Acta Oncologica

Testing the methodology for a dosimetric end-to-end audit of IMRT/VMAT: results of IAEA multicentre and national studies

Paulina Wesolowska, Dietmar Georg, Wolfgang Lechner, Pavel Kazantsev, Tomislav Bokulic, Asa Carlsson Tedgren, Emelie Adolfsson, Anna Maria Campos, Victor Gabriel Leandro Alves, Luo Suming, Wu Hao, Daniela Ekendahl, Irena Koniarova, Wojciech Bulski, Krzysztof Chelminski, José Luis Alonso Samper, Sumanth Panyam Vinatha, Sougata Rakshit, Srimanoroth Siri, Milan Tomsejm, Mikko Tenhunen, Julie Povall, Stephen F. Kry, David S. Followill, David I. Thwaites & Joanna Izewska

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Testing the methodology for a dosimetric end-to-end audit of IMRT/VMAT: results of IAEA multicentre and national studies

Paulina Wesolowska[a], Dietmar Georg[b,c], Wolfgang Lechner[b,c], Pavel Kazantsev[a] , Tomislav Bokulic[a], Asa Carlsson Tedgren[d,e], Emelie Adolfsson[d], Anna Maria Campos[f], Victor Gabriel Leandro Alves[f], Luo Suming[g], Wu Hao[h], Daniela Ekendahl[i], Irena Koniarova[i], Wojciech Bulski[j], Krzysztof Chelminski[j], José Luis Alonso Samper[k], Sumanth Panyam Vinatha[l], Sougata Rakshit[l], Srimanoroth Siri[m], Milan Tomsejm[n], Mikko Tenhunen[o], Julie Povall[p], Stephen F. Kry[q], David S. Followill[q], David I. Thwaites[p,r] and Joanna Izewska[a]

[a]International Atomic Energy Agency, Vienna, Austria; [b]Department of Radiation Oncology, Medical University of Vienna, Vienna, Austria; [c]Christian Doppler Laboratory for Medical Radiation Research for Radiation Oncology, Vienna, Austria; [d]Medical Radiation Physics, Department of Medical Physics and Department of Medical and Health Sciences, Linkoping University, Linköping, Sweden; [e]Department of Medical Radiation Physics and Nuclear Medicine, Karolinska University Hospital, Stockholm, Sweden; [f]Instituto Nacional De Câncer, Rio De Janeiro, Brazil; [g]Chinese Centre for Disease Control and Prevention, Beijing, China; [h]Beijing Cancer Hospital, Beijing, China; [i]National Radiation Protection Institute, Prague, Czech Republic; [j]Maria Sklodowska-Curie Memorial Cancer Centre and Institute of Oncology, Warsaw, Poland; [k]National Institute of Oncology and Radiobiology, Havana, Cuba; [l]Radiation Standards Section, Radiation Safety Systems Division, Bhabha Atomic Research Centre Trombay, Mumbai, India; [m]SSDL, Bureau of Radiation and Medical Devices, Department of Medical Science, Nonthaburi, Thailand; [n]CHU Charleroi, Hopital Andre Vesale, Montigny-le-Tilleul, Belgium; [o]Cancer Centre, Helsinki University Hospital, Helsinki, Finland; [p]University of Leeds, St James's University Hospital, Leeds, United Kingdom; [q]Imaging and Radiation Oncology Core Houston QA Centre, Anderson Cancer Centre, Houston, TX, USA; [r]Institute of Medical Physics, School of Physics, University of Sydney, Sydney, Australia

ABSTRACT

Introduction: Within an International Atomic Energy Agency (IAEA) co-ordinated research project (CRP), a remote end-to-end dosimetric quality audit for intensity modulated radiation therapy (IMRT)/volumetric arc therapy (VMAT) was developed to verify the radiotherapy chain including imaging, treatment planning and dose delivery. The methodology as well as the results obtained in a multi-centre pilot study and national trial runs conducted in close cooperation with dosimetry audit networks (DANs) of IAEA Member States are presented.

Material and methods: A solid polystyrene phantom containing a dosimetry insert with an irregular solid water planning target volume (PTV) and organ at risk (OAR) was designed for this audit. The insert can be preloaded with radiochromic film and four thermoluminescent dosimeters (TLDs). For the audit, radiotherapy centres were asked to scan the phantom, contour the structures, create an IMRT/VMAT treatment plan and irradiate the phantom. The dose prescription was to deliver 4 Gy to the PTV in two fractions and to limit the OAR dose to a maximum of 2.8 Gy. The TLD measured doses and film measured dose distributions were compared with the TPS calculations.

Results: Sixteen hospitals from 13 countries and 64 hospitals from 6 countries participated in the multicenter pilot study and in the national runs, respectively. The TLD results for the PTV were all within ±5% acceptance limit for the multicentre pilot study, whereas for national runs, 17 participants failed to meet this criterion. All measured doses in the OAR were below the treatment planning constraint. The film analysis identified seven plans in national runs below the 90% passing rate gamma criteria.

Conclusion: The results proved that the methodology of the IMRT/VMAT dosimetric end-to-end audit was feasible for its intended purpose, i.e., the phantom design and materials were suitable; the phantom was easy to use and it was robust enough for shipment. Most importantly the audit methodology was capable of identifying suboptimal IMRT/VMAT delivery.

Introduction

Since 1995, the International Atomic Energy Agency (IAEA) has conducted a series of co-ordinated research projects (CRPs) with the purpose of supporting its Member States to establish national audit activities through the development of audit methodologies and guidelines on how to structure and operate national dosimetry audit networks (DANs) [1–4]. In addition the IAEA, DAN database [5] was established to

monitor audit activities; moreover, it provides information on the coverage and operations of DANs in radiotherapy. The collected data suggest that the demand for audits significantly exceeds current capabilities [6].

It is generally agreed that complex radiotherapy techniques like intensity modulated radiation therapy (IMRT) including volumetric arc therapy (VMAT) need rigorous dose delivery verification procedures. Dedicated quality assurance is required during the radiotherapy units and treatment planning systems (TPS) commissioning, when IMRT/VMAT is introduced in clinics and finally as part of patient specific validation of the dose delivery [7–9]. To gain confidence that these complex treatments are accurately applied to patients, an independent verification, or a dosimetry audit, of the treatment chain should be performed [7,10]. The increased clinical use of IMRT and VMAT in IAEA Member States further supports the need for dosimetric audits of these treatment techniques for accuracy of dose delivery and safety purposes.

There are examples of IMRT verification projects [9,11–15] and several examples of pre-treatment QA performed at hospitals described in the literature [11,16–18]. These audit procedures include dosimetry ranging from point dose measurements, to 2D dosimetry with films and/or dosimeter arrays to 3D dosimetry with dosimeter arrays or gel. Even though IMRT verification is often implemented locally in a hospital there are studies demonstrating that local QA results do not necessarily concur with independent audit results [19,20]. The common approach for evaluating dose distributions is the use of a gamma analysis, which in some cases has shown inconsistencies between the results performed with the local QA equipment and from independent audits, which normally yields lower passing rates [19–23]. The inconsistencies noted above between local and independent audit QA results support the need for an independent dosimetry audit (beyond just relying on local QA measurements) for consistent, accurate and safe radiation therapy treatments utilizing complex modalities.

The purpose of this work is to present both the methodology of a new remote end-to-end audit for IMRT/VMAT head treatment delivery and the results obtained in a comprehensive testing phase. This audit aims to provide an independent verification of the radiotherapy chain including imaging, dose calculation, set-up and dose delivery, representing a continuation of previously developed audit methodologies [3,4,24,25]. This remote end-to-end audit methodology was developed under an IAEA CRP and included a multicentre pilot study and national trial runs conducted by various DANs in close cooperation with the IAEA.

A methodology for on-site IMRT head and neck audits, which is currently under implementation at the national level, was also recently developed by IAEA [26,27]. The advantage of an on-site audits is that any inconsistencies found can be immediately tracked to the cause. In a remote end-to-end test, error analysis is more challenging and if major deviations persist, an on-site visit is recommended. On the other hand, if an audit is to be performed on a large scale, remote audit is the modality of choice.

Material and methods

IMRT/VMAT phantom

A solid phantom specifically designed for this audit (Figure 1(a)) has dimensions of $15 \times 15 \times 15\,cm^3$ and contains an insert (denoted as IMRT QA insert) of $6.35 \times 6.35 \times 6\,cm^3$ outer dimensions. It contains an imageable planning target volume (PTV) and organ at risk (OAR) embedded into it. The distance between the PTV and OAR is 8 mm (Figure 1(b)). The phantom is made of polystyrene (density of $1.040\,g/cm^3$) while the PTV and OAR structures are made of Solid Water HE (GAMMEX RMI, Middleton, WI, USA) (density of $1.032\,g/cm^3$). For accurate positioning during CT scanning/imaging and during set-up on the treatment couch, the outer surface of the
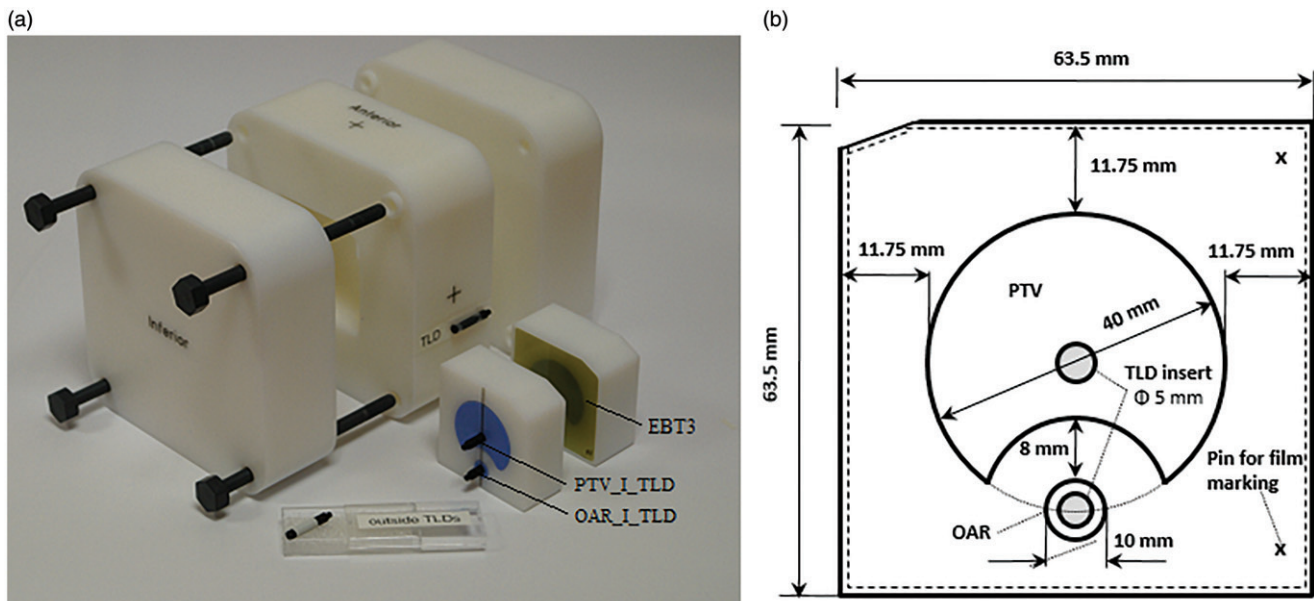

Figure 1 The IMRT polystyrene solid slab phantom with the IMRT QA insert disassembled and showing the location of PTV_I_TLD, OAR_I_TLD and EBT 3 film (a) and a cross section of the IMRT QA insert showing the TLD locations in the PTV and OAR (b).

phantom is engraved with crosses and markings for appropriate alignment.

The IMRT QA insert can be loaded with EBT3 Gafchromic film in the mid-trans axial plane along with four thermoluminescent dosimeters (TLDs), which are labeled as PTV_S_TLD and PTV_I_TLD for the superior and inferior positions in the PTV, respectively, and OAR_S_TLD and OAR_I_TLD for the superior and inferior positions in the OAR, respectively. Extra plugs and inserts enable additional measurements with a small volume ionization chamber (IC) (PTW Semiflex 31010, PTW PinPoint 31006 and Exradin A1SL) in the position of each TLD.

Imaging dose assessment

The imaging dose during CT scanning and its potential impact on the final audit dosimetric results was given a special focus in this study, since the contribution of CT dose to the OAR doses was expected to not be negligible. Therefore, the CT doses at the location of the OAR TLDs were determined for CT scans performed with different scanners and scanning parameters/protocols as used in several clinics. More specifically, for each CT scan evaluation four TLDs were positioned in the IMRT QA insert and two external TLDs were attached on the phantom surface (mid-lateral phantom sides); the locations of six TLDs are shown in Figure 1(a). The contribution of imaging doses to the PTV and OAR IMRT doses was derived from two outside TLDs attached to the phantom surface during CT scanning (Figure 1) and removed for IMRT irradiations. For this purpose, the phantom was filled with TLDs located in the PTV, OAR and outside (Figure 1) and CT scanned. The ratio of the signals from the inside TLDs to the signals from the outside TLDs was determined for a range of CT scanners and scanning parameters/protocols. The signal measured from the outside TLDs corrected for the inside/outside CT dose ratio was subtracted from the overall signal of TLDs at PTV and OAR.

A relation between CT dose index (CTDI) values and measured external TLD signal was investigated, to develop a method for estimating the imaging dose in order to subtract it from the dose received by inside TLDs during IMRT dose delivery.

End-to-end test methodology

The audit exercise begins with the local centre being instructed to perform a CT scan of the phantom using their clinical protocol for IMRT head (or head and neck) treatments. Next, the CT images are transferred to the TPS and an IMRT treatment plan is generated using the pre-defined dose prescriptions and dose constraints provided in the audit instructions. For the treatment plan preparation, the aim is to deliver 4 Gy in two fractions to the PTV, while limiting the dose to the OAR to below 2.8 Gy. The dose constraints per fraction are as follows: $D_{95\%}=2.0$ Gy and $V_{93\%}<1\%$, $D_{max}<2.2$ Gy (anywhere in the plan) and $D_{max}<1.4$ Gy in OAR. The treatment plan is to be verified using the local patient-related QA procedures and then the phantom irradiations

are to be performed with two fractions delivered one by one without the phantom repositioning.

A set of instruction sheets for this end-to-end audit, data sheets for reporting the details of irradiation, additional instructions for DANs on how to handle and evaluate the dosimeters, and the final format of the reporting of results were also developed.

Dosimetric evaluation

For the multicentre pilot study, the dosimeters were prepared and analyzed by the IAEA Dosimetry Laboratory (DOL). As mentioned above, two dosimeter types were used for the measurements within the phantom, i.e., TLD-100 (Harshaw, WI, USA) and Gafchromic EBT3 film (Ashland, OR, USA), one for determining the point doses and one for verifying the dose distribution, respectively.

For TLD, approximately 165 mg of LiF powder was encapsulated in plastic capsules (2.8 cm length, 0.5 cm diameter). TLD powder from the capsules was read/analyzed with a PCL3 TLD reader (Fimel, France) following the established IAEA protocol [28]. The uncertainty in the TLD measurements was 1.6% (1 standard deviation (SD)).

The scanning of the irradiated EBT3 films was performed with an EPSON 11000XL flat-bed scanner (EPSON, Nagano, Japan) using the transmission mode, 150 dpi resolution and 48-bit RGB color scale. The calibration films were irradiated with seven doses in the range from 0.5 to 6 Gy. The calibration films were irradiated within 2 weeks of a phantom's audit irradiation. The calibration curve and the dose distributions from the films were obtained using FilmQA Pro (Ashland, OR, USA) software with the triple channel method [29]. Gamma evaluation was performed for dose distributions matched using pin registration marks on the film and normalized to the same dose at the isocentre. The region of interest (ROI) was selected to include the whole PTV and OAR and exclude the pin marks on the film. The gamma analysis evaluation parameters were: 3% dose difference, 3 mm distance-to-agreement, 20% dose threshold, global gamma.

Multicentre pilot study

Participants from Austria, Belgium, Finland, Sweden, UK and USA (who were all closely involved in the methodology development) together with seven DANs from Brazil, China, Cuba, Czech Republic, India, Poland and Thailand took part in the multicentre pilot study. DANs work in close cooperation with local hospitals to perform all tests in a clinical environment. Each of the pilot study participants was provided with an audit set consisting of a phantom preloaded with dosimeters, instructions and datasheets. They were instructed to 'treat' this IMRT phantom as if it were one of their patients so that the entire IMRT treatment preparation and delivery process from end-to-end was carried out in accordance with their clinical process.

After the phantom irradiation was completed, the dosimeters (TLDs and film) were returned to the IAEA dosimetry laboratory for read-out, evaluation and analysis. Participants

were requested to report, among other things, the TPS doses calculated at the positions of each of the four TLDs as well as to provide the DICOM file with the TPS dose distribution.

In total 16 hospitals from 13 countries participated in the multicentre study; details of the equipment used are available in Supplementary Table S1.

In addition to the audit procedure of measuring the delivered doses with TLDs, IC measurements were also performed during the multicentre pilot study to verify whether any additional corrections were needed when using TLDs in the IMRT phantom. Furthermore, small isocentre shifts of 1 mm were simulated in the TPS to assess how potential phantom setup errors might affect the audit results. Five project participants performed these additional calculations for the phantom shifted laterally, and in vertical directions.

National trial runs

After the completion of the multicentre pilot study, the audit methodology was tested at the national level by DANs. The phantoms which each DAN has received during the pilot study were used for a circulation within the national runs. The same methodology description together with instructions were provided to DANs to initiate their national trial runs. They were performed in six countries (Brazil, China, Cuba, Czech Republic, India, Poland) where in total 64 hospitals received the phantom for irradiation (details of equipment used is available in Supplementary Table S2). Each DAN organizing a trial in the country was responsible for providing the dosimeters and performing the audit evaluation and analysis.

Methodology robustness

The one-sample Wilcoxon signed ranks test was used to determine whether the median of the distribution of the ratios of measured to hospital stated doses (Dmeas/Dstat) calculated for the four TLD locations was different from 1.

This would indicate a bias and could reveal, besides the problems with systematic errors related with beams calibration or equipment used, the need for an additional correction related to the phantom design. The choice of the test was based on the analysis of normality (Shapiro–Wilk test) and the analysis of Q-Q plots of the groups of D_{meas}/D_{stat} ratio results for the four TLD locations.

In addition, the non-parametric Kruskal–Wallis test, based on median scores and ranks, was conducted to investigate whether there was a significant difference between the D_{meas}/D_{stat} ratio data obtained in the multicentre pilot study versus the national trial runs.

The impact of treatment technique and underlying equipment was also reviewed.

Results

Imaging dose assessment

Evaluation of CT doses obtained from 11 scans of the phantom, suggested that TLD signals resulting from CT imaging were high enough to contribute to the overall TLD measured doses. Imaging dose contributions amounted to up to 2.6% of PTV dose and up to 20.6% of OAR dose.

The ratio of the signal from the inside to the outside TLDs ranged between 1.07 and 1.21, with an average value of 1.13 across the different scanning protocols.

The relation between the measured TLD signal and CTDI values can be seen in Figure 2. The results show some outlying CTDI values, therefore it was decided that the external TLD would be used to derive the TLD signal resulting from the planning CT scan rather than CTDI values.

Multicentre pilot study

The submitted details of the treatment plans showed that participants generally achieved the requested maximum dose constraint, except for two cases where the maximum dose for anywhere in the plan was above the 2.2 Gy limit
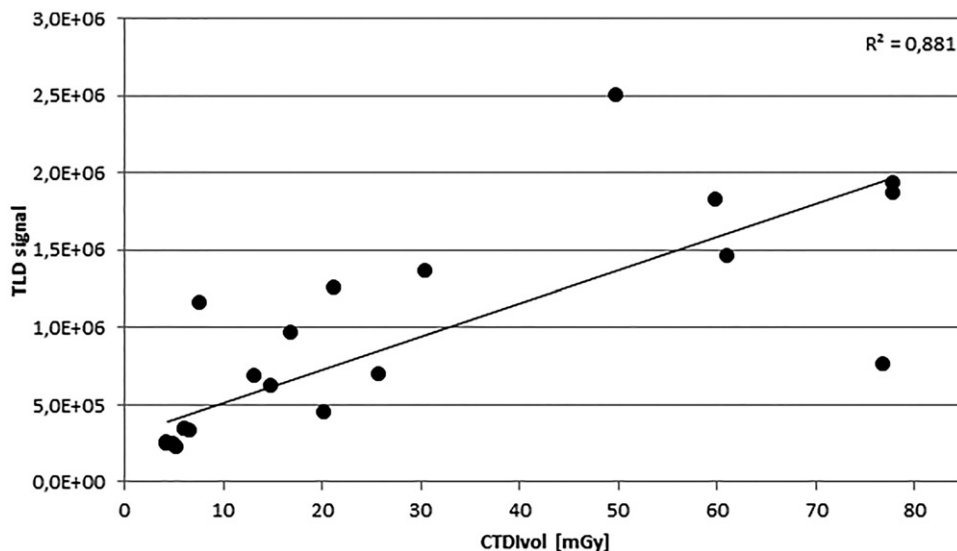

Figure 2. Relation between the TLD signal and the CTDI doses.

(both reported 2.27 Gy). The planned OAR doses were all below 1.4 Gy per fraction. For different participants the mean OAR doses ranged from 0.30 to 1.06 Gy per fraction.

Each treatment plan was verified using local procedures based on detector arrays, such as MatriXX (IBA Dosimetry, Schwarzenbruck, Germany), ArcCheck (Sun Nuclear Corporation, Melbourne, Australia), Octavius (PTW, Freiburg, Germany), Delta 4 (ScandiDos, Uppsala, Sweden) or EPID (Varian, Palo Alto, CA, USA). All but one participant successfully passed local QA verification of their treatment plans.

The TLD results for the 16 participants who contributed to the multicentre pilot study are shown in Figure 3(a). Three participants repeated the irradiation. Participant #2 repeated the irradiations due to a large discrepancy of doses in the OAR caused by a steep dose gradient. A new TPS plan was generated with the smaller gradient in the OAR volume, which resulted in good agreement for the repeated irradiation. Participant #5 made a mistake during the phantom irradiation and even though after TPS recalculations the results were acceptable the irradiation was repeated. Participant #10 had the average dose in the PTV 6.6% higher than planned; the follow-up irradiation with the same plan confirmed that the irradiation was performed correctly, and the discrepancy originated in the TPS calculations. This participant was about to introduce IMRT into clinical practice and thus was not sufficiently experienced. Other participants (#11 and #15) with outlier results in OAR only, were not requested to repeat irradiation and the reasons are discussed below.

The D_{meas}/D_{stat} ratios were calculated for all four TLD locations; the average and the SD values are 0.998 ± 0.028, 0.993 ± 0.023, 1.010 ± 0.063 1.003 ± 0.047 for PTV_S, PTV_I, OAR_S and OAR, respectively. The mean ratio of the TLD measured doses in the PTV to the stated ones was within $\pm 5\%$ for all participants except Participant #10. The $\pm 5\%$ acceptance limit was defined based on the SD of all D_{meas}/D_{stat} ratios in PTV. This limit corresponds to expanded SD of 5.1% ($k = 2$), giving approximately $\pm 5\%$ acceptance limit.

More variability was observed for the OAR results, with the minimum ratio of D_{meas}/D_{stat} as low as 0.82 and the maximum of 1.14 as shown in Figure 3(a). This was expected due to high dose gradients across OAR which were also more sensitive to phantom positioning uncertainties. All results were below the 2.8 Gy planning constraint (Figure 3(b)). The TPS simulation of the phantom isocentre shifts shown that the doses calculated for the TLD powder volume can change by 1.0% for PTV TLDs, while the 1 mm shift could make a difference to up to 11.7% for OAR TLDs.

The comparison of doses measured in the PTV with the IC and the TLD showed excellent agreement (average difference between dosimeters 0.5%, SD = 1.2%). The OAR IC dose results showed larger scatter with the SD of 9.9%, greater than that for TLD.

Gamma analysis of film evaluation resulted in very good agreement of the analyzed dose distributions. An example of the film measured and TPS calculated dose profiles can be seen in Figure 4. The average result was 99.5% of points passing the gamma criteria with the minimum value of 96.7%.
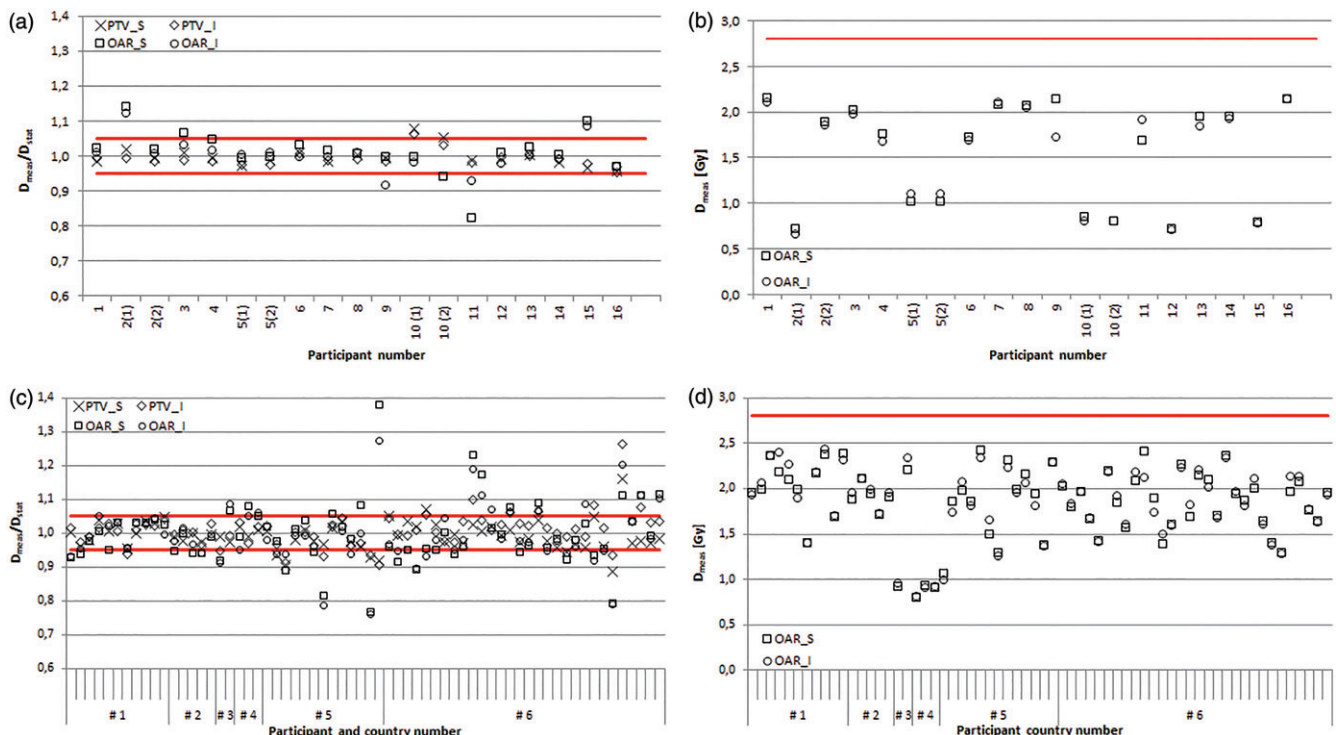

Figure 3. TLD results for multicentre study and national trials; TLD locations marked as: PTV_S_TLD and PTV_I_TLD for the superior and inferior positions in the PTV and OAR_S_TLD and OAR_I_TLD for the superior and inferior positions in the OAR. (a) ratio of the TLD measured to stated doses for each institution from the multicentre study (the ±5% acceptance level applicable for PTV only is shown in red), data for institutions 2, 5 and 10 include the initial (1) and repeated (2) irradiation results. (b) TLD dose measurement results in OAR for each institution from the multicentre study (the 2.8 Gy acceptance limit is shown in red), data for institutions 2, 5 and 10 include the initial (1) and repeated (2) irradiation results. (c) ratio of the TLD measured to stated doses for each institution from the national trials (the ±5% acceptance level applicable for PTV only is shown in red); (d) TLD dose measurement results in OAR for each institution from the national trial (the 2.8 Gy acceptance limit is shown in red).
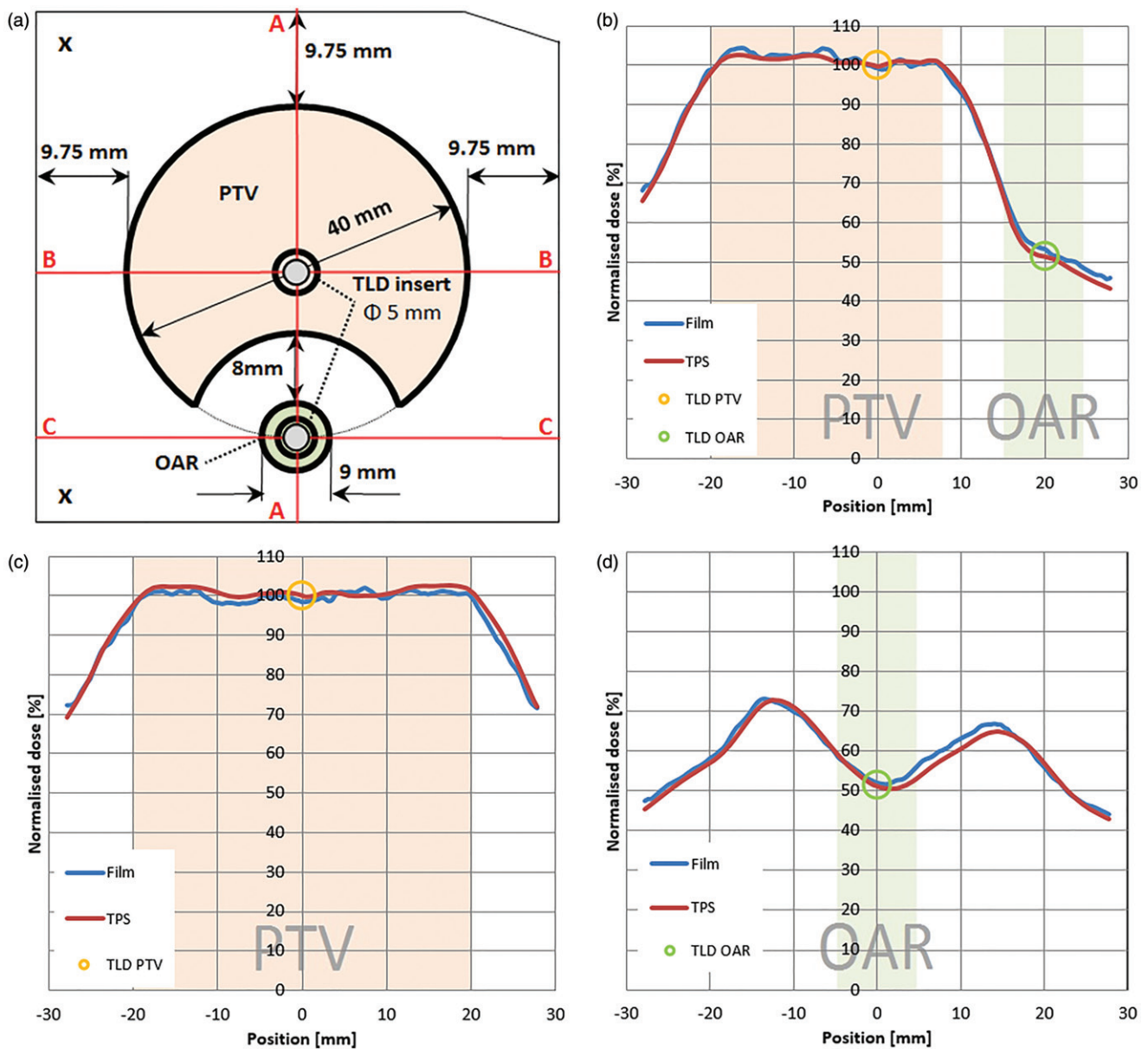
Figure 4. Comparison of measured and TPS calculated profiles; the IMRT QA insert with regions and cross-sections marked (a), vertical profile along the axis passing through the isocentre (A-A) (b), horizontal profile along the axis passing through the isocentre (B-B) (c) and a horizontal profile along the axis passing through the OAR midpoint (C-C) (d).

National trial runs

The details of the treatment plans developed for national audits identified three plans where the maximum dose anywhere in the plan was above the planning constraint limit and five plans for which OAR exceeded the maximum dose limit (in total seven different plans). All participants confirmed that their plans passed the local QA verification procedure. The measurements obtained from the national trial runs performed in six countries with 64 participating hospitals are summarized in Figure 3(c) and 3(d). The SD of the ratio of measured to stated dose for all results can be found in Table 1.

In 47 of 64 cases the TLD results in the PTV were within the 5% acceptance limits. Most of the unsatisfactory results (15/17) were observed in countries #5 and #6; for which the spread of

results was substantially higher than in other countries (see Table 1). The doses measured in the OAR were all below the 2.8 Gy threshold, with a minimum measured dose of 0.88 Gy and a maximum of 2.43 Gy for all cases.

The gamma analysis results identified seven centres who failed to achieve the 90% pass rate. One institution achieved only 10% of pixels passing this gamma criterion. Three of the EBT3 film results showed gamma passing rate below 50% but TLD results were within 5% limit. This suggests issues with film handling at the national level rather than dosimetric problems.

Some of centres with low gamma passing rates also had poor TLD results, with two centres failing both TLD and film criteria.

Table 1. Average results of TLD measurements for all national trial runs.

	TLD location					
	PTV_S	PTV_I	PTV	OAR_S	OAR_I	OAR
Country 1 $N = 11$						
D_{meas}/D_{stat}	1.007	1.006	**1.006**	0.992	1.003	**0.997**
SD	0.031	0.029	**0.029**	0.043	0.040	**0.041**
Country 2 $N = 5$						
D_{meas}/D_{stat}	0.988	1.000	**0.994**	0.963	0.979	**0.971**
SD	0.012	0.023	**0.018**	0.028	0.028	**0.028**
Country 3 $N = 2$						
D_{meas}/D_{stat}	0.980	0.971	**0.976**	0.993	1.000	**0.996**
SD	0.011	0.033	**0.021**	0.103	0.092	**0.092**
Country 4 $N = 3$						
D_{meas}/D_{stat}	1.007	1.007	**1.007**	1.040	1.020	**1.007**
SD	0.015	0.032	**0.023**	0.046	0.061	**0.023**
Country 5 $N = 13$						
D_{meas}/D_{stat}	0.970	0.974	**0.972**	0.998	0.968	**0.983**
SD	0.040	0.044	**0.041**	0.147	0.122	**0.133**
Country 6 $N = 30$						
D_{meas}/D_{stat}	0.998	1.030	**1.014**	0.998	1.003	**1.000**
SD	0.048	0.056	**0.054**	0.090	0.088	**0.088**
All $N = 64$						
D_{meas}/D_{stat}	**0.993**	**1.009**	**1.001**	**0.996**	**0.995**	**0.995**
SD	**0.042**	**0.050**	**0.047**	**0.093**	**0.086**	**0.089**

In total, 21 institutions (34%) failed the audit for either one or both criteria. The follow-up on the poor results is still ongoing at the national level and these results have not been reported so far.

Methodology robustness

The Shapiro–Wilk test showed that D_{meas}/D_{stat} ratios for different TLD locations in the phantom did not have the normal distribution and therefore a non-parametric statistical test was used. The one-sample Wilcoxon signed rank demonstrated that the median D_{meas}/D_{stat} ratios obtained in both the multicentre pilot study and the national trial runs were not different from 1 ($p > .01$) with the p values ranging from 0.02 to 0.9, i.e., no systematic difference was observed between the stated dose and the measured one.

Furthermore, in both the multicentre study and the national trial runs no significant difference was found between the results related to different locations within the phantom (with the p values of .05 for multicentre study and .18 for national runs) as well as between the results of both multicentre and national studies with p value of .63 (Mann–Whitney test, $p > .01$).

Additional tests did not detect any significant differences in D_{meas}/D_{stat} ratios between the linac manufacturers, delivery techniques or plans calculated on different calculation grids (0.4–5 mm). Even though the number of MU per 2 Gy fraction (302–1202 MU) differed substantially between plans, statistical tests did not detect any significant differences in D_{meas}/D_{stat} ratios.

Discussion

The remote end-to-end audit methodology described in this work aims to independently verify the entire IMRT/VMAT treatment chain, covering steps from imaging to dose delivery. The imaging dose assessment conducted during the initial phase of the CRP showed the importance of incorporating the CT imaging TLD measurements, i.e., imaging doses were found to potentially influence the final TLD audit result if not corrected. This is in agreement with the findings of Molineu et al. [13] for a similar audit methodology, reporting a CT dose contribution of approximately 1% to the PTV dose of 6.6 Gy. Also, their attempt to use the CTDI values for estimating the CT dose contribution to the PTV and OAR TLD doses was not successful, similar to our results. The CT dose vs. CTDI data scatter might be due to several reasons. First, the CTDI and TLD dose relationship is complex because the CT X-ray energy spectrum changes throughout the phantom and impacts on the signal recorded by the TLD. Next, there is a large impact on the TLD signal by the individual CT scanner parameters. For example, a similar TLD signal of approximately 0.7 million counts was associated with either 13 mGy, 26 mGy or 77 mGy in CTDIs (see Figure 2) depending on the CT scanner and protocol used. This largely depends on the type of phantom used for the CTDI determination as both head and body phantoms are in clinical use, resulting in very different CTDIs and variations in CT manufacturers and models dose reporting method (generally up to ±20% disagreement is allowed between the displayed CTDI value and measurements). The most striking finding of this imaging dose study was the unexpectedly wide range of CTDI values reported by participants, starting from abaut 5 mGy with an upper limit of approximately 78 mGy. This clearly indicates the optimization and standardization potential for planning CT in the head and head-and-neck regions to minimize patient exposure in healthy tissue.

After the design and development of an IAEA audit, the multicentre pilot study is a major milestone in the testing of the new prototype phantom, verifying whether the planning objectives and constraints can realistically be achieved within the phantom geometry, and checking the feasibility of irradiation conditions. Such a pilot study validates the remote audit process, provides feedback on the developed documents (e.g., instructions and data sheets) and gives an indication of the workload involved for both the centres to be audited and the DAN to provide the evaluation. Since the pilot study results are used as the benchmark for national audits, it is essential to include a representative number of centres with respect to equipment and overall experience. An extra effort is needed when participating in pilot studies for audit methodologies; for example, extra robustness tests such as the influence of positioning shifts on measurement points are needed and assessing dosimetry analysis parameters for radiochromic film. By doing so, the influence of positioning shifts was found to be very high in the OAR (up to 11.7% dose difference for a 1 mm shift), which lead to developing the acceptance criteria for OAR that are not based on D_{meas}/D_{stat} ratios as was done for the PTV. Instead it was decided to verify whether the planning constraints were fulfilled. The high sensitivity of the measurement points in the OAR to positioning also justifies the choice of orientation of the EBT3 film in the transverse plane, allowing a close examination of the dose gradient in the OAR region.

Based on the results of the multicentre study, the acceptance limits of 5% for the dose measurements in PTV and 2.8 Gy dose threshold for the dose measurements in OAR were considered feasible. For film gamma evaluation, the criterion of 90% of points passing selected criteria was adopted. These acceptance levels were recommended for testing during the national trial runs.

Centres participating in the multicentre pilot, who were experienced with IMRT/VMAT treatment planning and delivery, had very good audit results, while other centres who were just beginning their clinical implementation of IMRT did not do so well in this audit and benefited from participation in this study. As mentioned above, three participants repeated the exercise and in all these cases the local IMRT verification methods were passed with good results. The finding that local QA methods do not always correlate with the audit results has also been reported by others [19,20,30].

The results of the national trial runs indicated a much greater percentage of centres not meeting the audit acceptance criteria than observed in the multicentre study results, i.e., the SD of the D_{meas}/D_{stat} ratio was much larger. On the other hand, the audit implementation at the national level was done correctly by the respective DANs as the median results for the multicentre and national trial are similar. However, there were challenges with film handling in one country.

The follow-up of the unsatisfactory results includes at the first place verification whether all instructions were followed. Auditors request the participants to send additional information which is reviewed and the audit is repeated. The information on the linac output check (ionization chamber and electrometer details, calibration certificate, factors and corrections used), MLC performance (positioning accuracy test results) and the details of the TPS parameters related to the small fields and MLC (small fields output factors and profiles, detectors and corrections used, leaf gap and transmission) are requested. All of the parameters mentioned above could be a reason for deviations detected during the audit and should be carefully verified by a clinical medical physicist from the audit team.

Even though the IMRT/VMAT end-to-end audit presented in this work is based on a rather simple phantom which allows a dose distribution evaluation in a single plane only, it is able to detect shortcomings in the implementation of IMRT/VMAT. A phantom with a small PTV and OAR is especially challenging. The analyzed dose distributions using EBT3 film were relative and normalized to PTV. Therefore, any potential linac output problems cannot be detected by film measurements but by TLD only. In contrast, any problems with MLC characteristics, phantom positioning shifts or increased uncertainty in isocentre positioning and any other geometry inaccuracies, can be detected using the film [31].

The analysis of results obtained for this end-to-end audit did not detect any significant influence of linac type and manufacturer, delivery technique, number of monitor units and TPS calculation grid size. Nakamura et al. also did not detect any dependence on the delivery technique [32]. On the other hand, in other studies differences were found between TPS models, specifically when beam models were provided by the vendor [11,12,33]. Also, the treatment techniques or infrastructure of the department (number of medical physicists or age of equipment) have been shown to affect the results [12].

The overall aim when designing this IMRT/VMAT audit was to develop an end-to-end test that could be adopted at the national level by a DAN having a certain expertise gained during the implementation of the previous CRP's. The multicentre pilot study and the national runs proved that the IMRT/VMAT dosimetric end-to-end audit methodology can be implemented without major challenges. The phantom design and materials used proved to be suitable for the purpose; the phantom was easy to use and there were no issues during the shipment. More importantly, this audit is capable of identifying suboptimal IMRT/VMAT delivery in a radiotherapy department.

Development of such a remote audit is in line with recommendations [7] that regular independent peer review dosimetry audits should be made available to radiotherapy centres and that they are especially important when new treatment units are installed or when introducing new treatment techniques and/or TPS software updates.

ORCID

Pavel Kazantsev 🆔 http://orcid.org/0000-0001-7662-0992

References

[1] International Atomic Energy Agency (IAEA). Setting up a dosimetry audit centre: infrastructure and resources. SSDL Newsletter. 2017;66:9–18.

[2] Izewska J, Bera P, Azangwe G, et al., editors. IAEA support to national TLD audit networks for radiotherapy dosimetry. Standards, applications and quality assurance in medical radiation dosimetry (IDOS). Proceedings of an International Symposium; 2010 Nov 9–12; Vienna, Austria: IAEA; 2010.

[3] Izewska J, Georg D, Bera P, 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.

[4] Izewska J, Wesolowska P, Azangwe G, et al. Testing the methodology for dosimetry audit of heterogeneity corrections and small MLC-shaped fields: results of IAEA multi-center studies. Acta Oncol. 2016;3:1–8.

[5] Dosimetry Audit Networks (DAN) Database [Internet]. 2017. [cited 2018 May 9]. Available from: https://dosimetry-audit-networks.iaea.org/Home/AuditAvailability.

[6] Izewska J, Lechner W, Wesolowska P. Global availability of dosimetry audits in radiotherapy: the IAEA dosimetry audit networks database. Phys Imag Radiat Oncol. 2018;5:1–4.

[7] International Atomic Energy Agency (IAEA). Accuracy requirements and uncertainties in radiotherapy IAEA Human Health Series 31. Vienna (Austria): IAEA; 2016.

[8] Kron T, Lehmann J, Greer PB. Dosimetry of ionising radiation in modern radiation oncology. Phys Med Biol. 2016;61:167–205.

[9] Miften M, Olch A, Mihailidis D, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: recommendations of AAPM Task Group No. 218. Med Phys. 2018;45:53–83.

[10] Dunscombe P. Recommendations for safer radiotherapy: what's the message? Front Oncol. 2012;2:129.

[11] Clark CH, Hussein M, Tsang Y, et al. A multi-institutional dosimetry audit of rotational intensity-modulated radiotherapy. Radiother Oncol. 2014;113:272–278.

[12] Molineu A, Hernandez N, Nguyen T, et al. Credentialing results from IMRT irradiations of an anthropomorphic head and neck phantom. Med Phys. 2013;40:022101.

[13] Molineu A, Followill DS, Balter PA, et al. Design and implementation of an anthropomorphic quality assurance phantom for intensity-modulated radiation therapy for the Radiation Therapy Oncology Group. Int J Radiat Oncol Biol Phys. 2005;63:577–583.

[14] Budgell G, Berresford J, Trainer M, et al. A national dosimetric audit of IMRT. Radiother Oncol. 2011;99:246–252.

[15] Adolfsson E, Gustafsson H, Lund E, et al. A system for remote dosimetry audit of 3D-CRT, IMRT and VMAT based on lithium formate dosimetry. Radiother Oncol. 2014;113:279–282.

[16] Lafond C, Chiavassa S, Bertaut C, et al. DEMAT: a multi-institutional dosimetry audit of rotational and static intensity-modulated radiotherapy. Phys Med. 2016;32:664–670.

[17] Miri N, Lehmann J, Legge K, et al. Remote dosimetric auditing for intensity modulated radiotherapy: a pilot study. Phys Imag Radiat Oncol. 2017;4:26–31.

[18] Clark CH, Hansen VN, Chantler H, et al. Dosimetry audit for a multi-centre IMRT head and neck trial. Radiother Oncol. 2009;93:102–108.

[19] Kry SF, Molineu A, Kerns JR, et al. Institutional patient-specific IMRT QA does not predict unacceptable plan delivery. Int J Radiat Oncol Biol Phys. 2014;90:1195–1201.

[20] Weber DC, Vallet V, Molineu A, et al. IMRT credentialing for prospective trials using institutional virtual phantoms: results of a joint European Organization for the Research and Treatment of Cancer and Radiological Physics Center project. Radiat Oncol. 2014;9:123.

[21] Jornet N, Carrasco P, Beltrán M, et al. Multicentre validation of IMRT pre-treatment verification: comparison of in-house and external audit. Radiother Oncol. 2014;112:381–388.

[22] Jurado-Bruggeman D, Hernández V, Sáez J, et al. Multi-centre audit of VMAT planning and pre-treatment verification. Radiother Oncol. 2017;124:302–310.

[23] Hussein M, Clark CH, Nisbet A. Challenges in calculation of the gamma index in radiotherapy - towards good practice. Phys Med. 2017;36:1–11.

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

[25] Lechner W, Wesolowska P, Azangwe G, et al. A multinational audit of small field output factors calculated by treatment planning systems used in radiotherapy. Phys Imag Radiat Oncol. 2018;5:58–63.

[26] IAEA supported national "end-to-end" audit programme for dose delivery using intensity modulated radiation therapy through on-site visits to radiation therapy institutions [Internet]. Vienna, Austria: International Atomic Energy Agency (IAEA); c2018–2019 [cited 2019 Mar]. Available from: https://dosimetry-audit-networks.iaea.org/Content/end-to-end%20CIRS%20SHANE/National%20end-to-end%20IMRT-VMAT%20audit%20methodology.pdf.

[27] Kazantsev P, Clark C, Venencia D, et al. New IAEA end-to-end on-site IMRT audit methodology: pilot test results. Paper presented at International Conference on Advances in Radiation Oncology ICARO2; 2017 Jul 10–12; Vienna, Austria.

[28] Izewska J, Hultqvist M, Bera P. Analysis of uncertainties in the IAEA/WHO TLD postal dose audit system. Radiat Meas. 2008;43:959–963.

[29] Micke A, Lewis DF, Yu X. Multichannel film dosimetry with non-uniformity correction. Med Phys. 2011;38:2523–2534.

[30] Seravalli E, Houweling AC, Van Battum L, et al. Auditing local methods for quality assurance in radiotherapy using the same set of predefined treatment plans. Phys Imag Radiat Oncol. 2018;5:19–25.

[31] Pasler M, Hernandez V, Jornet N, et al. Novel methodologies for dosimetry audits: adapting to advanced radiotherapy techniques. Phys Imag Radiat Oncol. 2018;5:76–84.

[32] Nakamura M, Minemura T, Ishikura S, et al. An on-site audit system for dosimetry credentialing of intensity-modulated radiotherapy in Japanese Clinical Oncology Group (JCOG) clinical trials. Phys Med. 2016;32:987–991.

[33] Gershkevitsh E, Pesznyak C, Petrovic B, et al. Dosimetric inter-institutional comparison in European radiotherapy centres: results of IAEA supported treatment planning system audit. Acta Oncol. 2014;53:628–636.