Thorough design and pre-trial quality assurance (QA) decrease dosimetric impact of delineation and dose planning variability in the STRICTLUNG and STARLUNG trials for stereotactic body radiotherapy (SBRT) of central and ultra-central lung tumours

Hoffmann, L.; Persson, G. F.; Nygård, L.; Nielsen, T. B.; Borrisova, S.; Gaard-Petersen, F.; Josipovic, M.; Khalil, A. A.; Kjeldsen, R.; Knap, M. M.; Kristiansen, C.; Møller, D. S.; Ottosson, W.; Sand, H.; Thing, R.; Pøhl, M.; Schytte, T.

Published in:
Radiotherapy & Oncology

DOI (link to publication from Publisher):
10.1016/j.radonc.2022.04.005

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Radiotherapy and Oncology 171 (2022) 53–61

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Thorough design and pre-trial quality assurance (QA) decrease dosimetric impact of delineation and dose planning variability in the STRICTLUNG and STARLUNG trials for stereotactic body radiotherapy (SBRT) of central and ultra-central lung tumours

L. Hoffmann [a,b,*], G.F. Persson [c,d], L. Nygård [e], T.B. Nielsen [f], S. Borrisova [c], F. Gaard-Petersen [c], M. Josipovic [e], A.A. Khalil [a,b], R. Kjeldsen [g], M.M. Knap [a], C. Kristiansen [h], D.S. Møller [a,b], W. Ottosson [c], H. Sand [i], R. Thing [h], M. Pøhl [e,1], T. Schytte [j,k,1]

[a] Dept. of Oncology, Aarhus University Hospital, Denmark; [b] Dept. of Clinical Medicine, Faculty of Health Sciences, Aarhus University; [c] Dept of Oncology, Herlev-Gentofte Hospital, Copenhagen University Hospital; [d] Dept. of Clinical Medicine, Faculty of Health Sciences, University of Copenhagen; [e] Dept of Oncology, Rigshospitalet, Copenhagen University Hospital; [f] Laboratory of Radiation Physics, Odense University Hospital, Odense; [g] Dept. of Oncology, Aalborg University Hospital, Aalborg; [h] Dept of Oncology, Lillebaelt Hospital, University Hospital of Southern Denmark; [i] Dept of Medical Physics, Aalborg University Hospital; [j] Department of Oncology, Odense University Hospital; and [k] Department of Clinical Research, University of Southern Denmark, Odense, Denmark

ARTICLE INFO

Article history:
Received 1 July 2021
Received in revised form 28 March 2022
Accepted 5 April 2022
Available online 11 April 2022

Keywords:
Stereotactic body radiotherapy
Central tumours
Lung tumours
Quality assurance

ABSTRACT

Introduction: SBRT of central lung tumours implies significant risk of toxicity. We are initiating two phase II trials prescribing 56 Gy/eight fractions to PTV, allowing for dose escalation of GTV. We prioritize organs at risk (OAR) constraints over target coverage, making the treatment plans very sensitive to OAR delineation variations. The aim of this study is to quantify the dosimetric impact of contouring variations and to provide a thorough description of pre-trial quality assurance to be used in upcoming trials to provide consistent clinical care.

Materials and methods: Delineation: Seven physicians delineated OAR in three rounds, with evaluations in-between. For each patient case, seven treatment plans, repeatedly using each of the OAR structure sets from the seven physicians, were made and compared to evaluate the dosimetric effect of delineation variability.

Treatment planning: Treatment plans for seven cases were made at six departments in two rounds, with discussion in-between.

Results: OAR delineation variation between centres resulted in high variabilities in OAR dose for simulated plans and led to potential overdosage of the lobar bronchus (constraint: $D_{0.03cc} < 45$ Gy), with maximum doses ranging between 58 Gy (first round), and 50 Gy (third round). For mediastinal tissue, the constraint ($D_{0.03cc} < 45$ Gy) was violated for the majority of the delineations in all three rounds, with maximum doses of 84 Gy (first round), and 72 Gy (third round). For the treatment planning study, the range of the standard deviation for GTV mean dose was 12.8–18.5 Gy (first round) and 2.8–3.5 Gy (second round).

Conclusions: Even small variations in OAR delineation led to high OAR overdosage. The study demonstrates the importance of having extensive QA procedures in place before initiating clinical trials on dose escalation in SBRT.

Delivering high doses to the target volume while avoiding severe toxicity is a cornerstone in modern radiotherapy, particularly for stereotactic body radiotherapy (SBRT) and dose escalation studies using steep dose gradients toward organs at risk (OAR). High rates of severe and even lethal toxicity related to central and peri-hilar structures have been reported in former hypo-fractionated studies [1–8]. For the EORTC 22113-8113 Lungtech-trial on SBRT of central lung tumours, safety-related issues contributed to early closure of the trial [3]. In the HILUS-trial, delivering SBRT to central and ultra-central tumours, 15% of the patients died from causes possibly being treatment related, hereof eight from pulmonal haemorrhage [1]. A meta-analysis from 2019 estimated the risk of ≥grade 3 toxicity after

* Corresponding author at: Department of Medical Physics, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, entrance C, DK-8200 Aarhus N, Denmark.
E-mail address: Lone.Hoffmann@aarhus.rm.dk (L. Hoffmann).
[1] Shared last-author ship.

https://doi.org/10.1016/j.radonc.2022.04.005
0167-8140/

SBRT of central NSCLC to 12% [6].The high risk of lethal toxicity, calls for high demands to study design and quality assurance (QA). For clinical trials with multi-centre collaboration, it is important to secure consistent delineation of target and OAR in addition to consensus on the dose delivery as more coherent interpretation of the trial outcome is obtained by standardisation prior to trial initiation [9–15]. However, trial QA may be insufficient and is often unpublished [1,4,16]. A way to secure high consistency in multi-centre trials, is to setup pre-trial QA for the full treatment planning process and demand all centres to participate [10,17–19].

The risk of radiotherapy-induced toxicity increases when tumour location approaches the hilar and mediastinal structures. The proximity to airways, oesophagus and heart is a major challenge and has become the basis of distinction between peripherally and centrally located tumours [1,6,7,20,21]. For peripheral tumours, the ACROP-guideline recommends to use a fractionation schedule resulting in a biologically effective dose (BED, alpha/beta = 10 Gy) > 150 Gy for the mean dose to the gross tumour volume (GTV) [22]. No consensus has been reached for central tumours [6].

The Radiotherapy Committee of the Danish Oncological Lung Cancer Group (https://www.dolg.dk) is initiating two multi-centre phase II trials evaluating safety and efficacy of thoracic SBRT: STRICTLUNG for SBRT of central tumours and STARLUNG for on-line MR-guided SBRT of ultra-central tumours. Both studies engage a heterogeneous dose prescription of 56 Gy in eight fractions to the planning target volume (PTV), allowing for dose escalation to the GTV. OAR dose constraints have a higher priority than target dose and consequently limit the dose escalation. This is in contradiction to former studies where target coverage was prioritized at the expense of high toxicity levels [1,8,21].

In this work, we report on the protocol design and QA process, that aimed at ensuring consistency in OAR delineation and treatment planning. Former studies on delineation uncertainties report on geometric measures as e.g. Dice similarity coefficient (DSC) and mean surface distance [23–25]. In this study, we bring this a step further and report on the dosimetric consequences of OAR delineation variations in treatments with very steep dose gradients which may potentially lead to lethal toxicity. Heterogeneous dose escalation has been the subject of many clinical trials. However, the dosimetric consequence of delineation uncertainties has not been reported [1,4,26]. The aim is to provide a thorough description of a pre-trial setup to be used in upcoming trials in order to provide consistent clinical care.

Materials and methods

Consensus process for delineation (D)

To ensure consensus on study guideline interpretation, resulting in uniform OAR delineation and treatment planning across centres, a pre-study QA-programme was set up consisting of the following steps:

1. Guideline on the definition of OAR was prepared based on previous publications and circulated to all centres [27–30].
2. First OAR delineation round (D1). Seven physicians from six centres delineated OAR on one case (patient *a*).
3. Consensus meeting to discuss the results of step 2, followed by a refinement of the OAR delineation guidelines.
4. Second OAR delineation round (D2) with a new case (patient *b*).
5. Consensus meeting to discuss the results of step 4. Agreement to create a delineation atlas.
6. Circulation of OAR delineation atlas and consensus thereof.
7. Third delineation round (D3) using a new case (patient *c*).

Target and OAR delineation

The GTV was delineated as visible tumour using a diagnostic PET/CT and a free breathing planning 4DCT-scan with intravenous contrast. GTV and PTV were delineated on the mid-ventilation phase on all cases in advance. GTV to PTV margin was 4 mm transversal, and 5 mm longitudinal [31].

The OAR includes trachea, main bronchi including intermediate bronchus, lobar bronchi, oesophagus, aorta, heart, spinal cord, lungs, chest-wall, and connective-tissue. The latter defined as mediastinal and hilar tissue not otherwise categorized as OAR in order to avoid hot-spots in this area. OAR were delineated on the mid-ventilation phase of the 4DCT scans. Guidelines are included in the Supplementary materials.

Central versus ultra-central tumour

Tumours are considered ultra-central when GTV is located <0.5 cm in any direction from the main bronchi, intermediate bronchus or the oesophagus at the diagnostic CT-scan. Ultra-centrally located tumours will be treated on MR-linacs enabling online plan adaptation. Tumours are considered central when located within 0.5–2.5 cm from the trachea, main bronchi, intermediate bronchus or the oesophagus, in addition to tumours located <0.5 cm from the lobar bronchi, spinal cord, heart and aorta. Central tumours will be treated on standard-linacs using pre-treatment setup with daily image guidance based on GTV-match. Daily review of OAR positional changes is required and in case dosimetric changes are suspected, plan adaptation must be performed.

For each of patient *a-c*, seven physicians investigated if the tumour was considered ultra-central or central. The patients were selected based on tumour being approximately 5 mm from one OAR.

Consensus process for treatment planning (TP)

1. First treatment planning round (TP1) of five cases (patients *1–5*).
2. Consensus meeting to discuss the results of step 1, followed by refined treatment planning constraints.
3. Second treatment planning round (TP2) using two new cases (patients *6–7*).
4. Consensus meeting to discuss the results of step 3.

Treatment planning

Constraints for OAR were derived from published work [26,32–35], see Table 1. GTV, PTV and OAR were delineated before the cases were sent to all centres for treatment planning. The plans were made using either IMRT or VMAT. The treatment planning systems (and algorithms) used were, Eclipse (Acuros), Pinnacle (Collapsed cone) and RayStation (Monte Carlo). The grid size was 2.5 mm. During TP1, the dose to the GTV should be as high as possible and no constraints were set for maximum dose to the GTV. This strategy was applied to investigate the upper limit achievable for dose to GTV respecting dose to OAR. Doses above clinical meaningful limits were accepted for TP1. At least 95% of PTV should be covered by 95% of the prescription dose of 56 Gy. Dose was planned to be delivered in eight fractions. Constraints to OAR had higher priority than target dose and coverage. After the first consensus meeting additional constraints were set: a maximum dose of 110 Gy to $D_{0.3cc}$ of GTV ($GTV_{D0.3cc}$) and a maximum mean GTV (GTV_{mean}) dose of 85 Gy (BED_{10} = 175 Gy) following the ACROP recommendations suggesting BED_{10} > 150 Gy [22]. This was applied for TP2. All plans were collected for analysis at one centre.

Table 1
Dose constraints for organs at risk.

OAR constraints	α/β	D_{max} (Gy)	EQD_2 (Gy)
Spinal cord	2	$D_{0.1cc}$ < 32	48
Oesophagus	3	$D_{0.3cc}$ < 40	64
Trachea	3	$D_{0.3cc}$ < 42	69.3
Main bronchi including interm. bronchus	3	$D_{0.3cc}$ < 42	69.3
Heart	3	$D_{0.3cc}$ < 43	74
Lobar bronchi	3	$D_{0.3cc}$ < 45	77.6
Connective-tissue	3	$D_{0.3cc}$ < 45	77.6
Aorta	3	$D_{0.3cc}$ < 50	77.6
Chest-wall	3	$D_{0.3cc}$ < 53	102

Maximum dose constraints for OAR.

Data analysis for delineation of OAR

For each of patient *a–c*, the delineations of the OAR were collected in one structure set.

For patient *a*, contour overlap mapping (MIM Software Inc v6.8.9) was used to illustrate differences between the delineated structures using a heat map for trachea, main bronchi, intermediate bronchus, and five lobar bronchi and a heat map for connective-tissue.

For all patients, the delineations of the trachea, main bronchi, intermediate bronchus and lobar bronchi were compared between centres (MIM Software Inc v6.8.9). The DSC, and mean and maximum undirected Hausdorff distances (H_{mean} and H_{max}) were calculated [36,37]. All metrics were calculated pairwise between all possible pairs for each patient.

For patients *a*(D1), *b*(D2) and *c*(D3), treatment plans with six IMRT beams were created and calculated using Acuros (Eclipse v15.6). Firstly, the plans were optimized respecting the constraints for the OAR as delineated by one centre. GTV_{mean} was increased as much as possible respecting GTV_{mean} < 85 Gy and $GTV_{D0.3cc}$ < 110 Gy. Secondly, doses to OAR for these plans were assessed for delineations from each of the other centres to estimate the dosimetric effect of delineation variations. Thirdly, the plans were copied and new optimizations were conducted using the OAR delineated by a second centre, still respecting the constraints for the OAR, while dose to the GTV was increased as high as possible (respecting GTV_{mean} < 85 Gy). The process was repeated until seven treatment plans were made for each patient, each of them based on the delineations of one centre and only optimized to the OAR as delineated by that specific centre. All plans were made by the same medical physicist using the exact same treatment planning parameters. Mean dose and dose to 95% of GTV and PTV volume ($D_{95\%}$) were compared between the seven treatment plans using box plots. Dose to selected OAR was compared between all 49 (seven treatment plans times seven OAR structure sets) plans using box plots, to visualize the dosimetric effect of the diversity in delineations between centres. The standard deviation (SD) of GTV_{mean} and PTV_{mean} doses were considered a measure of the dosimetric effect of the inter-observer variation. The lower the SD, the higher concordance between observers.

Data analysis for treatment planning

Doses to GTV, PTV, main and lobar bronchi, heart, and connective-tissue were compared for the treatment plans for the seven patients (*1–7*) and the results were displayed using boxplots. Each of the patients had six plans made at six centres.

Statistics

A union of trachea, main, intermediate, and lobar bronchi was made for the delineated structures of each centre. Differences in DSC, H_{mean}, and H_{max} between D3 and D1 or D2 was investigated with a Mann-Whitney *U* test for continuous variables using Matlab (version 2019a). A *p*-value <0.05 was considered statistically significant.

Results

Based on D1, D2 and consensus meetings, an atlas patient case was generated for the OAR (Fig. 1). The heatmap (Fig. 2) illustrates the delineations of the airways and the connective-tissue for patient *a*. For the airways, deviations between the centres were primarily observed at the lateral ends of the lobar bronchi. For the connective-tissue, larger differences were detected on delineation of tissue abutting the tumour. The median DSC for the unified structure of trachea, main, intermediate and lobar bronchi was high. It was significantly higher ($p < 0.05$) for D3 (0.88), than for both D1 (0.87) and D2 (0.86). The median value of H_{mean} was small. It was significantly lower/shorter ($p < 0.05$) for D3 (0.66 mm), than for both D1 (1.04 mm) and D2 (1.10 mm). The median H_{max} was significantly lower/shorter ($p < 0.05$) for D3 (12.9 mm), than for both D1 (21.0 mm) and D2 (18.8 mm). Boxplots of DSC and H_{mean} are shown in the Supplementary materials.

Differences in delineation of the OAR between centres had an impact on both mean dose and $D_{95\%}$ to GTV and PTV leading to high variability in target coverage and mean dose. This is illustrated for the three delineation rounds (D1–3) in Fig. 3. The SD of GTV_{mean} decreased for each delineation round 7.5 Gy (D1), 4.0 Gy (D2), and 1.2 Gy (D3) showing increased concordance between observers.

The dose to ipsilateral and main bronchi, connective-tissue and chest-wall are presented in Fig. 3 for all plans ($n = 7$) created based on the delineations of all physicians ($n = 7$) showing high variability. In six plans for each of D1, D2 and D3 the variability in the delineation lead to potential overdosage of the lobar bronchus (constraint: $D_{0.3cc}$ < 45 Gy). The maximum dose was 58.0 Gy, 49.8 Gy and 49.6 Gy in D1, D2 and D3, respectively. Likewise, connective-$tissue_{D0.3cc}$ > 45 Gy for approximately 60% of the delineations in each of the three delineation rounds. The maximum dose was 84.3 Gy, 84.7 Gy and 71.9 Gy in D1, D2 and D3, respectively. Ten, fourteen and four plans led to $D_{0.3cc}$ > 70 Gy in D1, D2 and D3, respectively. In all cases, each plan complied with the constraints for the OAR used for optimization of the plan. The deviations were solely caused by the variations in the delineation from the other physicians. All other OAR had doses <30 Gy in all plans.

The definition of ultra-central or central tumour is highly dependent on the specific delineations. Patient *a* fulfilled the definition of an ultra-central tumour based on the OAR delineations from three physicians and the definition of a central tumour based on OAR delineation from four physicians. For the remaining patients, the scorings were: patient *b* (7 ultra-central), patient *c* (1 ultra-central, 6 central).

The tumour positions for patients *1–7* are shown in Fig. 4 in addition to dose distributions from one of the centres.

For the treatment planning (TP) study, no constraints were set for $GTV_{D0.3cc}$ or GTV_{mean} dose in TP1, resulting in a large variability in $GTV_{D0.3cc}$ (range 65–195 Gy) and GTV_{mean} (range 59–126 Gy). All plans complied with the maximum doses to the OAR except for the connective-tissue and chest-wall, where overdosage was seen (maximum 54 Gy and 64 Gy). In TP2, less variability in $GTV_{D0.3cc}$ (range 64–109 Gy) and GTV_{mean} (range 54–85 Gy) was observed, due to an added constraint for these structures (Fig. 5). All plans complied with the maximum doses to the OAR except for the connective-tissue, where slight overdosage was seen (maximum 47 Gy). SD in the GTV_{mean} decreased at TP2. SD ranged from 12.8

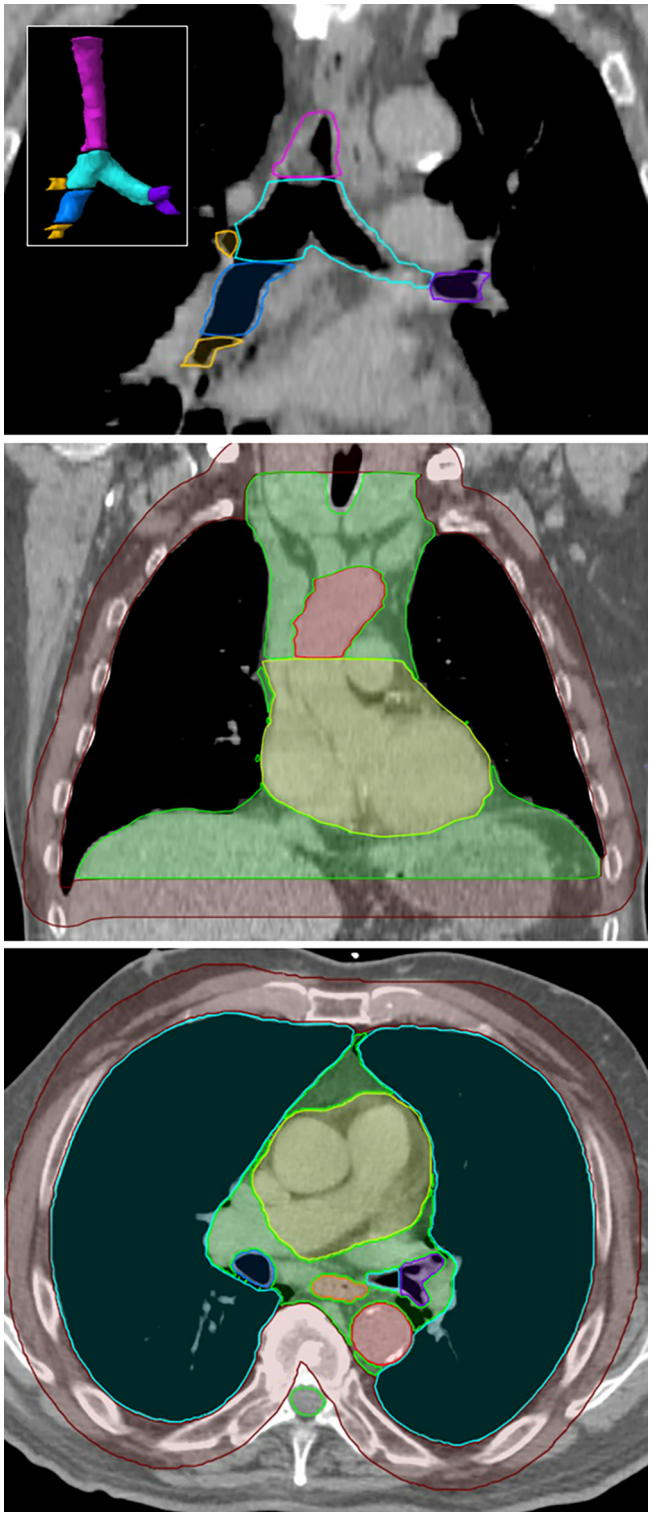

Fig. 1. Delineation of organs at risk. Upper panel: trachea (purple), main bronchus (turquoise), intermediate bronchus (blue), left lobar bronchi (lilac), right lobar bronchi (orange). Centre panel: heart (yellow), aorta (red), connective-tissue (green), chest-wall (brown). Lower panel: heart, connective-tissue, aorta, chest-wall, main bronchi, intermediate bronchus, left lobar bronchus, oesophagus (orange), spinal cord (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

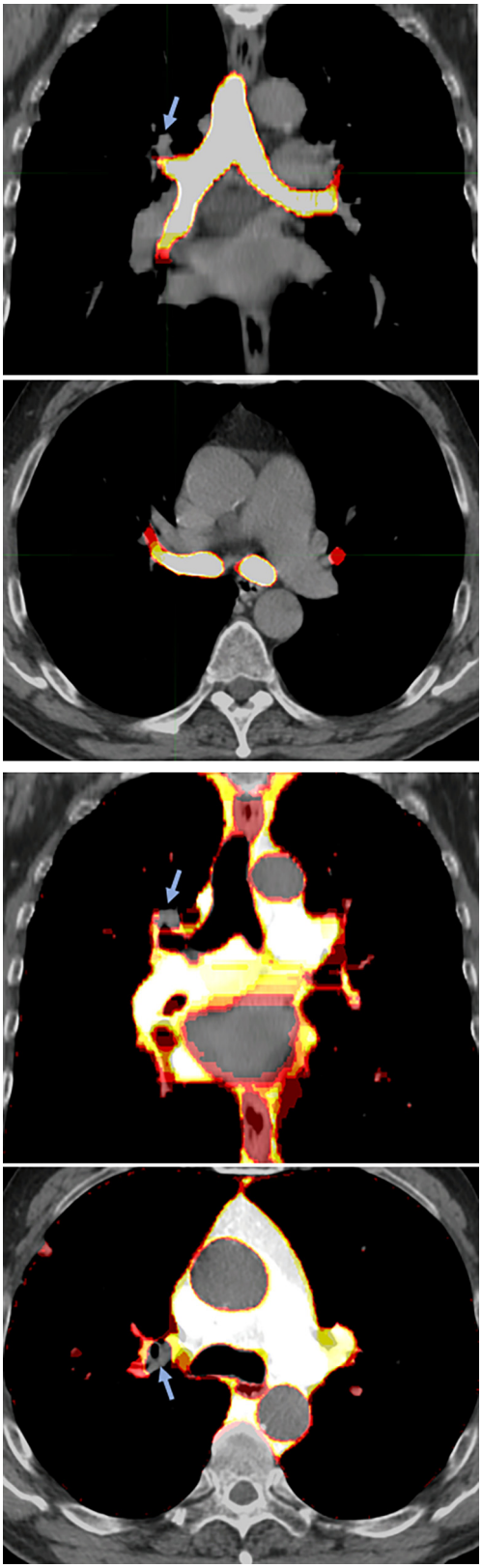

Fig. 2. Heat map going from white (overlap of all delineations) to dark red (only delineated by one centre) of the seven delineations of the trachea, main bronchi, intermediate bronchus and lobar bronchi (upper panels) and the connective-tissue (lower panel). Blue arrow shows position of the tumour in the three images, were tumour is in the selected view. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

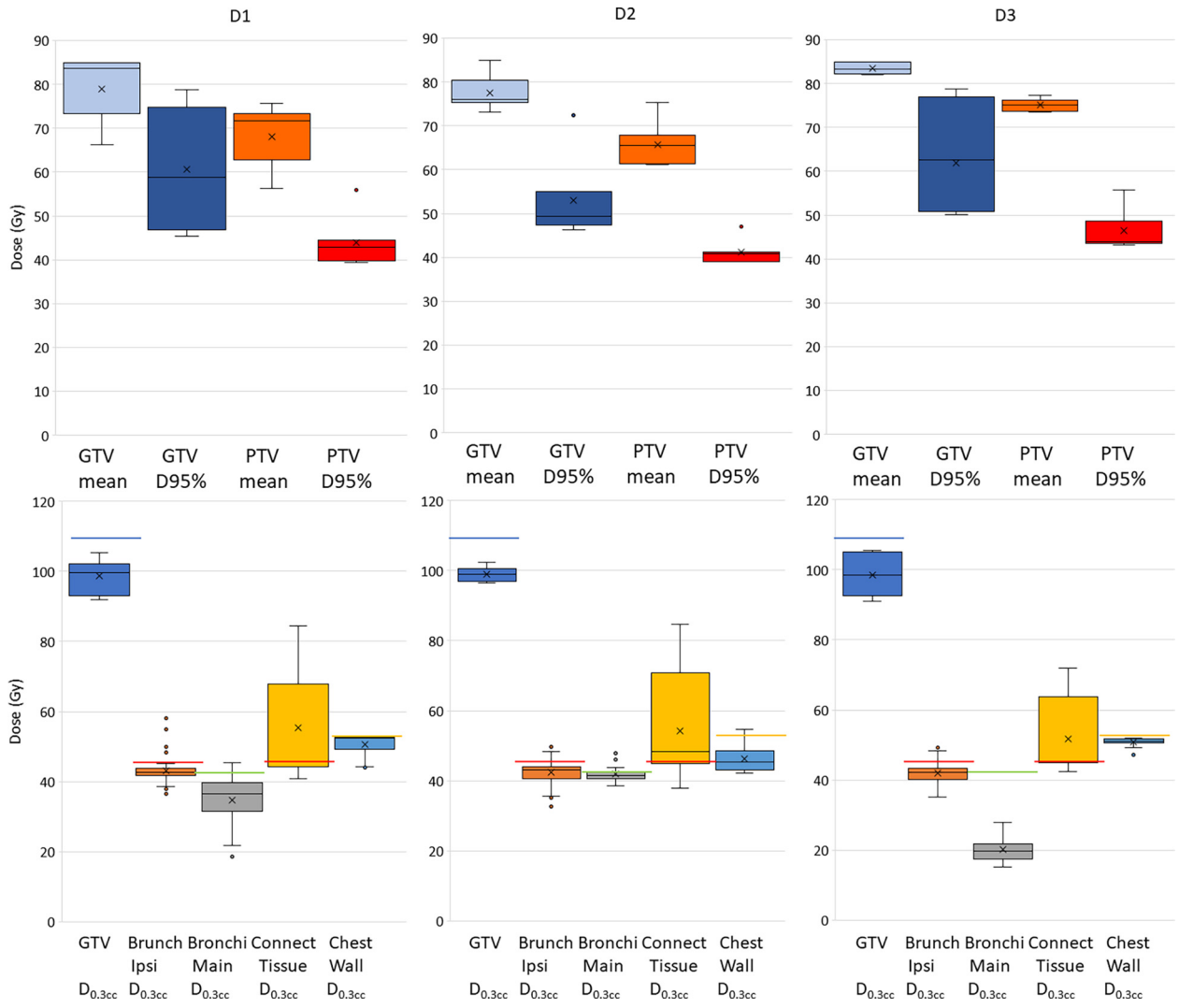

Fig. 3. Upper pane: Boxplot showing mean and $D_{95\%}$ dose to GTV and PTV resulting from the seven dose plans for patient *a* (delineation round one, D1), patient *b* (delineation round two, D2) and patient *c* (delineation round three, D3). Lower pane: Boxplot showing $D_{0.03cc}$ to selected OAR for the same patients. Each boxplot constitutes seven times seven measures. Upper limit at 110 Gy (GTV $D_{0.3cc}$) is shown as a blue line, at 45 Gy ($D_{0.3cc}$ for ipsilateral bronchus and connective-tissue) as red line, at 42 Gy ($D_{0.3cc}$ for main bronchi) as green line, and 50 Gy ($D_{0.3cc}$ for chest-wall) as orange line. Box plot: median (horizontal line), first and third interquartile ranges (box), minimum/maximum (whiskers), and outliers (o). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to 18.5 Gy in TP1 and from 2.8 to 3.5 Gy in TP2. Likewise, the SD for connective-tissue$_{D0.3cc}$ decreased from 2.0 (TP1) to 0.5 Gy (TP2). Doses to target and selected OAR are shown in Fig. 5 for patients *1–7*. In all plans, PTV $V_{95\%}$ was not fulfilled. For all plans, the maximum dose was located in the GTV.

Discussion

We hereby present the protocol design and pre-trial QA for the multi-centre phase II clinical trials STRICTLUNG and STARLUNG. Pre-trial QA and standardisation of delineation and treatment planning is essential [11–13,15]. A high variation in fractionation, dose prescription, distribution and delivery requirements, including definition and delineation of target and OAR for SBRT has been described previously [38–41]. This represents a challenge when comparing and interpreting the results from multiple studies, but also between radiotherapy centres within a single clinical study [1,41]. The current SBRT practice and guidelines, lacking a clear consensus on the above-mentioned parameters will therefore inevitably lead to large inter- and intra-institutional variation. This highlights the importance of commencing a thorough QA-programme before initiating multi-centre trials, minimizing these variabilities. Additionally, the QA-programme should be followed by checks throughout the trial period [9,11,14]. In this study, we prospectively store clinical data and treatment plans, making the data immediately available for analysis in case high rates of severe toxicity require the trial to be halted.

In the present trial, we focus on strict constraints to OAR having higher priority than target coverage and hereby limit dose escalation. This is contradictory to former trials which have shown a high level of severe toxicity including fatal toxicity [1,8,21]. In the recently published HILUS-trial, only dose constraints to spinal cord, trachea and contralateral main bronchus had higher priority than PTV-coverage [1]. The median value of the minimum equivalent dose in 2 Gy-fractions (EQD2) to the hottest 0.2 cc ($D_{0.2cc}$) of the lumen of trachea plus ipsilateral main bronchus was 65 Gy (1.0–207 Gy). In the present trial, $D_{0.3cc}$ < 69.3 (EQD2) for trachea and main bronchus was set as constraint to prevent severe toxicity.

There is no definite consensus about the characterization of central versus ultra-central tumours [1,6,40]. In the present trial, the treatment modality was determined from the tumour position.

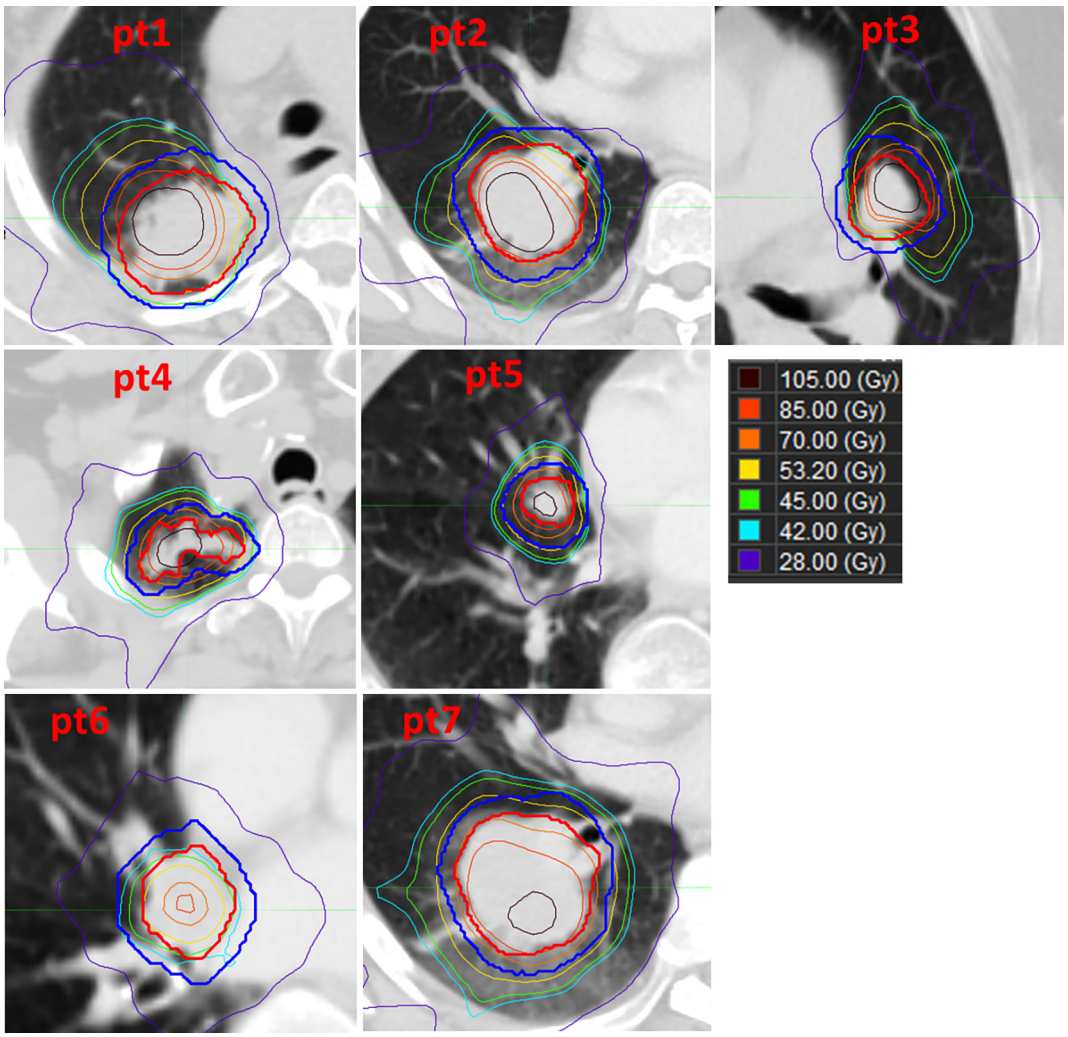

Fig. 4. Tumour position and dose distribution for a treatment plan from one of the six centres. Iso dose levels at 105 Gy, 85 Gy, 70 Gy, 53.2 Gy, 50 Gy, 45 Gy, 42 Gy and 28 Gy are shown (thin lines). Images are shown for patient *1–5* in TP1 and patient *6–7* in TP2. GTV contour is shown in red (thick line). PTV contour is shown in blue (thick line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Ultra-central tumours will be selected for treatment on MR-linacs using on-line adaptation. As MR-linacs are available at a few centres only, consistent referral from all centres requires that the distinction between ultra-central and central tumour location is clearly described. We saw a rather large discrepancy in classification of the tumours of patient *a-c*. In the prospective trial, clinicians will be able to discuss classification of tumours at multi-centre teleconferences.

Three delineation rounds (D1–D3) were accomplished, with the creation of a consensus atlas between D2 and D3 entailing a statistically significant improvement in DSC, H_{mean}, and H_{max}. In all three delineation rounds high DSC and low H_{mean}/H_{max} values were found, showing good agreement on the delineation of the airways. Major differences were only observed at the lateral ends of the lobar bronchi (Fig. 2). The atlas was introduced as written guidelines still lead to differences in the delineations. The convergence in delineation of the lobar bronchi manifested in less variability in dose to the ipsilateral bronchus at D3 compared to D1 (Fig. 3). Hereby showing, that even small variations in delineations may result in dosimetric differences in central SBRT of lung tumours. This is due to coincidence of the region of highest geometrical uncertainty being located exactly at the sharpest dose gradient.

The physicians judge tissue in the hilum differently, either as lung or connective-tissue. This variability manifested in high dosimetric variability. In all delineation rounds, the connective-tissue constraint was violated for some delineations. However, the maximum dose decreased from D1 to D3 showing higher agreement between the delineations.

We brought the effect of the geometric contour variability a step further than suggested in guidelines and former studies by converting the geometrical variations into their dosimetrical impact [11,15,23–26]. We observe, that in heterogeneous dose escalation with steep gradients, even minor deviations in OAR contours lead to large dosimetric variations and hereby risk of severe toxicity. Thus, delineation uncertainty should be reduced as much as possible. The dosimetric effect of variations in tumour delineations, was investigated as part of the pre-trial QA of the Lungtech trial. It was found that delineation variations lead to higher dosimetric variability than variations in the multi-centre treatment plans [9,14]. Over-dosage of the heart was observed in a retrospective evaluation of delineations submitted for the CONVERT trial. The heart was re-outlined according to the gold standard trial protocol. Differences in delineation of the OAR may result in inconclusive outcome for the toxicity evaluation which is the primary outcome in the STRICTLUNG and STARLUNG trials. Daily patient

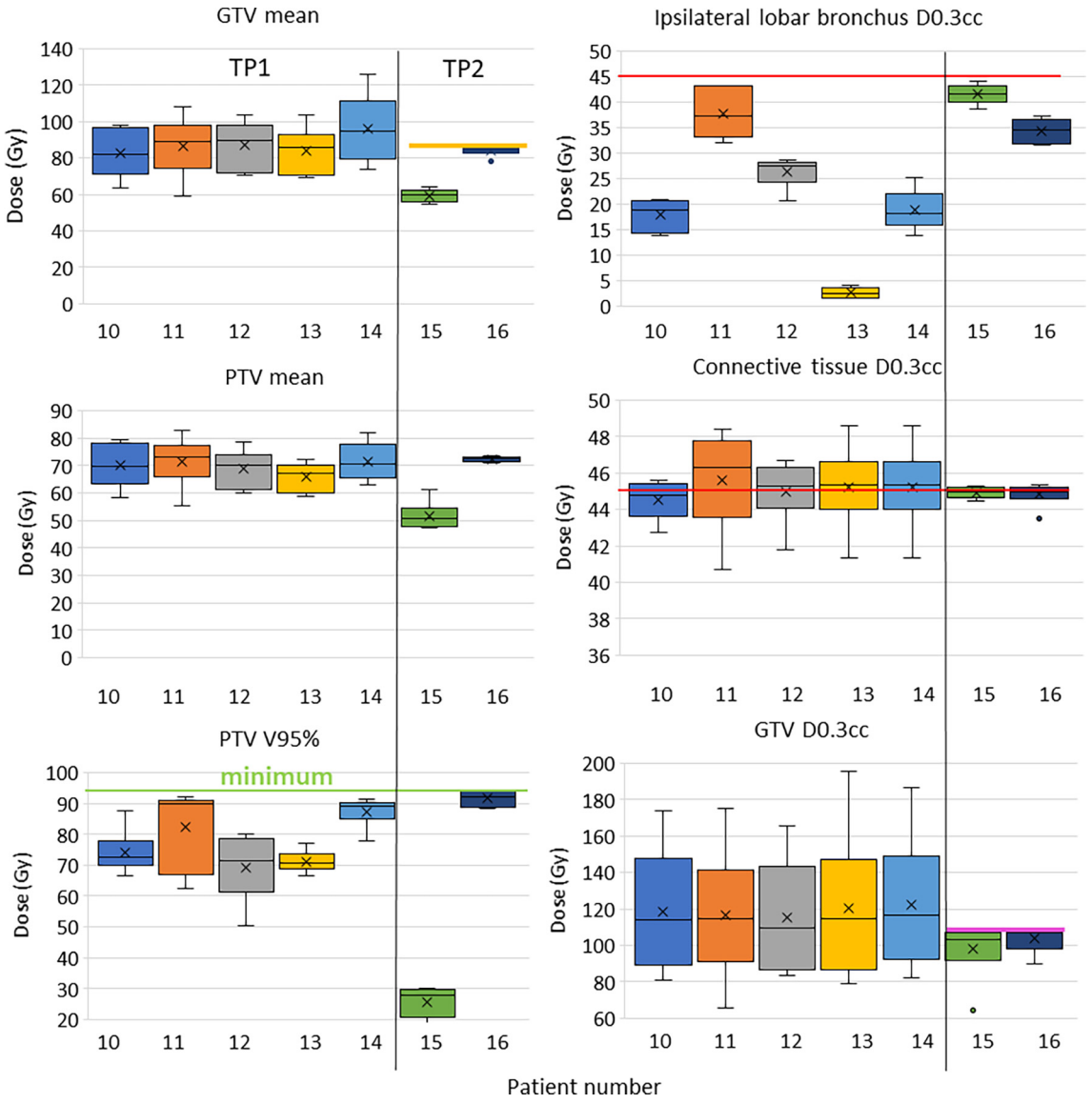

Fig. 5. Boxplot showing dose to target (right) and OARs (left) for patients *1–5* (TP1) and patients *5–6* (TP2). Red line illustrates 45 Gy (upper limit for connective-tissue and lobar bronchus). Green line illustrates minimum requirement for V95% to PTV. Yellow line (85 Gy) and pink line (110 Gy) illustrates upper limit for GTV_{mean} dose and $GTV_{D0.3cc}$ dose. This limit was only required in TP2. Box plot: median (horizontal line), first and third interquartile ranges (box), minimum/maximum (whiskers), and outliers (o). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

setup based on tumour matching is required in addition to treatment adaptation in case of suspicion of OAR over-dosage in order to secure safe dose delivery [42–44].

For the treatment planning study TP1, no constraints were set for GTV_{mean} and $GTV_{D0.3cc}$ dose to investigate how high the tumour dose could be increased. The expectation was that maximum dose would be clinically acceptable due to the strict OAR constraints, meaning that no constraint was necessary for the mean dose to GTV. However, it resulted in maximum GTV_{mean} dose of 125 Gy, far above clinical acceptable limits. A meeting in-between TP1 and TP2, added GTV_{mean} < 85 Gy, leading to reduced variability in dose between the centres. Furthermore, it was agreed that no over-dosage was accepted for OAR leading to compliance of all OAR constraints in TP2. In both TP1 and TP2, PTV coverage was compromised due to higher priority of dose to OAR. This may impact the local control obtained in the trial. Patient *6* had an ultra-central tumour which led to median GTV_{mean} dose of 59.4 Gy and $GTV_{D0.3cc}$ dose of 80 Gy, demonstrating the importance of strict constraints for OAR, whereby dose to the target was reduced. In all centres, the highest doses were located in the part of the GTV farthest from the OARs (Fig. 4). The size of the PTV margin used, will impact the dose to OAR and possibility to cover the PTV with the prescribed dose. In this study, margins of 4–5 mm were used. A PRV margin to the OAR may be used to minimize risk of over-dosage.

This study illustrates the importance of thorough and stringent protocol design and QA-procedures with clear guidelines before initiating multi-centre clinical trials in order to achieve consistent results and deliver a safe and efficient treatment for the patient particularly for heterogeneous dose distributions with steep dose gradients between target and OAR. It is well known that workshops on delineation of OAR result in less variability [27–29]. However, this study demonstrates that the variability may influence the dose escalation of GTV and even more important, the dose to critical OAR with resultant increased risk of toxicity. The current study may serve as a guideline for delineation, atlas

creation, and treatment planning rounds with workshops in-between to complete before initiating of multi-centre clinical trials.

Weaknesses of the present study are firstly, the use of three different patients in the three delineation rounds. Other studies have used only one patient throughout the study [9,10,18,22]. However, the QA process was so fast that the oncologists felt that they would be biased when delineating and discussing the same case. Secondly, only CT images and treatment planning of central tumours were investigated. A similar study will be conducted for MR images. Thirdly, minor variability in the reported near-maximum doses was observed between the treatment planning systems. This results from differences in defining the voxels and reporting dose to each voxel. Hereby, slight over-dosage of some OAR may incorrectly be observed at the centre collecting all the plans even though the plans complied with the constraints when evaluated at the centre responsible for the planning. Additionally, differences up to ~3% in dose calculation algorithms exist [45]. These inconsistencies account for the over-dosage observed for connective-tissue in TP2 (up to 2 Gy) which was not seen in the treatment planning system actually used to generate the plan.

In conclusion, we have setup a thorough pre-trial QA-programme which may serve as a guideline for delineation, treatment planning and workshops to be organized before initiating of multi-centre clinical trials. In heterogeneous dose-escalation trials care should be taken to secure consistent OAR delineation as even small variations may lead to high OAR over-dosage. Extensive QA decreased variability in OAR delineation and dose planning and emphasises the need of thorough QA while conducting a multi-centre trial.

Conflict of interest statement

The authors have no conflicts of interests to report.

Acknowledgement

The study was supported by DCCC Radiotherapy – The Danish National Research Center for Radiotherapy, Danish Cancer Society [grant no. R191-A11526] and Danish Comprehensive Cancer Center.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2022.04.005.

References

[1] Lindberg K, Grozman V, Karlsson K, et al. The HILUS-trial - a prospective Nordic multi-center phase II study of ultra-central lung tumors treated with stereotactic body radiotherapy. J Thorac Oncol 2021;21:10.
[2] Cannon DM, Mehta MP, Adkison JB, Khuntia D, Traynor AM, Tomé WA, et al. Dose-limiting toxicity after hypofractionated dose-escalated radiotherapy in non-small-cell lung cancer. J Clin Oncol 2013;31:4343–8.
[3] Adebahr S, Liu Y, Colette S, Faivre-Finn C, Ahmad S, Ahmed M, et al. OC-0061 EORTC 22113–8113 Lungtech trial on SBRT of central lung tumors. Radiother Oncol 2019;133:S24–5.
[4] Bradley JD, Paulus R, Komaki R, et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): a randomized, two-by-two factorial phase 3 study. Lancet Oncol 2015;16:187–99.
[5] Holloway CL, Robinson D, Murray B, Amanie J, Butts C, Smylie M, et al. Results of a phase I study to dose escalate using intensity modulated radiotherapy guided by combined PET/CT imaging with induction chemotherapy for patients with non-small cell lung cancer. Radiother Oncol 2004;73:285–7.
[6] Rim CH, Kim Y, Kim CY, Yoon WS, Yang DS. Is stereotactic body radiotherapy for ultra-central lung tumor a feasible option? A systemic review and meta-analysis. Int J Radiat Biol 2019;95:329–37.
[7] Timmerman R, McGarry R, Yiannoutsos C, Papiez L, Tudor K, DeLuca J, et al. Excessive toxicity when treating central tumors in a phase II study of stereotactic body radiation therapy for medically inoperable early-stage lung cancer. J Clin Oncol 2006;24:4833–9.
[8] Haseltine JM, Rimner A, Gelblum DY, Modh A, Rosenzweig KE, Jackson A, et al. Fatal complications after stereotactic body radiation therapy for central lung tumors abutting the proximal bronchial tree. Pract Radiat Oncol 2016;6:e27–33.
[9] Lambrecht M, Clementel E, Sonke J-J, Nestle U, Adebahr S, Guckenberger M, et al. Radiotherapy quality assurance of SBRT for patients with centrally located lung tumours within the multicentre phase II EORTC Lungtech trial: Benchmark case results. Radiother Oncol 2019;132:63–9.
[10] Groom N, Wilson E, Lyn E, Faivre-Finn C. Is pre-trial quality assurance necessary? Experiences of the CONVERT Phase III randomized trial for good performance status patients with limited-stage small-cell lung cancer. Br J Radiol 2014;87:20130653.
[11] Fairchild A, Bar-Deroma R, Collette L, Haustermans K, Hurkmans C, Lacombe D, 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.
[12] Thor M, Apte A, Haq R, Iyer A, LoCastro E, Deasy JO. Using auto-segmentation to reduce contouring and dose inconsistency in clinical trials: the simulated impact on RTOG 0617. Int J Radiat Oncol Biol Phys 2021;109:1619–26.
[13] Spoelstra FOB, Senan S, Le Péchoux C, Ishikura S, Casas F, Ball D, et al. Lung Adjuvant Radiotherapy Trial Investigators Group. Variations in target volume definition for postoperative radiotherapy in stage III non-small-cell lung cancer: analysis of an international contouring study. Int J Radiat Oncol Biol Phys 2010;76:1106–13.
[14] Lambrecht M, Melidis C, Sonke J-J, Adebahr S, Boellaard R, Verheij M, et al. Lungtech, a phase II EORTC trial of SBRT for centrally located lung tumours - a clinical physics perspective. Radiat Oncol 2016;11:7.
[15] 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.
[16] Grønberg BH, Killingberg KT, Fløtten Ø, Brustugun OT, Hornslien K, Madebo T, et al. High-dose versus standard-dose twice-daily thoracic radiotherapy for patients with limited stage small-cell lung cancer: an open-label, randomised, phase 2 trial. Lancet Oncol 2021;22:321–31.
[17] Tsang Y, Hoskin P, Spezi E, Landau D, Lester J, Miles E, et al. Assessment of contour variability in target volumes and organs at risk in lung cancer radiotherapy. Tech Innov Patient Support Radiat Oncol 2019;10:8–12.
[18] Petkar I, McQuaid D, Dunlop A, Tyler J, Hall E, Nutting C. Inter-observer variation in delineating the pharyngeal constrictor muscle as organ at risk in radiotherapy for head and neck cancer. Front Oncol 20219;11:644767.
[19] Naismith O, Mayles H, Bidmead M, Clark CH, Gulliford S, Hassan S, et al. Radiotherapy quality assurance for the CHHiP trial: conventional versus hypofractionated high-dose intensity-modulated radiotherapy in prostate cancer. Clin Oncol (R Coll Radiol) 2019;31:611–20.
[20] Song SY, Choi W, Shin SS, Lee S-W, Ahn SD, Kim JH, et al. Fractionated stereotactic body radiation therapy for medically inoperable stage I lung cancer adjacent to central large bronchus. Lung Cancer 2009;66:89–93.
[21] Bral S, Gevaert T, Linthout N, Versmessen H, Collen C, Engels B, et al. Prospective, risk-adapted strategy of stereotactic body radiotherapy for early-stage non-small-cell lung cancer: results of a Phase II trial. Int J Radiat Oncol Biol Phys 2011;80:1343–9.
[22] de Jong EEC, Guckenberger M, Andratschke N, Dieckmann K, Hoogeman MS, Milder M, et al. Variation in current prescription practice of stereotactic body radiotherapy for peripherally located early stage non-small cell lung cancer: Recommendations for prescribing and recording according to the ACROP guideline and ICRU report 91. Radiother Oncol 2020;142:217–23.
[23] Thomas M, Mortensen HR, Hoffmann L, Møller DS, Troost EGC, Muijs CT, et al. Proposal for the delineation of neoadjuvant target volumes in oesophageal cancer. Radiother Oncol 2021;156:102–12.
[24] Aznar MC, Girinsky T, Berthelsen AK, Aleman B, Beijert M, Hutchings M, et al. Interobserver delineation uncertainty in involved-node radiation therapy (INRT) for early-stage Hodgkin lymphoma: on behalf of the Radiotherapy Committee of the EORTC lymphoma group. Acta Oncol 2017;56:608–13.
[25] Cox S, Cleves A, Clementel E, Miles E, Staffurth J, Gwynne S. Impact of deviations in target volume delineation – Time for a new RTQA approach? Radiother Oncol 2019;137:1–8.
[26] Møller DS, Nielsen TB, Brink C, Hoffmann L, Lutz CM, Drøgemüller Lund M, et al. Heterogeneous FDG-guided dose-escalation for locally advanced NSCLC (the NARLAL2 trial): design and early dosimetric results of a randomized, multi-centre phase-III study. Radiother Oncol 2017;124:311–7.
[27] Milo MLH, Offersen BV, Bechmann T, Diederichsen ACP, Hansen CR, Holtved E, et al. Delineation of whole heart and substructures in thoracic radiation therapy: National guidelines and contouring atlas by the Danish Multidisciplinary Cancer Groups. Radiother Oncol 2020;150:121–7.
[28] Feng M, Moran JM, Koelling T, Chughtai A, Chan JL, Freedman L, 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.
[29] Kong F-M, Ritter T, Quint DJ, Senan S, Gaspar LE, Komaki RU, 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.

[30] Mir R, Kelly SM, Xiao Y, Moore A, Clark CH, Clementel E, et al. Organ at risk delineation for radiation therapy clinical trials: Global Harmonization Group consensus guidelines. Radiother Oncol 2020;150:30–9.
[31] Møller DS, Holt MI, Alber M, Tvilum M, Khalil AA, Knap MM, et al. Adaptive radiotherapy for advanced lung cancer ensures target coverage and decreases lung dose. Radiother Oncol 2016;121:32–8.
[32] De Rose F, Franceschini D, Reggiori G, Stravato A, Navarria P, Ascolese AM, et al. Organs at risk in lung SBRT. Phys Med 2017;44:131–8.
[33] Xue J, Kubicek G, Patel A, Goldsmith B, Asbell SO, LaCouture TA. Validity of current stereotactic body radiation therapy dose constraints for aorta and major vessels. Semin Radiat Oncol 2016;26:135–9.
[34] Khalil AA, Knap MM, Møller DS, et al. Local control after stereotactic body radiotherapy of centrally located lung tumours. Acta Oncol 2021;14:1–5.
[35] Hanna GG, McDonald F, Murray L et al. UK Consensus on Normal Tissue Dose Constraints for Stereotactic Radiotherapy: Reply to Ghafoor et al. Clin Oncol (R Coll Radiol) 2018;30:456.
[36] Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26:297–302.
[37] Ruckligde W. Efficient visual recognition using the Hausdorff distance. Lect. Notes Comput. Sci., Springer 1996; pp. 27-42.
[38] Daly ME, Perks JR, Chen AM. Patterns-of-care for thoracic stereotactic body radiotherapy among practicing radiation oncologists in the United States. J Thorac Oncol 2013;8:202–7.
[39] Guckenberger M, AllgÃ¤uer M, Appold S, et al. Safety and efficacy of stereotactic body radiotherapy for stage 1 non-small-cell lung cancer in routine clinical practice: a patterns-of-care and outcome analysis. J Thorac Oncol 2013;8:1050–8.
[40] Corso CD, Park HS, Moreno AC, Kim AW, Yu JB, Husain ZA, et al. Stage I lung SBRT clinical practice patterns. Am J Clin Oncol 2017;40:358–61.
[41] Chen H, Laba JM, Zayed S, Boldt RG, Palma DA, Louie AV. Safety and effectiveness of stereotactic ablative radiotherapy for ultra-central lung lesions: a systematic review. J Thorac Oncol 2019;14:1332–42.
[42] Møller DS, Holt MI, Alber M, et al. Adaptive radiotherapy for advanced lung cancer ensures target coverage and decreases lung dose. Radiother Oncol 2016;121:32–8.
[43] Korreman SS, Juhler-Nøttrup T, Persson GF, et al. The role of image guidance in respiratory gated radiotherapy. Acta Oncol 2008;47:1390–6.
[44] Ottosson W, Baker M, Hedman M, et al. Evaluation of setup accuracy for NSCLC patients; studying the impact of different types of conebeam CT matches based on whole thorax, columna vertebralis, and GTV. Acta Oncol 2010;49:1184–91.
[45] Hoffmann L, Alber M, Söhn M, Elstrøm UV. Validation of the Acuros XB dose calculation algorithm versus Monte Carlo for clinical treatment plans. Med Phys 2018.