam characteristics. *Z Med Phys* 2016;26(3):200–8.
- [15] Nakaguchi Y, Ono T, Maruyama M, Shimohigashi Y, Kai Y. Validation of a method for in vivo 3D dose reconstruction in SBRT using a new transmission detector. *J Appl Clin Med Phys* 2017;18(4):69–75.
- [16] Vikraman S, Manigandan D, Karrthick K, Sambasivasselli R, Senniandavar V, Ramu M, et al. Quantitative evaluation of 3D dosimetry for stereotactic volumetric-modulated arc delivery using Compass. *J Appl Clin Med Phys* 2014;16(1):192–207.
- [17] Nakaguchi Y, Araki F, Ono T, et al. Validation of a quick three-dimensional dose verification system for pre-treatment IMRT QA. *J Radiol Phys Technol* 2014;8(1):73–80.
- [18] Feygelman V, Forster K, Opp D, et al. Evaluation a biplanar diode array dosimeter for quality assurance of step-and-shoot IMRT. *J Appl Clin Med Phys* 2009;10(4):64–78.
- [19] Sadagopan R, Bencomo JA, Martin RL, Nilsson G, Matzen T, Balter PA. Characterization and clinical evaluation of a novel IMRT quality assurance system. *J Appl Clin Med Phys* 2009;10(4):104–19.
- [20] Bedford JL, Lee YK, Wai P, South CP, Warrington AP. Evaluation of the Delta<sup>4</sup> phantom for IMRT and VMAT verification. *Phys Med Biol* 2009;54:N167–76.
- [21] Arumugan S, Xing A, Ypung T, Thwaites D, Holloway L. Comparison of three commercial dosimetric systems in detecting clinically significant VMAT delivery errors. *Phys Med* 2016;32:1238–44.
- [22] Du W, Cho SH, Zhang X, Hoffman KE, Kudchadker RJ. Quantification of beam complexity in intensity-modulated radiation therapy treatment plans. *Med Phys* 2014;41(2):21716.
- [23] Vieilleveigne L, Molinier J, Brun T, Ferrand R. Gamma index comparison of three VMAT QA systems and evaluation of their sensitivity to delivery errors. *Phys Med* 2015;31(7):720–5.
- [24] Millin AE, Windle RS, Lewis DG. A comparison of electronic portal dosimetry verification methods for use in stereotactic radiotherapy. *Phys Med* 2016;32(1):188–96.
- [25] Stasi M, Bresciani S, Miranti A, Maggio A, Sapino V, Gabriele P. Pretreatment patient-specific IMRT quality assurance: a correlation study between gamma index and patient clinical dose volume histogram. *Med Phys* 2012;39:7626–34.
- [26] Kim JJ, Park SY, Kim HJ, Kim JH, Ye SJ, Park JM. The sensitivity of gamma-index method to the positioning errors of high-definition MLC in patient-specific VMAT QA for SBRT. *Radiat Oncol* 2014;28(9):167.
- [27] Waghorn BJ, Meeks SL, Langen KM. Analyzing the impact of intrafraction motion: Correlation of different dose metrics with changes in target D95%. *Med Phys* 2011;38(8):4505–11.
- [28] Rajasekaran D, Jeevanandam P, Sukumar P, Ranganathan A, Johnjothi S, Nagarajan V. A study on the correlation between plan complexity and gamma index analysis in patient specific quality assurance of volumetric modulated arc therapy. *Rep Pract Oncol Radiother* 2015;20(1):57–65.
- [29] Hussein M, Clark CH, Nisbet A. Challenges in calculation of the gamma index in radiotherapy – towards good practice. *Phys Med* 2016;36:1–11.