Radiotherapy and Oncology 132 (2019) 63–69

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Radiotherapy quality assurance of SBRT for patients with centrally located lung tumours within the multicentre phase II EORTC Lungtech trial: Benchmark case results

Marie Lambrecht [a,c,*], Enrico Clementel [b], Jan-Jacob Sonke [c], Ursula Nestle [d,e], Sonja Adebahr [e,f,g], Mathias Guckenberger [h], Nicolaus Andratschke [h], Damien C. Weber [h,i], Marcel Verheij [c], Coen W. Hurkmans [a]

[a] Department of Radiation Oncology, Catharina Hospital, Eindhoven, The Netherlands; [b] Department of Quality Assurance, EORTC Headquarters, Brussels, Belgium; [c] Department of Radiation Oncology, The Netherland Cancer Institute, Amsterdam, The Netherlands; [d] Department of Radiation Oncology, Kliniken Maria Hilf GmbH Mönchengladbach; [e] Department of Radiation Oncology Medical Center, Faculty of Medicine, University of Freiburg; [f] German Cancer Consortium (DKTK) Partner Site Freiburg; [g] German Cancer Research Center (DKFZ), Heidelberg, Germany; [h] Department of Radiation Oncology, University Hospital Zurich; and [i] Center for Proton Therapy, Paul Scherrer Institute, ETH Domain, Villigen, Switzerland

ARTICLE INFO

Article history:
Received 30 November 2017
Received in revised form 14 September 2018
Accepted 23 October 2018
Available online 21 December 2018

Keywords:
Benchmark
SBRT
Lung
Planning exercise
Clinical trial

ABSTRACT

Purpose: To report on the benchmark case (BC) study performed in the context of the European Organisation for Research and Treatment of Cancer prospective multicentre Lungtech trial of SBRT for patients with inoperable centrally located lung tumours.

Methods and materials: Target volume and organs at risk (OARs) delineations first needed to be acceptable before the treatment plan was reviewed. Retrospectively, Dice similarity coefficients of the OARs and the target volumes were calculated and a set of gold standard contours adapted for each institution margins was applied on the accepted dose submissions to evaluate the influence of acceptable delineation variations on dosimetry.

Results: Twenty-five institutions participated. Five BCs were accepted at the first attempt. Twenty institutions had to revise their delineation at least once and seven had to revise their planning once. The V_{60} Gy dose coverage improved significantly ($p = 0.05$) between the first and final submissions from median (range) 94.8% (22.5–97.8) to 95.3% (70.5–99.3). The median Dice coefficient varied significantly between OARs: The lowest values were found for the brachial plexus 0.25 (0.01–0.54) and the highest for the spinal cord 0.89 (0.71–0.95). The mean PTV Dice coefficient was 0.82 (0.48–0.92). Applying the gold standard contours, only one institution remained compliant with the dose coverage criteria with V_{60} Gy median (range) of 83.4% (54.2–93.9).

Conclusions: Clinical guidelines and radiotherapy protocols are not a substitute for timely radiotherapy quality assurance procedures, which improve dose coverage significantly. Delineation remains the main source of BC rejection and plan review without first reviewing delineation may not be efficient. Our results show that delineation variations seem to have a larger influence on PTV coverage than variations in planning and irradiation techniques and thus suggest that dose tolerance criteria should preferably take into account the accuracy of delineation.

Radiotherapy Quality assurance (RTQA) has become a necessary and valuable intrinsic tool for conducting clinical trials and poor protocol compliance has been shown to have a critical impact on the trial outcome [1–3]. The EORTC has defined various RTQA levels which can be adopted in specific trials depending on e.g. trial complexity and trial end-points. Within the EORTC 22113-08113 prospective phase II Lungtech trial on stereotactic radiotherapy for centrally located non-small cell lung cancer (NSCLC), an RTQA programme was initiated that entailed not only all the standard RTQA levels but also included new quality assurance measurements and procedures not previously used within the EORTC [4,5]. This study reports here on one of the steps within the Lungtech RTQA process; the Benchmark Case (BC). A BC is often used to test the clarity of the protocol and to evaluate the inter-institution variation in delineation and treatment planning [6].

The aim of the BC within this trial was to first retrieve an overview of the treatment planning techniques and to quantify the dosimetric variations as well as the variability in target volume and organ at risk delineations. Furthermore, the objective was to quantify institutional PTV coverage when a gold standard (GS)

* Corresponding author at: Department of Radiation Oncology, Catharina Hospital, Michelangelolaan 2, 5623EJ, Eindhoven, The Netherlands.
E-mail address: marie.lambrecht@catharinaziekenhuis.nl (M. Lambrecht).

https://doi.org/10.1016/j.radonc.2018.10.025
0167-8140/

delineation adapted to institution specific margins is used to calculate the DVHs. This provides insight into the influence of delineation variations relative to variations in treatment planning on DVHs.

The data from the BC will give valuable insights into the techniques currently employed for treating this patient group and the level of congruence for delineation and treatment planning of such patients, which to our knowledge, has not been presented thus far for stereotactic body radiotherapy (SBRT) for centrally located lung tumours. It is therefore also of interest to institutions treating such patients outside of this trial.

Materials and methods

The Lungtech trial is a phase II EORTC trial on SBRT for centrally located NSCLC (T1-T3, N0, M0) launched in 2014. The trial aimed to evaluate the efficacy and toxicity using an SBRT regimen of 8 times 7.5 Gy. The primary end point was freedom from local progression at 3 years. Twenty-five institutions distributed over 6 European countries participated in the benchmark procedure.

Benchmark process and evaluation

All institutions willing to participate in the Lungtech trial were sent the BC data. These comprise a 3D-CT, a 4DCT and a 3D PET-CT of an example patient which would be eligible for this trial. The patient is a 54 years old female with histologically confirmed NSCLC stage IA who was considered medically inoperable. The patient did not have prior radiotherapy or chemotherapy within 3 months before the start of SBRT. The patient was known with severe chronic obstructive pulmonary disease and with a concurrent hepatitis infection. The patient had a WHO performance status of 2. Institutions were asked to plan this case according to the trial guidelines using the planning technique they would also use within the trial. The trial guidelines included requierements for window and level settings used during delineation. Institutions were required to submit their DICOM-RT structures, RT-dose and RT-plan files through the EORTC digital platform for central evaluation in the Vodca^{TM} software.

The anonymizsd BCs were reviewed and classified as "per protocol", with "acceptable variations" or with "unacceptable variations" according to the current Global Harmonisation Group classification in a two-step process: for each complete submission, the delineations were reviewed by a radiation oncologist and if judged per protocol or with acceptable variations, a medical physicist then reviewed the BCs treatment plans and dose distributions. Institutions with a BC that was judged unacceptable were requested to resubmit a new BC with the proper adjustments. All institutions were informed about the results of the BC within three working days after submission with suggestions for improvements when applicable. Institutions were not allowed to recruit patients until satisfactory completion of the exercise.

The BCs were reviewed according to the RTQA guidelines defined in the protocol. In the Annex the dosimetric and the delineation criteria for the OARs are defined (Annex Tables 1 and 2). The dose specification criteria for the target volumes are given in Table 1. Additional descriptions for the delineations of the OARs and target volumes can be found in the trial protocol (ClinicalTrials.gov Identifier: NCT01795521) and in the article introducing the Lungtech trial from Adebahr et al. [4]. Delineation guidelines were based on the delineation atlases previously published by Feng et al. and Kong et al. [7,8]. To aid the reviewers, a set of gold standard (GS) volumes was defined before the benchmark exercise took place. These volumes were delineated by one experienced radiation oncologist and reviewed and agreed upon by the radiation oncologists defining the protocol.

All submissions were retrospectively reviewed in order to evaluate the level of protocol understanding and to identify the main causes of protocol deviations. Furthermore, the evaluation aimed to determine the remaining variations in OARs and target delineations and dose distribution within the accepted submissions. Within the 25 institutions that submitted a BC, 18 planned on the same CT (3D) while the remaining planned on either one phase of the 4DCT or the average 4DCT images of the same patient. As the geometrical data linking the 3D dataset to the 4D datasets were not incorporated in the data sent back from the institutions and we could not assume that the institutions actually used the provided link, we limited our analysis of common and encompassing volumes and Dice coefficients to the 18 submissions planned on the same 3D dataset as requested by the guidelines even if dose calculation and therefore delineations were tolerated on the 4DCT set.

Evaluation of the delineation variations and influence on DVH results

Common and encompassing volumes were generated for the ITVs and PTVs (Fig. 1). To evaluate if the gold standard (GS) CTV

Table 1
Target volume coverage criteria.

	Per protocol	Acceptable variation	Unacceptable variation
PTV[&] (in case OAR constraints can be met)	V60 Gy ≥95% and V54 Gy ≥99% 110% ≤ Dmax < 120%	**x**	V60 Gy <95% or V54 Gy <99%
PTV (reduced prescription dose 8*7 Gy when OAR dose constraints are reached)	V56 Gy ≥95% V50.4 Gy ≥99%	x	V56 Gy <95% V50.4 Gy <99%
PTV (in case inhomogeneous PTV dose is used when OAR overlap with PTV)[1]	V48 Gy ≥95%	**x**	V48 Gy <95%
CTV[*] or ITV (in case inhomogeneous PTV dose is used when OAR overlap with PTV)	V60 Gy ≥95% and V54 Gy ≥99%	**x**	V60 Gy <95% or V54 Gy <99%
GTV max to 0.5 cc	110%–120%	100–110% or 120%-130%	<100% OR >130%

[1] Dose within the PTV may be reduced until the planned dose to the OAR meets the acceptable variation criteria for the OAR but without reaching unacceptable variation criteria for the PTV constraints.
[&] PTV is formed from ITV adding a minimum margin of 3 mm and maximum margin of 5 mm for set-up uncertainties. In case a GTV/CTV to PTV expansion is used additional margins are needed.
[*] In the Lungtech protocol CTV is defined to be equal to the GTV.

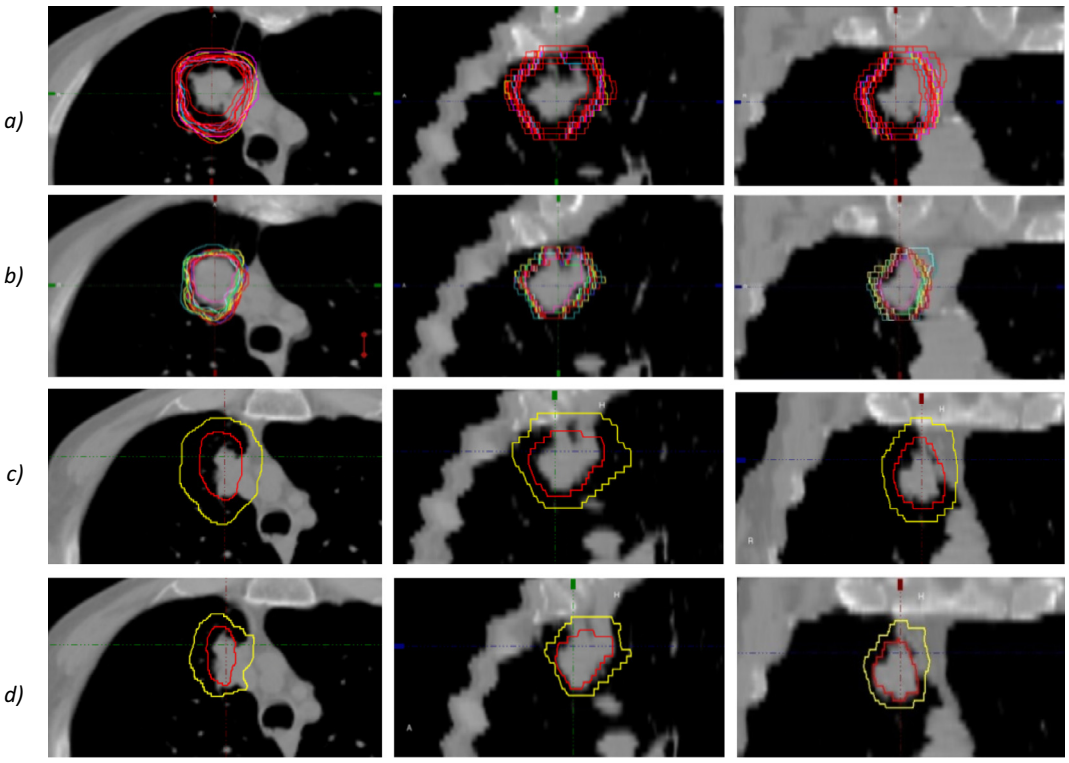

Fig. 1. Axial, Sagittal and coronal images of the target volumes. Delineated PTV and ITV volumes are shown (a and b), as well as their respective common (red) and encompassing (yellow) volumes (c and d). (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this article.)

volume deviated significantly from the submitted CTV volumes, the distribution of the volumes was first tested for normality using a Shapiro-Wilks test. A two-sided t-test was thereafter used to test if the GS volume deviated from the average remaining CTV volumes. Furthermore, Dice coefficients (Eq. (1)) (DC) were calculated for each combination of two volumes. If all delineations are identical the DC value would be 1. The median and range of the DC per volume was calculated as well as the median and range of the DC comparing each submitted volume only to the GS volume. A statistical signed rank Wilcoxon test was performed to test if these results were statistically different.

Equation1. Dice coefficient

$$DC = 2 * \frac{|A_m \cap A_n|}{|A_m| + |A_n|}$$

A_m = target volume of institution m, A_n = target volume of institution n.

From the GS CTV contours, GS PTV contours were derived for each CTV to PTV margin used in the accepted delineations. As such, the GS PTVs were defined as it would have been in the individual institutions if they would have delineated exactly according to the GS CTVs (neglecting possible small differences in expansion algorithms). Using the submitted dose distributions, the dosimetric coverage of the GS CTVs and PTVs was calculated. As such the influence of the remaining accepted delineation variations among the submissions on the dosimetric coverage of the gold standard target could be quantified.

Results

Initial benchmark evaluations

An overview of the treatment techniques and equipment used within the institutions is given in Annex Table 3. Besides the institution using a Tomotherapy unit which employs the mid-ventilation concept, all others used the ITV concept to generate a treatment plan. At the exception of one institution (institution 5) using an 8 MV beam, all institutions used an energy of 6 MV. One institution reported its intention to adopt a motion restriction approach by abdominal compression in the near future.

Five institutions passed both steps at their first attempt. Five institutions succeeded the delineation step after one resubmission and subsequently passed the dosimetry step at their first attempt. One institution succeeded the delineation after one resubmission and had to resubmit the dosimetry once. One institution failed at the first attempt and did not resubmit their benchmark and did not participate to the study. Four institutions needed two delineation resubmissions but passed the dosimetry at their first attempt. Two institutions resubmitted their delineation twice and their dosimetry once. Two institutions had to resubmit three times their delineation but passed successfully the dosimetry at their first attempt. Two other institutions had to resubmit three times their delineation and once their dosimetry. Two institutions had to resubmit their delineations four times, one passed the dosimetry at their first attempt the other had to resubmit once. One institution had to resubmit their delineations 5 times and had to resubmit dosimetry once. Thus, in total 20 institutions had to revise their delineation at least once and only 7 institutions needed to revise their treatment plan.

Delineations graded as unsatisfactory were related to the proximal bronchial tree in 7 cases, the trachea in 4 cases and the spinal cord, the brachial plexus, the chest wall and the lung in 5 cases. For the proximal bronchial tree, the most frequently observed unacceptable variation was the non-inclusion of the distal 2 cm of the trachea. In one submission the margin from CTV to PTV was 7 mm, exceeding the protocol limit of 5 mm. One submission was unacceptable because the delineated GTV insufficiently covered the tumour volume, while in another submission it was the ITV that was insufficiently covering the cranial and caudal extend of the tumour trajectory.

To accommodate some of the planning challenges presented by centrally located tumours, the protocol stipulates that the dose within the PTV may be reduced until the OARs meet the dose acceptable variation criteria (Annex Table 1). In the benchmark exercise 22 institutions succeeded to reach the OARs criteria without reducing the dose to the PTV, demonstrating thus the feasibility of the regular approach. Nonetheless, two institutions (1 and 5) presented in the annex, felt the need to reduce the prescribed dose in order to better spare the large vessels, although no dose constraints for this OAR were formulated in the protocol. These institutions have been informed that the regular prescription could be obtained for this patient by other institutions. However, their plans were not rejected as institution policies on OAR dose may be followed for the OARs for which the Lungtech protocol did not define criteria.

The inter-institution comparison is performed using the regular prescription of 60 Gy as dose normalisation factor for the DVH analysis as this was the dose level that could be reached according to the protocol. We included the 2 institutions which used the reduced prescription dose in this analysis as this better reflects what happens in clinical practice. The resulting DVHs of all the first plan submissions and of the eventually approved submissions are shown in Fig. 2. Six dose submissions were rejected for an insufficient dose coverage of the PTV. One submission was rejected for a high Dmax (>120% of the prescribed dose) to a too large volume (14 cc).

The mean and range of Dmax values (represented by the $D_{0.5\ cc}$) for the OARs are shown in Table 2a. All OARs Dmax values are significantly lower than their Dmax limit. The variations observed over institutions are displayed as the range of Dmax. The proximal bronchial tree shows the largest variations among institutions with a ratio (highest Dmax)/(lowest Dmax) of 12.

The median and range of target volumes and dosimetric coverage are given in Table 2b. The PTV delineation was always acceptable at first submission. The median coverage by the prescription dose before the review was 94.8% (22.5–97.8) and improved to 95.3% (70.5–99.3) through the exercise. The two distributions were significantly different based on a Wilcoxon-rank test ($p = 0.05$).

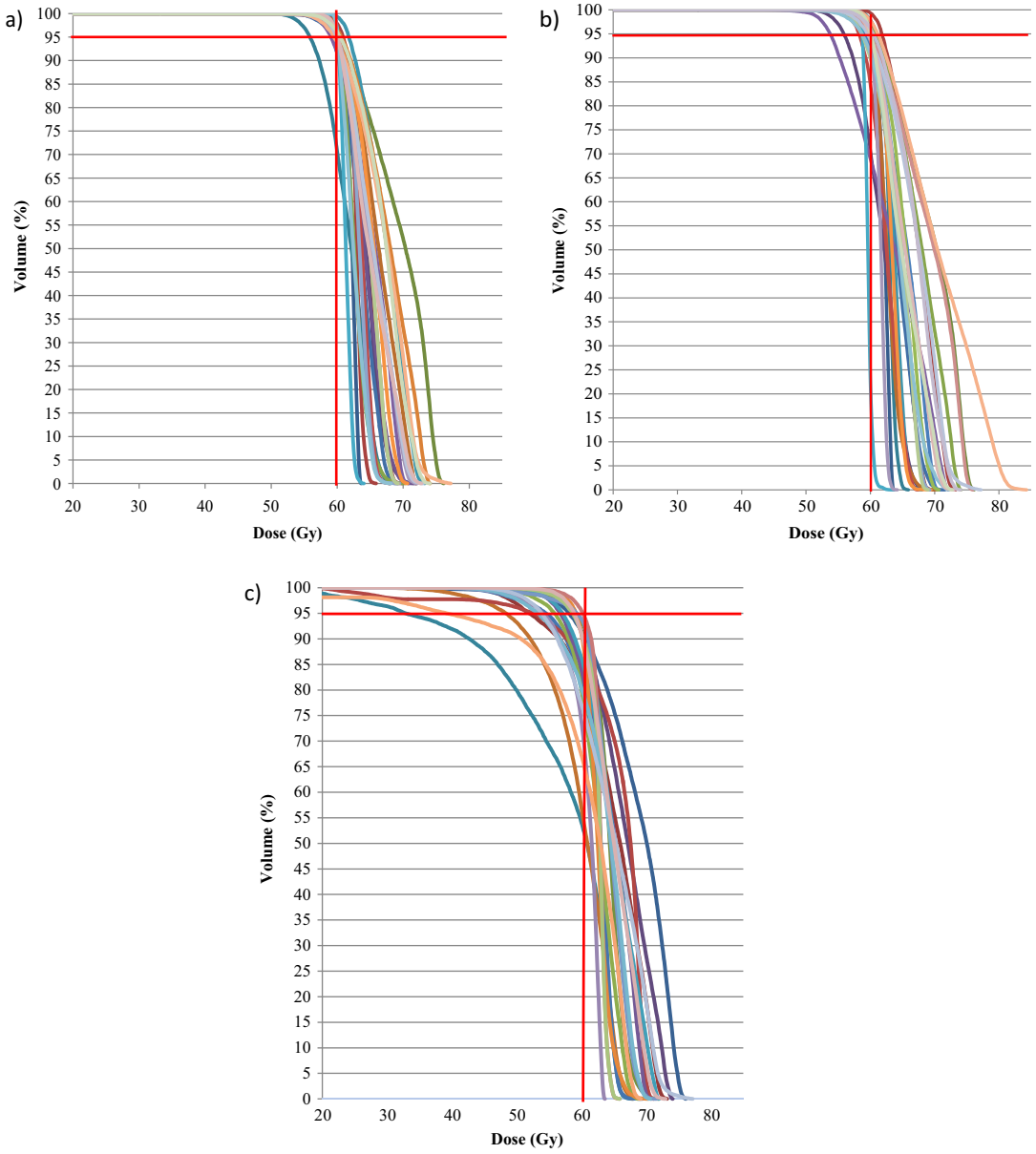

Fig. 2. (a) PTV DVHs of the 24 accepted submissions, (b) PTV DVHs at first submission. (c) DVHs of the 18 accepted dose submissions within the gold standard contours.

Table 2
Dose to OARs (a) and target volumes (b).

OAR	Dmax (Gy)		Dmax (Gy) mean (range)
(a)			
Spinal cord	32		10.2 (0.6–26.1)
Oesophagus	40		18.4 (8.5–23.8)
Brachial plexus	38		0.8 (0.3–2.8)
Proximal trachea	44		37.9 (15.7–41.4)
Proximal bronchus tree	44		28.5 (2.6–36.7)
Lungs-CTV	no restriction but recording of DVH data for toxicity evaluation		66.9 (46.7–73.8)
Chest wall	No restriction but recording DVH data for toxicity evaluation		50.7 (35.27–66.7)
	Volume (cc) median (range)	V60 Gy (%) median (range)	V54 Gy (%) median (range)
(b)			
First review PTV	27.8 (15.2–46.5)	Not applicable	Not applicable
First review plans PTV	27.4 (14.7–46.7)	94.8 (22.5–97.8)	100 (94.4–100)
Accepted plans PTV	27.4 (14.7–46.7)	95.3 (70.5–99.3)	100 (98.4–100)
Accepted plans ITV	12.7 (6.5–18.3)	99.7 (94.9–100-)	100 (100–100)
Gold standard PTV	28.09	83.3 (54.2–93.9)	95.2 (71.5–99.6)

All institutions passed the dose specifications based upon their own contours in their final submission (Fig. 2a), considering a tolerance of 1% in dose and 1% in volume to account for software differences in DVH calculation between the institutions and the central reviewing platform.

Evaluation of the delineation variations and influence on DVH results

The PTV volumes were normally distributed (Shapiro–Wilks test) and the GS CTV volume was not significantly different from the CTV volumes of the other institutions (two-sided student test $p = 0.41$).

As presented in Fig. 2c, using the gold standard contours only one institution among the accepted submissions remained compliant with the dose coverage criteria. These GS DVHs show that the effect of target delineation variations on the DVHs is larger than of variations in planning techniques.

As an example, the DVHs of the PTV of the rejected and accepted treatment planning submissions of one institution are displayed in Fig. 3. Based on their own delineations the dosimetric improvement between the two plan submissions is considerable (from V60 Gy = 84.1% to V60 Gy = 95.1%). However, the improvement is drastically less within the gold standard PTV (from V60 Gy = 70.2% to V60 Gy = 74.4%). The larger improvement seen in the DVH based on the submitted volumes compared to the improvement seen in the GS volumes is caused by an increase in the dose in the volume located outside of the GS contours but inside the submitted volumes.

In addition, a study was performed to see if the size of the institutions own PTV correlated with the dosimetric coverage within the adapted GS-PTV contours (Fig. 4). As can be expected, the coverage of the GS-PTV generally increased with increasing institution planned PTV, although this correlation was not strong.

Delineation analysis

The delineated PTVs and ITVs are depicted in Fig. 1 as well as the encompassing and common volumes. The common volume is less than a quarter of the volume encompassing all contours for both PTV (12.5 cc vs. 52.1 cc) and ITV (5.3 cc vs. 21.8 cc) structures. The range of accepted target volumes reflects not only the variation of submissions, but also the uncertainty around what the reviewing clinicians thought was acceptable when reviewing.

In Table 3 the median Dice coefficients are presented. The median Dice coefficients are 0.82 and 0.79 for the PTVs and the ITVs, respectively. The median Dice coefficients of the OARs vary widely

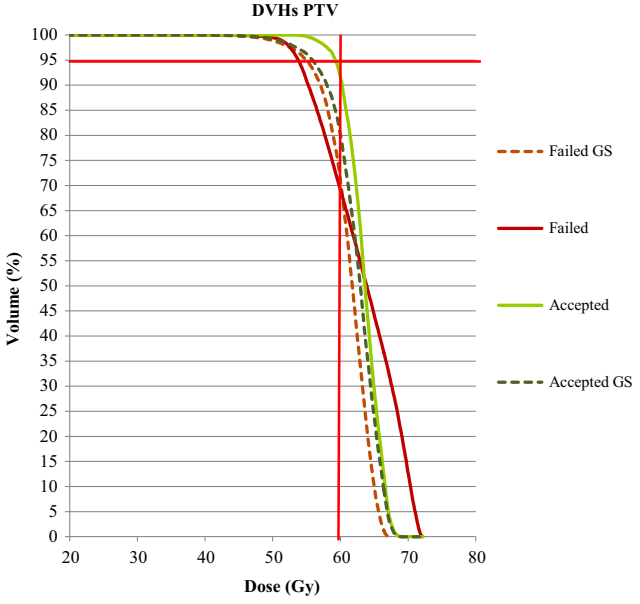

Fig. 3. Failed first dosimetric attempt and final accepted submission respectively within the institution own contours and within the gold standard contours.

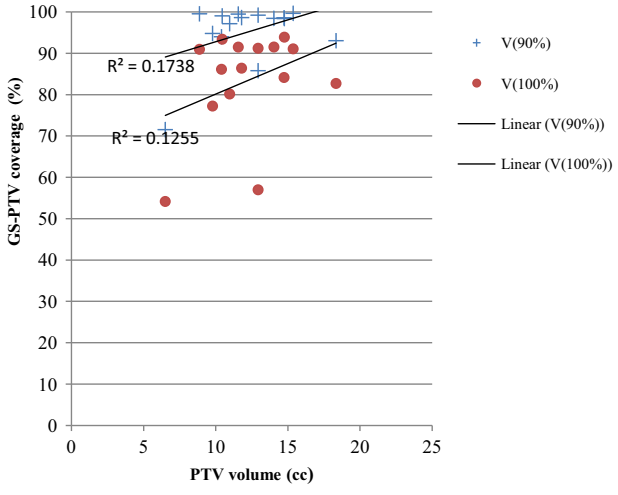

Fig. 4. PTVs' initial volume as a function of the coverage of the GS-PTV.

Table 3
Median and range of Dice coefficients.

	Spinal cord	Proximal bronchial tree	Trachea	Oesophagus	Brachial plexus	ITV	PTV
Over all contours	0.89 (0.71–0.95)	0.62 (0.35–0.83)	0.80 (0.63–0.92)	0.71 (0.61–0.82)	0.25 (0.01–0.54)	0.79 (0.52–1.00)	0.82 (0.48–0.92)
Compared to gold standard	0.76 (0.71–0.80)	0.74 (0.44–0.81)	0.82 (0.72–0.90)	0.71 (0.65–0.75)	0.24 (0.01–0.47)	0.80 (0.69–0.84)	0.85 (0.67–0.91)

and are lowest for the brachial plexus (0.25). The Wilcoxon signed rank test showed no significant difference between the Dices coefficients comparing each institution to each other one and the Dice coefficient calculated between each institution and the GS delineation ($p = 0.54$).

Discussion

This study reports on a two-step review of the Lungtech benchmark case performed by 25 institutions. Foremost, it should be acknowledged that the V60 dose coverage improved significantly ($p = 0.05$) between the first and final submissions, showing the efficacy of the RTQA procedure in improving the trial quality.

The results show that the delineation step was the most challenging with 20 out of 25 institutions failing at their first attempt. In comparison, the planning exercise was graded unacceptable in only 7 cases. Planning was mainly rejected because of a too low dose covering the PTV. This was also the most common dosimetry comment given in the recently published benchmark case results of the Chisel trial on SBRT for early stage lung cancer [9]. Therefore, the delineation was the main source of plan rejection as it is in many other benchmark case exercises [10,11]. The fact that the planning acceptance rate was higher than the delineation acceptance rate might be related to the fact that no OARs with constraints were in direct proximity of the tumour, making planning constraints relatively easy to be met. However, the tumour was in reasonable proximity to the large vessels and some institutions reduced their prescription dose based on this. The Dice coefficient in relation to the gold standard volumes of the OARs varied widely depending on the OAR. For example, the range of Dice coefficients for the brachial plexus delineation was much wider than the range of Dice coefficients for the spinal cord. It should be recognised that the low Dice score of this organ is probably also caused by the far distance of the organ from the target volume, making this delineation not an important part of the dosimetric evaluation of the benchmark. This might have led to a lower attention paid to the contouring of this volume by the institutions. Nonetheless the distance from the tumour did not influence the attention paid to the review of this OAR as 4 institutions were rejected based on an incorrect or absent delineation of the brachial plexus.

For this trial, delineation guidelines were based on a recent and detailed delineation atlas, constructed in supine position with the arms above the head as similar to the BC patient position. Several reasons may have contributed to the high rate of unacceptable delineations. As institutions participating in the trial were not required to have previous experience with treatment of this patient group (central NSCLC) or with the use of the specified delineation atlases, the delineations requested might have been more complex than their clinical routine delineations. However, we did not collect any information on the level of experience of the participants to support our suggestion nor did we ask weather or not the PET-CT images were used during delineation. As an example, in the article of Peulen et al. on target delineation variability in lung SBRT [12], the inter observer variations were small but the observers were all radiation oncologists from institutions with ample experience. In addition, the differences seen in the brachial plexus case could probably have been reduced using MR imaging [13]. Nonetheless, other trials reported similar failure rates despite strict and detailed delineation guidelines [14]. The EORTC has implemented delineation teaching sessions in various new trials to provide training and hopefully reduce the percentage of unacceptable delineations.

In addition to the analysis of the BCs itself, which has been done for many previous trials as well, we also generated DVHs based on gold standard PTVs. While with the submitted DVHs the differences in plan quality could be quantified, the gold standard DVHs were used to quantify the impact of acceptable delineation variations within the submission on target coverage. As a result, it was found that the differences in accepted delineations significantly increased the spread of the DVHs and would have led to unacceptable PTV dose variations if such analysis would have been part of the prospective pass/fail criteria. How this would impact the dose actually delivered to the CTV is certainly also of interest but is considered beyond the scope of this article. The benchmark case exercise conduced in preparation for the PORTEC-3 trial, also large differences in contouring of the CTV were observed [15]. However, these did not translate into large variations in dosimetry. A correlation study was performed comparing the institution initial PTV volume and the plan dosimetric performances using the golden standard contours. The exercise demonstrated as expected that larger initial PTV volumes showed better agreement with the dose specifications when using the golden standard contours. This information might be useful to determine if the review process could be accelerated. To fast track benchmark reviews, distance thresholds to the gold standard contours could be set to allow plans to be reviewed simultaneously for delineation and planning assuming the dosimetric acceptance criteria then take into account the influence of acceptable delineation variations.

The number of patients needed in a trial does not usually explicitly take into account the effect of contouring and planning uncertainties on the trial results. Nonetheless, several articles have demonstrated that the dose delivered in multicentre trials may vary significantly among institutions, pleading for more QA and arguing that toxicity and local control rates currently reported are hampered by such variations [2]. These findings demonstrate the need to enter benchmark exercises in trial routine and to standardise the review of this exercise in terms of parameters reported. Under the hypothesis that delineation and dosimetric variations could uniformly and systematically be reported and correlated to tumour control and toxicity, the number of patients needed for a trial could be tailored according to this quantified variation. Proper prospective RTQA combined with delineation workshops will reduce this variation and thus smaller number of patients would be needed to answer the trial question. As such, trial RTQA would not only improve the treatment of the patients but would also reduce overall trial costs and time.

In conclusion, the data from the BC gave valuable insights into the techniques currently employed for treating centrally located NSCLC and the level of congruence for delineation and treatment planning of such patients. The study demonstrates that QA improved trial quality significantly. Nonetheless the study also shows that delineation was the main source of unacceptable variations in this trial. Moreover, the variation in tumour delineation

within the accepted contours resulted in a substantial increase in the spread of the target coverage compared to the spread resulting from the variation in planning and delivery techniques used. As a very conformal plan is shown through the use of the gold standard contour to have limited value if a high level of accuracy in delineation is not achieved, we suggest to modify the dose tolerance criteria in future trials taking into consideration the accuracy of delineation.

Conflict of interest statement

We have no conflict of interest to disclose.

Acknowledgements

The Lungtech trial was supported by the Deutsche Krebhilfe, Germany. A grant from Elekta, Sweden to perform extensive QA is gratefully acknowledged. We thank all the participating institutions for their fruitful cooperation. David Eaton is kindly acknowledged for reviewing the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2018.10.025.

References

[1] Fairchild A, Bar-Deroma R, Collette L, et al. Development of clinical trial protocols involving advanced radiation therapy techniques: the European Organisation for Research and Treatment of Cancer Radiation Oncology Group approach. Eur J Cancer 2012;48:1048–54.
[2] Weber DC, Tomsej M, Melidis C, Hurkmans CW. QA makes a clinical trial stronger: evidence-based medicine in radiation therapy. Radiother Oncol 2012;105:4–8.
[3] Peters LJ, O'Sullivan B, Giralt J, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: results from TROG 02.02. J Clin Oncol 2010;28:2996–3001.
[4] Adebahr S, Collette S, Shash E, et al. LungTech, an EORTC Phase II trial of stereotactic body radiotherapy for centrally located lung tumours: a clinical perspective. Br J Radiol 2015;88. https://doi.org/10.1259/bjr.20150036.10.1186/s13014-015-0567-5. 20150036.
[5] Lambrecht M, Melidis C, Sonke JJ, et al. Lungtech, a phase II EORTC trial of SBRT for centrally located lung tumours – A clinical physics perspective. Radiat Oncol 2016;11. https://doi.org/10.1186/s13014-015-0567-5. 7.
[6] Melidis C, Bosch WR, Izewska J, et al. Global harmonization of quality assurance naming conventions in radiation therapy clinical trials. Int J Radiat Oncol Biol Phys 2014;90:1242–9.
[7] Feng M, Moran JM, Koelling T, et al. Development and validation of a heart atlas to study cardiac exposure to radiation following treatment for breast cancer. Int J Radiat Oncol Biol Phys 2011;79:10–8.
[8] Kong FM, Ritter T, Quint DJ, et al. Consideration of dose limits for organs at risk of thoracic radiotherapy: atlas for lung, proximal bronchial tree, esophagus, spinal cord, ribs, and brachial plexus. Int J Radiat Oncol Biol Phys 2011;81:1442–57.
[9] Kron T, Chesson B, Hardcastle N, et al. Credentialing of radiotherapy centres in Australia for TROG 09.02 (Chisel), a Phase III clinical trial on stereotactic ablative body radiotherapy of early stage lung cancer. Br J Radiol 2018;91. 20170737.
[10] Trada Y, Kneebone A, Paneghel A, et al. Optimizing radiation therapy quality assurance in clinical trials: a TROG 08.03 RAVES substudy. Int J Radiat Oncol Biol Phys 2015;93:1045–51.
[11] Versteijne E, Lens E, van der HA, et al. Quality assurance of the PREOPANC trial (2012-003181-40) for preoperative radiochemotherapy in pancreatic cancer: the dummy run. Strahlenther Onkol 2017;193:630–8.
[12] Peulen H, Belderbos J, Guckenberger M, et al. Target delineation variability and corresponding margins of peripheral early stage NSCLC treated with stereotactic body radiotherapy. Radiother Oncol 2015;114:361–6.
[13] De Ruijsher RD, Faivre-Finn C, Moeller D, et al. European Organization for Research and Treatment of Cancer (EORTC) recommendations for planning and delivery of high-dose, high precision radiotherapy for lung cancer. Radiother Oncol 2017;124:1–10.
[14] Fairchild A, Langendijk JA, Nuyts S, et al. Quality assurance for the EORTC 22071-26071 study: dummy run prospective analysis. Radiat Oncol 2014;9. 248.
[15] Jameson MG, McNamara J, Bailey M, et al. Results of the Australasian (Trans-Tasman Oncology Group) radiotherapy benchmarking exercise in preparation for participation in the PORTEC-3 trial. J Med Imaging Radiat Oncol 2016;60:554–9.