Acta Oncologica, 2014; 53: 628–636

informa
healthcare

ORIGINAL ARTICLE

Dosimetric inter-institutional comparison in European radiotherapy centres: Results of IAEA supported treatment planning system audit

EDUARD GERSHKEVITSH[1], CSILLA PESZNYAK[2], BORISLAVA PETROVIC[3], JOSEPH GREZDO[4], KRZYSZTOF CHELMINSKI[5], MARIA DO CARMO LOPES[6], JOANNA IZEWSKA[7] & JACOB VAN DYK[7,8]

[1]*North Estonia Medical Centre, Department of Radiotherapy, Tallinn, Estonia,* [2]*National Institute of Oncology, Budapest, Hungary,* [3]*Institute of Oncology Vojvodina, Sremska Kamenica, Serbia,* [4]*St. Elizabeth Institute of Oncology, Bratislava, Slovakia,* [5]*M. Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland,* [6]*Portuguese Institute of Oncology, Coimbra, Portugal,* [7]*International Atomic Energy Agency, Vienna, Austria and* [8]*Western University, London, Ontario, Canada*

Abstract

Background and purpose. One of the newer audit modalities operated by the International Atomic Energy Agency (IAEA) involves audits of treatment planning systems (TPS) in radiotherapy. The main focus of the audit is the dosimetry verification of the delivery of a radiation treatment plan for three-dimensional (3D) conformal radiotherapy using high energy photon beams. The audit has been carried out in eight European countries – Estonia, Hungary, Latvia, Lithuania, Serbia, Slovakia, Poland and Portugal. The corresponding results are presented. *Material and methods.* The TPS audit reviews the dosimetry, treatment planning and radiotherapy delivery processes using the 'end-to-end' approach, i.e. following the pathway similar to that of the patient, through imaging, treatment planning and dose delivery. The audit is implemented at the national level with IAEA assistance. The national counterparts conduct the TPS audit at local radiotherapy centres through on-site visits. TPS calculated doses are compared with ion chamber measurements performed in an anthropomorphic phantom for eight test cases per algorithm/beam. A set of pre-defined agreement criteria is used to analyse the performance of TPSs. *Results.* TPS audit was carried out in 60 radiotherapy centres. In total, 190 data sets (combination of algorithm and beam quality) have been collected and reviewed. Dosimetry problems requiring interventions were discovered in about 10% of datasets. In addition, suboptimal beam modelling in TPSs was discovered in a number of cases. *Conclusions.* The TPS audit project using the IAEA methodology has verified the treatment planning system calculations for 3D conformal radiotherapy in a group of radiotherapy centres in Europe. It contributed to achieving better understanding of the performance of TPSs and helped to resolve issues related to imaging, dosimetry and treatment planning.

Quality assurance (QA) in all aspects of radiotherapy reduces the number of instances of accidental dose misadministration. Reduction of such errors and uncertainties plays an important role in the outcome of radiotherapy [1]. It is generally agreed that, based on clinical dose-response curves, the overall accuracy of the dose delivery should be better than 5% [2,3].

The International Atomic Energy Agency (IAEA) has a long standing history of conducting dosimetry audits in radiotherapy using thermoluminescence dosimeters (TLD). TLD audit to verify the radiotherapy beam output in reference conditions was started in 1969 [4]. Since then, the dosimetry audits have gradually evolved from the basic beam calibration audit that is the most essential step, by including other beam parameters [5], progressing to complex treatment techniques, verifying treatment planning and delivery [6], and finally auditing the complete process of radiation treatment using a comprehensive audit within the framework of Quality Assurance Team for Radiation Oncology (QUATRO) [7]. These

Correspondence: E. Gershkevitsh, North Estonia Medical Centre, Department of Radiotherapy, J. Sütiste tee 19, 13419 Tallinn, Estonia. Tel: + 372 6172451. Fax: + 372 6172460. E-mail: eduardger@yahoo.co.uk

(Received 15 July 2013; accepted 26 August 2013)

ISSN 0284-186X print/ISSN 1651-226X online
DOI: 10.3109/0284186X.2013.840742

developments bring the dosimetry audit scope closer to the patient treatment setting and therefore have the potential to increase possible benefits. The IAEA, based on its Technical Report Series 430 [8] has developed a Technical Document (TECDOC 1583, [9]) outlining a set of practical clinical tests for treatment planning systems (TPSs) to help the users to verify the dosimetry part of the TPS. This methodology was successfully tested in a number of different centres and forms the foundation of the IAEA TPS audit programme [6]. The TPS audit reviews the dosimetry, treatment planning and radiotherapy delivery processes in radiotherapy centres using an 'end-to-end' approach, i.e. following the pathway similar to that of the patient, through imaging, treatment planning and dose delivery.

The main focus of the audit is the dosimetry verification of the delivery of a radiation treatment plan for three-dimensional conformal radiotherapy (3D-CRT) using high energy photon beams. The audit has been carried out in eight European countries – Estonia, Hungary, Latvia, Lithuania, Serbia, Slovakia, Poland and Portugal. The data presented here reflect the status of radiotherapy physics practices in participating centres and provide information on the accuracy of dose delivery that is realistically achievable in these centres.

Material and methods

Audit operation

The audit was implemented at the national level with IAEA assistance. The IAEA provides equipment (a semi-anthropomorphic dosimetry phantom) and the detailed audit methodology (based on the IAEA TECDOC 1583 [9]), as well as expert services to help the national auditing organisation to introduce the TPS audit in a particular country. The auditing organisation was usually a medical physics group or a department of a recognised radiotherapy centre nominated by the national authorities or formally endorsed by the medical physics society in the country. At first, the local dosimetry equipment was checked against the auditor equipment and then the audit was performed by the IAEA expert in a radiotherapy department of the national auditing organisation. This was used to transfer the audit methodology to the national level and to ensure that it would be consistently applied in other centres in the country. Before starting the audit programme in the country, a one-day workshop was organised for medical physicists from different radiotherapy centres to present the methodology and discuss the audit operation and logistics. Following the workshop, the auditing organisation conducted the TPS audit at local radiotherapy centres through on-site visits. The participation of the centres in the audit was voluntary. The audit required about 12–14 hours per centre including about three hours of a linac time.

Phantom and its computed tomography (CT) imaging

For the measurements of clinical test cases, a commercially available semi-anthropomorphic CIRS Thorax (CIRS Inc., Norfolk, VA, USA) phantom has been used. The phantom consisted of a thoracic body made of plastic water-equivalent material (plastic-to-water relative electron density of 1.003), lung-equivalent material (relative electron density 0.207) and bone-equivalent material (relative electron density 1.506) with 10 holes to hold interchangeable rod inserts for an ionisation chamber. The position of the holes was labelled, as shown in Figure 1. The phantom was supplemented with a set of five electronic density reference plugs simulating muscle, bone, dense bone, lung and adipose equivalent tissue.

The phantom was CT scanned twice in participating centres. For the first scan, the relative electron density reference plugs were inserted to obtain the CT number to relative electron density (CT-to-RED) conversion curve. The tolerance value of ± 20 HU was used throughout the study [8]. For the second scan the reference plugs were replaced with the corresponding material plugs and it was used for the planning of clinical test cases. The local scanning protocol was used and the scanning parameters for both scans were kept the same. The local CT-to-RED conversion curve was kept for calculations.

Clinical test cases

CT images of the phantom in the DICOM format have been transferred to the TPS either through the local network or via the digital media. A set of clinical test cases was created to verify a range of basic treatment

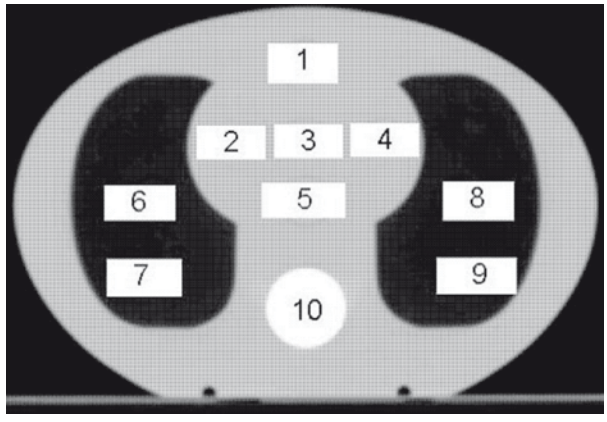

Figure 1. Position of measurement points in CIRS Thorax phantom.

techniques applied in the clinical practice. The test cases, reference and measurement points have been described elsewhere [9] and for clarity are given in Table I. One measurement point for each test was used as the reference point. The outline of the holes and structures created around them (test 5) were used to set-up the isocentre/measurement point location as well as shape of the aperture. The locations of measurement points for each test case were selected such that high dose gradients and measurements in the penumbra region were avoided. The total number of measurement points in eight test cases was 15. The same set of clinical test cases was applied in all participating centres. Test cases were planned with the TPS and the number of monitor units or time to deliver the prescribed dose of 2 Gy to the reference point was calculated. The dose calculations were performed based on the grid size normally used in the centre's clinical practice. The TPS calculation algorithms in this study were all model based and they were grouped similarly to Knöös et al. [10] as follows:

> Type (a) Model-based algorithms, which use pencil beam convolution model and primarily equivalent path length corrections to account for heterogeneities. Effects of lateral electron and photon transport are not modelled (no lateral transport);
> Type (b) Model-based algorithms, which primarily use point kernel convolution/superposition model and account for density variation in 3D. Changes in lateral electron and photon transport are approximately modelled (with lateral transport).

List of TPS makes and algorithms is given in Table II.

Table I. Description of clinical test cases.

Description of test cases	Test case no.	Reference point		Measurement point/field	Agreement criteria (%)
Single field: SSD = 100 cm, field size 10×10 cm^2, gantry 0°, coll. 0°				3	2
Points 1,3,5 and 10 on central axis, point 9 outside the field	1	3		9	4
				10	3
Tangential field, oblique incidence and lack of scattering: SAD set-up at point 1, field size 15×10 cm^2, gantry 90°, coll. 90°, wedge 45°	2	1		1	3
Significant blocking of the field corners: SSD = 100 cm, field size 14×14 cm^2 blocked to a 10×10 cm^2, gantry 0°, coll. 45°	3	3		3	3
				F1: 0°	2
				F2: 90°	3
			5	**F3: 180°**	3
				F4: 270°	3
				Σ	3
				F1: 0°	4
Four field box: SAD set-up at point 5; Field1: field size 15×10 cm^2, gantry 0°, coll. 0°; Field2: field size 15×8 cm^2, gantry 90°, coll. 0°; Field3: field size 15×10 cm^2, gantry 180°, coll. 0°; Field4: field size 15x8 cm^2, gantry 270°, coll. 0°	4	5	6	**F2: 90°**	3
				F3: 180°	4
				F4: 270°	3
				Σ	4
				F1: 0°	3
				F2: 90°	4
			10	**F3: 180°**	3
				F4: 270°	4
				Σ	4
Customised blocking of cylinder Ø 10 cm: SAD set-up at point 2, gantry 0°, coll.0°	5	2		2	3
				7	4
Oblique incidence with irregular L-shaped field (blocking off the centre of the field): SAD set-up at point 5, field size 20×10 cm^2 (blocked to 12×6 cm^2), gantry 45°, coll. 0°	6	3		3	3
				7	5
				10	5
Plan with **asymmetrically wedged fields:** SAD set-up at point 3; Field1: field size 10×12 cm^2, gantry 0°, coll. 0°; Field2: field size 10×6 cm^2 asymmetrical (half beam), gantry 90°, coll. 90°; physical wedge 30°; Field3: field size 10×6 cm^2 asymmetrical (half beam), gantry 270°, coll. 270°; soft wedge 30°	7	5	5	**F1: 0°**	2
				F2: 90°	4
				F3: 270°	4
				Σ	3
Plan with **non-coplanar field:** SAD set-up at point 5; Field1: field size 4×4 cm^2, gantry 30°, coll.0°, couch 270°; Field2: field size 4×16 cm^2, gantry 90°, coll. 330° couch 0°; Field3: field size 4×16 cm^2), gantry 270°, coll. 30°, couch 0°	8	5	5	**F1: 0°**	3
				F2: 90°	3
				F3: 270°	3
				Σ	3

Tests 3,5 and 6 involves blocking of the rectangular field either by block or MLC (whichever was available and used for patient treatment).

Table II. List of TPSs.

TPS manufacturer	TPS model	Distribution of TPS (%)	Type (a) algorithm	Type (b) algorithm
Elekta	XiO	35	Fast Fourier Transformation Convolution, Clarkson	Multi Grid Superposition, Fast Multi Grid Superposition
Elekta	Oncentra Masterplan	14	Pencil Beam	Collapsed Cone
Elekta	Plato	3	Pencil Beam	
Elekta	Precise Plan	4	Pencil Beam	
Philips	Pinnacle	5	Pencil Beam	Collapsed Cone
Siemens	Prowess Panther	2	Pencil Beam	Collapsed Cone
Varian	Eclipse	37	Pencil Beam Convolution	Anisotropic Analytical Algorithm

Measurements

Measurements were performed in 60 centres using different makes and models of clinical accelerators and Co-60 units listed in Table III. Photon beams were divided according to the energy into three groups: Co-60 beams, lower energy x-ray (4–6 MV) and higher energy x-ray (10–20 MV) beams.

Farmer type ionisation chambers and reference class electrometers with the calibration traceable to the national or international secondary standards laboratories were employed throughout the audit. All measured doses were calculated using the IAEA TRS 398 code of practice [11] and reported in terms of the absorbed dose to water. When measuring in lung and bone equivalent materials, it was assumed that the doses were measured in small water cavities within these materials.

In addition to the measurements, relative output factors and wedge factors in TPSs were checked against the datasets published in the literature [12,13]. According to published data [12] the mean wedge transmission data for the manufacturer's standard wedges on most makes and models of accelerators exhibited a Gaussian distribution with a standard deviation of $\pm 2\%$, therefore a $\pm 4\%$ action level corresponding to two standard deviations (SD) of the mean value was chosen to evaluate the outliers for both relative output factors and wedge factors.

Analysis of the results

For evaluation of the measured (D_{meas}) and TPS calculated (D_{cal}) values the same criteria as specified in the IAEA TRS 430 were employed [8]. However, due to the limited number of available positions for the dose measurements in the CIRS phantom, dose differences were normalised to the dose measured at the reference point for each test case (rather than to the measurement point – local difference), i.e. the following equation was used:

$$\Delta[\%] = 100 * \left(\frac{D_{cal} - D_{meas}}{D_{meas,ref}} \right) \quad (1)$$

where $D_{meas,ref}$ was the dose value measured at the reference point. The agreement criteria for each test case are listed in Table I and they depend on the complexity of the test case geometry.

Results

Participating countries

The percentage of audited centres was on average 64% (range 29–100%) and included the major centres in the countries (Table IV). The highest participation rate was in Estonia and Portugal and the lowest in Poland and Latvia. The low number of audited centres in Poland is due to the fact that the audit programme was planned for an extended time period and it is still going on in the country. The reason that some other centres have not participated in the audit was mostly related to the lack of TPS in the radiotherapy department or facilities for 3D-CRT at those institutions, in particular CT scanners. On average, one third of equipment in participating

Table III. List of clinical accelerators and Co-60 units.

Treatment unit manufacturer	Treatment unit model	Energy (MV)	Distribution of beams (%)
Baltiets	Agat-R1	Co-60	2
Elekta	Precise, Synergy	4, 6, 10, 15, 18	28
Siemens	Mevatron, Primus, Oncor, Artiste	6, 10, 15, 18	25
Best Medical	Theratronics 780	Co-60	1
UJP	Terabalt	Co-60	1
Varian	Clinac 600, 2100/2300 C/D, DHX, iX	4, 6, 10, 15, 18, 20	44

countries was older than 10 years, although it varied widely between countries (range 8–51%). 2D treatments presented in the Table IV were characterised by the manual calculations without using blocks/MLCs, i.e. treatments with rectangular beams. 3D-CRT was the most common treatment approach in participating centres with the average percentage of treatment plans being 83%. Majority of the treatments were done with curative intent (average 70%). Most prevalent anatomical sites treated were breast, prostate, lung, and head and neck. The percentage of treatment plans in the participating centres where the dose calculation in the lung was required (breast, lung) was on average 33.3% (range 30–38%).

Measurement results

All data collected in the audit were analysed and feedback was provided to the centre through the auditing organisation. The results collected at the national level were sent to the global international database maintained by the IAEA.

About two thirds of the centres used the generic or TPS manufacturer supplied CT-to-RED conversion curves. In the majority of cases ($>95\%$) the difference was only seen in denser materials which have been estimated to have an impact of less than 2% on the dose calculation [9]. However, in several test cases the difference was also seen in the low density material. In the test case 5 point 7 (customised block, lung material) the difference between the measurements and the dose calculation due to the inaccuracies in the electron density conversion exceeded 4%. Simulator-based cone beam CT scans were used in one centre without proper CT-to-RED conversion curve, which resulted in all of the measurement points inside the lung being outside the agreement criteria.

One hundred and ninety datasets (combination of algorithm and beam quality) have been collected in 60 centres over the period of 2007–2012. The percentage differences between the measured and calculated doses for the various measurement points and test cases are presented in Figure 2. The results are grouped by the calculation algorithm type and the mean relative difference between the calculated and measured doses $\Delta\%$) is given with the error bars corresponding to two standard deviations. The value of the agreement criteria for each measurement point is shown as a thick black line. As expected, the mean Δ and SD are smaller for type (b) algorithms (mean 0.5%, 2SD 3.3%) when compared to type (a) algorithms (mean 0.8%, 2SD 5.6%) [9]. The differences are particularly pronounced for points within the lung material (test case 1 point 9; test case 4 point 6; test case 5 point 7; test case 6 point 7). The largest deviations observed in the study were in the customised block test case 5 and point 7, which is located within lung equivalent material. The differences up to 17% were discovered for some of type (a) algorithms within the group of higher energy x-ray beams.

The histogram in Figure 3 summarises the results for different algorithm types and the beam energy groups. For Co-60 beams, type (a) model-based dose calculation algorithms generated major discrepancies between the calculations and the measurements. Many centres using older Co-60 units were equipped with TPS with type (a) algorithm only. In general, for linac beams the number of measurements with results outside the agreement criteria was higher for higher beam energies independent of the algorithm used.

Various dosimetry problems were identified in approximately 10% of datasets. The largest source of discrepancies was the TPS input data and inaccuracies in beam modelling (50%), and discrepancies in the beam calibration (39%). Various mechanical problems on treatment units accounted for about 11% of discrepancies. The output and wedge factor checks during the audit helped to identify some

Table IV. Overview of radiotherapy infrastructure in eight participating countries at the time of the audit.

Country	Number of participating institutions	% of audited/total centres in the country	% of audited/total beams in the country	Number of treatment units per centre	% of treatment units being older than 10 years	Medical Physicists per treatment unit	% of 2D treatments in participating institutions	% of 3D treatments in participating institutions	% of IMRT treatments in participating institutions
Estonia	2	100%	67%	1.5	33%	2.0	5%	95%	0%
Hungary	9	75%	26%	2.9	51%	1.6	24%	75%	1%
Latvia	1	33%	36%	2.3	43%	1.7	10%	85%	5%
Lithuania	2	50%	36%	2.3	44%	1.8	10%	90%	0%
Poland	8	29%	15%	3.9	10%	2.3	0%	90%	10%
Portugal	24	100%	62%	1.8	11%	1.3	2%	93%	5%
Serbia	3	50%	23%	2.2	8%	2.3	20%	80%	0%
Slovakia	11	79%	46%	1.9	31%	1.9	23%	75%	2%

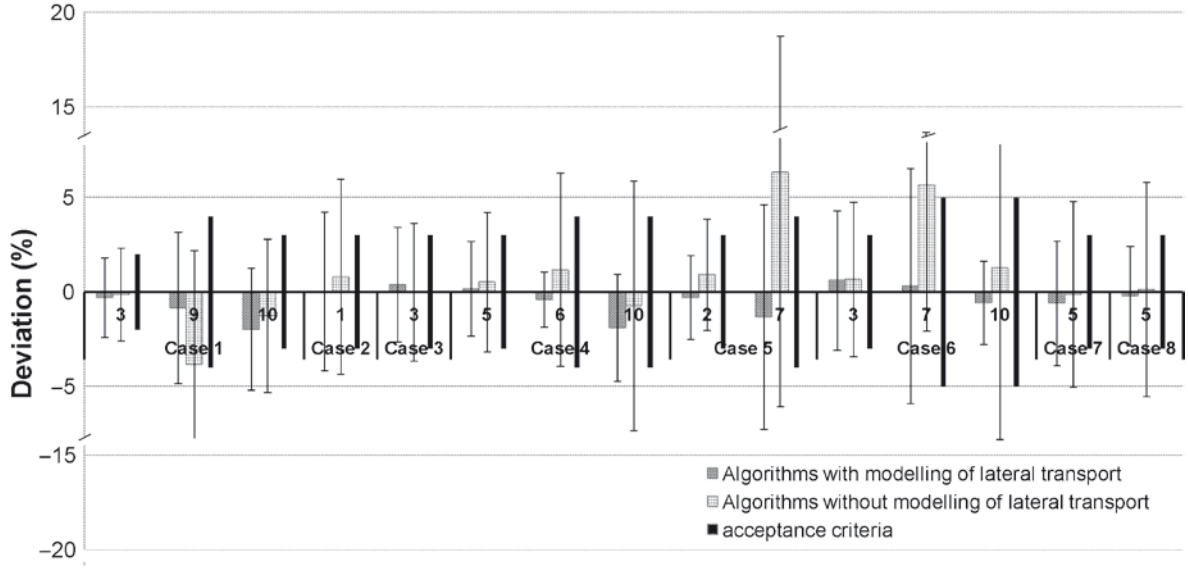

Figure 2. Percentage difference between the measured and the calculated doses for the different test cases [dark grey – Algorithms with modelling of lateral transport (average data of 100 datasets), light grey – Algorithms without modelling of lateral transport (average data of 90 datasets), black – acceptance criteria; error bars represent 2SD].

typographical errors in the input beam data. Figure 4 show the percentage of measurement results outside the agreement criteria in 12 audited centres that used similar combination of the TPS algorithm and the beam data, i.e. Varian Eclipse TPS with anisotropic analytical algorithm (AAA) algorithm and low and high energy x-ray beams generated by a dual energy Varian accelerators. In Figure 4, the measurement data have been normalised to the reference point 3 (test case 1) to focus only on the local TPS implementation and exclude from the analysis any impact of the machine output variation or the beam calibration problems (N.B. this was not done for the data of Figures 2 and 3). As can be seen, the local implementation of the TPS, including the input beam data and beam modelling, could lead to deviations up to 39% in a number of failing measurements points. In the analysed data the high energy beams showed less variation on average when compared to low energy beams in the same institutions.

Discussion

The majority of centres participating in the audit were from Central and Eastern Europe where various inadequacies in radiotherapy infrastructure have been identified [14]. These include insufficient numbers of radiotherapy equipment per cancer

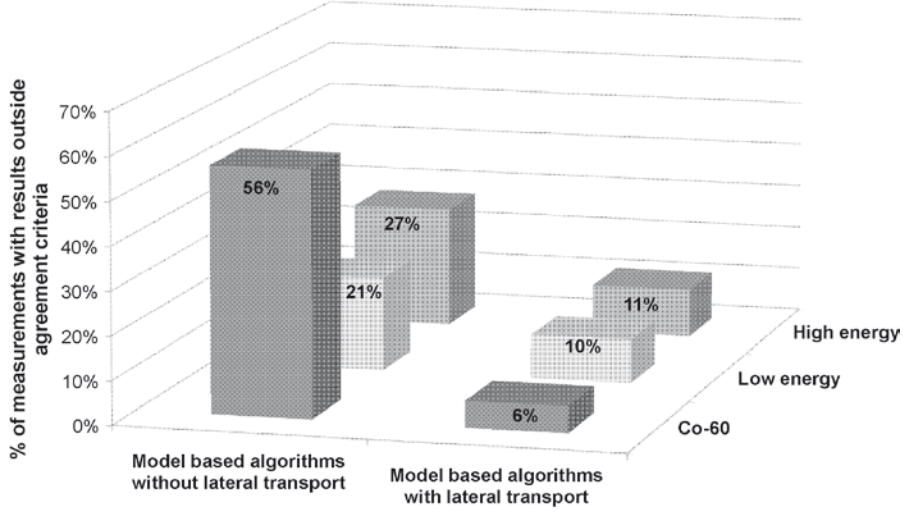

Figure 3. Percentage of measurement points with results outside agreement criteria for different algorithm types and energy groups.

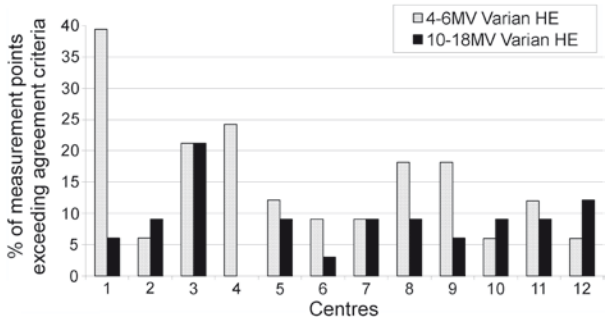

Figure 4. Percentage of measurement points with results outside agreement criteria for Varian Eclipse AAA algorithm, 4–6 MV photon beams (grey) and 10–18MV photon beams (black) from high energy Varian accelerators (data normalised to case 1 point 3).

incidence resulting in high patient throughput on treatment units and, in addition, ageing radiotherapy equipment in some countries. Also, the dosimetry equipment and TPSs in those centres that have older treatment units are usually older models using simpler dose calculation algorithms. The percentage of deviations was somewhat higher in centres with older equipment than in those with the newer equipment. As well, no significant correlations have been found between the deviations in TPS audit results and patterns of radiotherapy infrastructure including staffing levels.

Following the findings of the TPS audit, there was a change in practice in those centres that had been using type (a) calculation algorithms before the audit and had a choice to use different algorithms [15]. CT-to-RED calibration curves have been updated in some centres which improved the agreement between measured and calculated doses. This has been confirmed through follow-up audits in some centres where deviations occurred. The follow-up visits were performed in those centres where resolution of the problems was taken more time than the auditor visit duration (typically 1–2 days). One centre did not use the inhomogeneity correction for treatment planning, but started to apply it after the results of the audit demonstrated the magnitude of deviations in the dose calculation. By appreciating the measurement results, the more accurate dose calculation algorithms have been employed for treatment planning, in particular for anatomical sites involving lung tissue.

Dosimetric inter-institution comparison using an anthropomorphic phantom is a useful approach to verify the treatment planning and delivery of specific radiotherapy techniques for clinical trials [15–17]. These studies are helpful in identifying various deviations between the planned and delivered doses that can lead to improvements in treatment planning practices in radiotherapy centres. The range of dose deviations in the different centres performing 3D conformal radiotherapy observed in this study reflects upon the relative dosimetric accuracy of the treatment planning process from the CT scanning to the dose delivery. Deviations between the measurement and the calculation may arise for a number of reasons. These include: 1) inaccuracies in TPS beam data input, 2) suboptimal beam modelling, 3) limitations of dose calculation algorithms, 4) dosimetry errors in the calibration of treatment beams, 5) differences in CT-to-RED calibration curves, and 6) various other reasons. Some deviations are related to systematic errors associated with the algorithm limitations, while others may indicate the deficiencies in the different parts of the treatment planning process and the dose delivery chain. The differences between algorithms have been extensively described in the literature [18–20]. The number of measurements with results outside the agreement criteria and the magnitude of dose differences between the calculated and the measured dose values for each test case are lower for more advanced algorithms which is in line with the previous publication [6]. Most of the large deviations ($>5\%$) occurred in the lung equivalent material which is similar to the findings by other authors [18,19]. The results confirmed the inadequacy of the type (a) algorithms to manage the dose calculation in the presence and inside the low density heterogeneities. Type (b) algorithms showed better performance in the applied test cases with most results within the specified agreement criteria. The dose difference inside the bone equivalent material was underestimated on average by 3% for high energy x-ray beams for all algorithm types. Similar results for some TPSs as used in this study were found in the study performed by Carrasco et al. [20].

However, the main advantage of the independent audit is to identify problems that are not related to well-known limitations of the calculation algorithm. Problems in the local TPS implementation including inaccurate TPS input data and poor quality beam modelling were the greatest contributors to faulty results obtained in this study. Errors in machine calibrations and the inaccurate calibration dose value in the TPS was another major contributor to observed deviations. Although none of the differences in the beam calibration exceeded 5% (current IAEA threshold in postal TLD audit), the fact that there was a deviation in the reference dose (point 3 of test case 1) by more than 2% had an impact on the results for all other tests. While typographical errors may be limited to a particular field size range or a wedge, the beam modelling problems could have influence on the dose calculation for every field (e.g. if it involves incorrect energy spectrum). As can be seen from Figure 4, errors in input beam data

and suboptimal beam modelling could result in as much as six times more results being outside the agreement criteria than for more optimal beam data representation in TPS. In this subset of data for one particular manufacturer the number of measurement points with results outside agreement criteria are higher for lower energy beams than for higher energy beams which could be partly attributed to the fact that beam modelling of lower energies may be more prone to discrepancies due to steeper depth dose curves and penumbras than higher energies. Similar findings for the same type of accelerators and the same TPS make and algorithm were observed in the study by Breitman et al. [21] where they compared the results of measurements in the water phantom with TPS calculations based on the "golden beam data" sets. Re-commissioning of the TPS with carefully measured data had improved the agreement between calculated and measured doses in their case, which also emphasises the importance of proper input data and beam modelling process. This would be even more pronounced if the same model was used for intensity-modulated radiotherapy (IMRT). In addition, some problems related to treatment units were detected such as a loose wedge and treatment couch table top not taken into account in the dose calculation, which resulted in up to 8% lower dose than expected for 180 degree beam (test case 4). The subset of results presented in Figure 4 showed an opposite trend to the global findings where the number of measurement points with results outside agreement criteria was higher for high energy beams (Figure 3). In the global results the proportion of "problematic" datasets were smaller, therefore the findings presented in Figure 3 could be partially attributed to the fact that the ranges of secondary particles created by high energy x-ray beams are larger and therefore integration of dose contribution by these secondary particles may be more problematic for the current dose calculation algorithms when compared with low energy beams.

Another finding of the study was that the largest number of measurement points with results outside the agreement criteria was for Co-60 gamma beams. More detailed analysis of the data suggested that those were primarily related to the inaccuracies of the beam calibration and various problems related to treatment units. This is in line with IAEA/WHO TLD audit observations that a larger number of dose audit results falling outside the 5% acceptance limit is for Co-60 gamma beams whereas x-ray beams from linacs are generally calibrated more accurately [22].

Overall, for the particular phantom and methodology employed in the current study, the majority of participating centres that used type (b) calculation algorithms have demonstrated the ability to deliver and report the dose correctly to within 2–5% depending on the complexity of the treatment plan. This finding is similar to the uncertainties estimated as practically achievable by AAPM and others [23,24]. However, in realistic clinical situations where not a dosimetry phantom, but a patient is irradiated, the overall uncertainty in the dose delivery is likely to be higher due to a number of reasons including geometrical uncertainties resulting from the patient breathing and organ motion, set-up reproducibility, etc. To evaluate the overall uncertainty in the radiotherapy dose delivery process including both dosimetric and geometrical components, more complex audit methodology would be needed to design which should also include tests for image guidance system and phantom movement simulating patient breathing, phantom positioning reproducibility as well as measurements of the dose distribution, particularly for more advanced techniques such as IMRT.

One of the advantages of the current TPS audit approach is that all audit steps involving CT imaging, planning and measurements are closely followed at the local centre because the auditor is present on-site and ensures the consistent and accurate execution of the audit procedure among the participating centres. Another advantage is that the measurements with ionisation chamber provide real-time reading and discrepancies that occur in the measurements can be resolved while auditor is still on-site. However, some of the audit aspects related to the dose distribution measurement could be better explored with other dosimetry systems such as TLD in conjunction with films used in the remote audit [16], but the process of resolving discrepancies in the remote audit is more difficult than during on-site audit.

As the rate of introduction of new equipment and treatment methods is growing, there is a need to continue the independent verification of radiotherapy planning and delivery using the end-to-end approach, as described in this paper. As delivery methods are becoming more complex there may be a larger number of centres that would not be able to pass the initially set criteria [16]. However, conducting similar dosimetry audits and repeating them over the time have been shown to improve the consistency between the measured and calculated doses and to increase the confidence in the new dose planning and delivery methods [16,25].

In summary, the methodology for TPS audit described in the IAEA TECDOC 1583 [9] has been applied in 60 radiotherapy centres in Europe. The range of observed dose deviations was presented and discussed. Errors have been identified in nearly 10% of datasets and corrective actions have been taken.

Most of errors were related to the beam data input to TPS and suboptimal beam modelling as well as inaccurate beam output dosimetry.

Following the findings of the audits, several radiotherapy centres have improved their treatment planning procedures by employing type (b) algorithms for the dose calculation in lung tissue rather than using type (a) algorithms that produce larger discrepancies. Overall, the IAEA supported national TPS audit has contributed to achieving better understanding of the performance of TPSs and their limitations.

Acknowledgements

Special thanks to Mr. Stanislav Vatnitsky who drafted a proposal for this TPS audit project. To Technical Cooperation Department of IAEA in particular, Mrs. Sandra Steyskal (RER6018), Mr. Ivan Videnovic (POL6009) and Mr. Arnaud Atger (EST6004) for their support during the project implementation. To the physicists of the different radiotherapy departments who helped with the measurements and in particular to Mr. Sergey Popov (Latvian Oncology Centre), Mr. Albert Miller (Lithuanian Institute of Oncology) and Mr. Antanas Vaitkus (Kaunas Medical University Hospital, Lithuania).

Declaration of interest: J. Van Dyk has a license agreement with Modus Medical Devices, Inc. for the development and sale of QUASAR Phantoms.

References

[1] World Health Organization (WHO). Radiotherapy risk profile. Geneva: WHO; 2008.

[2] International Commission on Radiological Units (ICRU) Report 24. Determination of absorbed dose in a patient irradiated by beams of X or gamma rays in radiotherapy procedures. Washington DC: ICRU; 1976.

[3] Pettersen MN, Aird E, Olsen DR. Quality assurance of dosimetry and the impact on sample size. Radiother Oncol 2008;86:195–9.

[4] Izewska J, Andreo P. The IAEA/WHO TLD postal programme for radiotherapy hospitals. Radiother Oncol 2000;54:65–72.

[5] Izewska J, Georg D, Bera P, Twaites D, Arib M, Saravi M, et al. A methodology for TLD postal dosimetry audit of high-energy radiotherapy photon beams in non-reference conditions. Radiother Oncol 2007;84:67–74.

[6] Gershkevitsh E, Schmidt R, Velez G, Miller D, Korf E, Yip F, et al. Dosimetric verification of radiotherapy treatment planning systems: Results of IAEA pilot study. Radiother Oncol 2008;89:338–46.

[7] International Atomic Energy Agency (IAEA) comprehensive audits of radiotherapy practices: A tool for Quality Improvement, Quality Assurance Team for Radiation Oncology (QUATRO), Vienna: IAEA; 2007.

[8] IAEA Technical Report Series 430. Commissioning and quality assurance of computerized planning systems for radiation treatment of cancer. Vienna: IAEA; 2005.

[9] IAEA TECDOC 1583. Commissioning of radiotherapy treatment planning systems: Testing for typical external beam treatment techniques. Vienna: IAEA; 2008.

[10] Knöös T, Wieslander E, Cozzi L, Brink C, Fogliata A, Albers D, et al. Comparison of dose calculation algorithms for treatment planning in external photon beam therapy for clinical situations. Phys Med Biol 2006;51:5785–807.

[11] IAEA Technical Report Series 398. Absorbed dose determination in external beam radiotherapy. An international code of practice for dosimetry based on standards of absorbed dose to water. Vienna: IAEA; 2000.

[12] Lowenstein J, Kry S, Molineu A, Alvarez P, Aguirre J, Summers P, et al. High-energy photon standard dosimetry data: A quality assurance tool. Med Phys 2012;39:3754.

[13] Followill DS, Kry SF, Qin L, Lowenstein J, Molineu A, Alvarez P, et al. The Radiological Physics Center's standard dataset for small field size output factors. J App Clin Med Phys 2012;13:282–9.

[14] Rosenblatt E, Izewska J, Anacak Y, Pynda Y, Scalliet P, Boniol M, et al. Radiotherapy capacity in European countries: An analysis of the Directory of Radiotherapy Centres (DIRAC) database. Lancet Oncol 2013;14: e79–86.

[15] Rutonjski L, Petrović B, Baucal M, Teodorovic M, Cudic O, Gershkevitsh E, et al. Dosimetric verification of radiotherapy treatment planning systems in Serbia: National audit. Radiat Oncol 2012;7:155.

[16] Molineu A, Hernandez N, Nguyen T, Ibbott G, Followill D. Credentialing results from IMRT irradiations of an anthropomorphic head and neck phantom. Med Phys 2013; 40:022101.

[17] Lopes MC, Cavaco A, Jacob K, Madureira L, Germano S, Faustino S, et al. Treatment planning systems dosimetry auditing projects in Portugal. Phys Med Epub 2013 Apr 23.

[18] Carrasco P, Jornet N, Duch MA, Weber L, Ginjaume M, Eudaldo T, et al. Comparison of dose calculation algorithms in phantoms with lung equivalent heterogeneities under conditions of lateral electronic disequilibrium. Med Phys 2004;31:2899–911.

[19] Venselaar JLM, Welleweerd J. Application of a test package in an intercomparison of the performance of treatment planning systems used in a clinical setting. Radiother Oncol 2001;60:203–13.

[20] Carrasco P, Jornet N, Duch MA, Weber L, Ginjaume M, Eudaldo T, et al. Comparison of dose calculation algorithms in slab phantoms with cortical bone equivalent heterogeneities. Med Phys 2007;34:3323–33.

[21] Breitman K, Rathee S, Newcomb C, Murray B, Robinson D, Field C, et al. Experimental validation of the Eclipse AAA algorithm. J App Clin Med Phys 2007;8:76–92.

[22] Izewska J, Andreo P, Vatnitsky S, Shortt KR. The IAEA/WHO TLD postal dose quality audits for radiotherapy: A perspective of dosimetry practices at hospitals in developing countries. Radiother Oncol 2003;69:91–7.

[23] American Association of Physicists in Medicine (AAPM) Report 85. Tissue inhomogeneity correction for megavoltage photon beams. Madison: AAPM; 2004.

[24] Venselaar J, Welleweerd H, Mijnheer B. Tolerances for the accuracy of photon beam dose calculations of treatment planning systems. Radiother Oncol 2001;60:191–201.

[25] Palmer A, Mzenda B, Kearton J, Wills R. Analysis of regional radiotherapy dosimetry audit data and recommendations for future audits. Br J Radiol 2011;84:733–42.