Planning benchmark study for SBRT of liver metastases

Moustakis, C; Blanck, O; Chan, Mark; Boda-Heggemann, J; Andratschke, N; Duma, M-N; Albers, D; Bäumer, C; Fehr, R; Körber, S A

Published in:
International Journal of Radiation Oncology, Biology, Physics

DOI:
10.1016/j.ijrobp.2022.01.008

Publication date:
2022

INTERNATIONAL JOURNAL OF
RADIATION ONCOLOGY • BIOLOGY • PHYSICS
www.redjournal.org

PHYSICS CONTRIBUTION

Planning Benchmark Study for Stereotactic Body Radiation Therapy of Liver Metastases: Results of the DEGRO/DGMP Working Group on Stereotactic Radiation Therapy and Radiosurgery

Christos Moustakis, PhD,[1] Oliver Blanck, PhD,[2,3] Mark ka heng Chan, PhD,[2] Judit Boda-Heggemann, MD, PhD,[4] Nicolaus Andratschke, MD,[5] Marciana-Nona Duma, MD,[6,7] Dirk Albers, MSc,[8] Christian Bäumer, PhD,[9] Roman Fehr, PhD,[10] Stefan A. Körber, MD,[11] Daniel Schmidhalter, MSc,[12] Manfred Alraun, PhD,[13] Wolfgang W. Baus, PhD,[14] Eric Beckers, MSc,[15] Mathias Dierl, PhD,[16] Stephan Droege, PhD,[17] Fatemeh Ebrahimi Tazehmahalleh, PhD,[1,14] Jens Fleckenstein, PhD,[4] Matthias Guckenberger, MD,[5] Christian Heinz, PhD,[18] Christoph Henkenberens, MD,[19] Andreas Hennig, MSc,[20] Janett Köhn, PhD,[3,21] Christine Kornhuber, PhD,[22] Thomas Krieger, PhD,[23] Britta Loutfi-Krauss, MSc,[3,21] Manfred Mayr, MSc,[24] Markus Oechsner, PhD,[6] Tina Pfeiler, PhD,[9] Gerhard Pollul, MSc,[25] Jürgen Schöffler, MSc,[26] Heiko Tümmler, PhD,[27] Claudia Ullm, MSc,[28] Mathias Walke, PhD,[29] Rocco Weigel, MSc,[30] Martin Wertman, MSc,[10,31] Rolf Wiehle, PhD,[32] Tilo Wiezorek, PhD,[7] Lotte Wilke, PhD,[5] Ulrich Wolf, PhD,[33] Hans Theodor Eich, MD,[1] and Daniela Schmitt, PhD[11,34]

[1]Department of Radiation Oncology, University Hospital Muenster, Muenster, Germany; [2]Department of Radiation Oncology, University Medical Center Schleswig Holstein, Kiel, Germany; [3]Saphir Radiosurgery Center, Frankfurt and Northern Germany, Kiel, Germany; [4]Department of Radiation Oncology, University Medicine Mannheim, Medical Faculty Mannheim, University Heidelberg, Mannheim, Germany; [5]Department of Radiation Oncology, University Hospital Zurich, University of Zurich, Zurich, Switzerland; [6]Department of Radiation Oncology, University Hospital Rechts der Isar, Technical University, Munich, Germany; [7]Department of Radiation Oncology, University Hospital Jena, Jena, Germany; [8]Department of Radiotherapy and Radiation Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [9]West German Proton Therapy Centre, Essen, Germany; [10]Department of Radiation Oncology, University Medicine Rostock, Rostock, Germany; [11]Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany; [12]Division of Medical Radiation Physics and Department of Radiation Oncology, Inselspital, Bern University Hospital, and University of Bern, Bern, Switzerland; [13]Department of Radiation Oncology, Schwarzwald-Baar Hospital, Villingen-Schwenningen, Germany; [14]Department of Radiation Oncology, University Hospital Cologne, Cologne, Germany; [15]Gamma Knife Center Krefeld, Krefeld, Germany; [16]Department of Radiation Oncology, Hospital Bayreuth, Bayreuth, Germany; [17]Department of Radiation Oncology, Lung Clinic Hemer, Hemer, Germany; [18]Department of Radiation Oncology, University Hospital Munich, LMU Munich, Munich, Germany; [19]Department of Radiation Oncology, Medical University Hannover, Hannover, Germany; [20]Private Practice for Radiation Oncology, Distler Bautzen, Germany; [21]Department of Radiation Oncology, University Hospital Frankfurt, Frankfurt, Germany; [22]Department of Radiation Oncology, University Hospital Halle, Halle (Saale), Germany; [23]Department of Radiation Oncology, University Hospital Wuerzburg, Wuerzburg, Germany; [24]Private Practice for Radiation Oncology South, Kaufbeuren, Germany; [25]Department of Radiation Oncology, University Hospital Mainz, Mainz, Germany; [26]Private Practice for Radiation Oncology, Boeblingen, Germany; [27]Private Practice for Radiation Oncology, Dresden, Germany; [28]Radiologie Muenchen, Private Practice for Radiation Oncology, Munich, Germany; [29]Department of Radiation Oncology, University Hospital

Corresponding author: Christos Moustakis, PhD; E-mail: christos.moustakis@ukmuenster.de

Disclosures: none.

Data sharing statement: Research data are stored in an institutional repository and will be shared upon request to the corresponding author.

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijrobp.2022.01.008.

Acknowledgments—We thank the independent experts Dr Marco Esposito (Medical Physics Unit, AUSL Toscana Centro, Florence, Italy) and Dr Victor Hernandez (Department of Medical Physics, Hospital Universitari Sant Joan de Reus, Tarragona, Spain) for critically reading the work and discussing and validating the RATING score.

Int J Radiation Oncol Biol Phys, Vol. 113, No. 1, pp. 214–227, 2022
0360-3016/$ - see front matter
https://doi.org/10.1016/j.ijrobp.2022.01.008

Magdeburg, Magdeburg, Germany; [30]*Department of Radiation Oncology, University Hospital Innsbruck, Innsbruck, Austria;* [31]*Department of Radiation Oncology, Ernst von Bergmann Hospital, Potsdam, Germany;* [32]*Department of Radiation Oncology, University Hospital Freiburg, Freiburg, Germany;* [33]*Department of Radiation Oncology, University Hospital Leipzig, Leipzig, Germany; and* [34]*Department of Radiation Oncology, University Medical Center Goettingen, Goettingen, Germany*

Received Jun 26, 2021; Revised Dec 19, 2021; Accepted for publication Jan 7, 2022

Purpose: Our purpose was to investigate whether liver stereotactic body radiation therapy treatment planning can be harmonized across different treatment planning systems, delivery techniques, and institutions by using a specific prescription method and to minimize the knowledge gap concerning intersystem and interuser differences. We provide best practice guidelines for all used techniques.

Methods and Materials: A multiparametric specification of target dose (gross target volume $[GTV]_{D50\%}$, $GTV_{D0.1cc}$, $GTV_{V90\%}$, planning target volume $[PTV]_{V70\%}$) with a prescription dose of $GTV_{D50\%} = 3 \times 20$ Gy and organ-at-risk (OAR) limits were distributed with computed tomography and structure sets from 3 patients with liver metastases. Thirty-five institutions provided 132 treatment plans using different irradiation techniques. These plans were first analyzed for target and OAR doses. Four different renormalization methods were performed (PTV_{Dmin}, $PTV_{D98\%}$, $PTV_{D2\%}$, PTV_{Dmax}). The resulting 660 treatments plans were evaluated regarding target doses to study the effect of dose renormalization to different prescription methods. A relative scoring system was used for comparisons.

Results: $GTV_{D50\%}$ prescription can be performed in all systems. Treatment plan harmonization was overall successful, with standard deviations for D_{max}, $PTV_{D98\%}$, $GTV_{D98\%}$, and PTV_{Dmean} of 1.6, 3.3, 1.9, and 1.5 Gy, respectively. Primary analysis showed 55 major deviations from clinical goals in 132 plans, whereas in only <20% of deviations GTV/PTV dose was traded for meeting OAR limits. $GTV_{D50\%}$ prescription produced the smallest deviation from target planning objectives and between techniques, followed by the PTV_{Dmax}, $PTV_{D98\%}$, $PTV_{D2\%}$, and PTV_{Dmin} prescription. Deviations were significant for all combinations but for the PTV_{Dmax} prescription compared with $GTV_{D50\%}$ and $PTV_{D98\%}$. Based on the various dose prescription methods, all systems significantly differed from each other, whereas $GTV_{D50\%}$ and $PTV_{D98\%}$ prescription showed the least difference between the systems.

Conclusions: This study showed the feasibility of harmonizing liver stereotactic body radiation therapy treatment plans across different treatment planning systems and delivery techniques when a sufficient set of clinical goals is given.

Introduction

Stereotactic body radiation therapy (SBRT) is defined as highly precise hypofractionated radiation therapy[1] and is well-established for many extracranial tumor manifestations like primary and secondary lung and liver malignancies[2-7] and spinal metastases,[8] among others. Owing to knowledge and technology gains in recent years,[9] the treatment of liver metastases with SBRT found its way into clinical routine, especially for those originating from colorectal, breast, or lung cancer.[7,9,10] However, currently, SBRT for liver metastases is not often performed and treatment planning as well as techniques may vary widely.[7,11] To harmonize liver SBRT practice, also in the context of multicenter multiplatform clinical trial preparations, several investigations on minimally needed biological effective dose (BED)[7,12-14] and technology comparison on phantom and in vivo dose delivery accuracy[13] have been published. The aim of this study was to fill the knowledge gap concerning the intersystem and interuser differences for treatment techniques and treatment planning systems (TPSs).

For SBRT in general, machine/technology and planner experience variability have been investigated in the past.[14-17] An overview of planning benchmark studies has also recently been published.[18] The results from those studies show that treatment plan quality only slightly depends on machine/technology while the highest variability seems to originate from planner experience.[15,19] One way to overcome those differences can be addressed through benchmark studies and crowd knowledge-based learning.[16,19-21] Another way is the strict specification of planning goals and dose prescription, like a recent study on lung SBRT demonstrated.[22] In the lung, tissue heterogeneities are mostly responsible for discrepancies in tumor dose distribution between different patients using common circumferential planning target volume (PTV) prescription methods.[23,24] However, there is a lack of benchmark planning studies for larger tumors in more homogeneous tissue closer to critical organs at risk (OAR), like liver metastases.

In this joint planning benchmark study from the German Society for Radiation Oncology working group for radiosurgery and stereotactic radiation therapy and the German Society for Medical Physics working group for physics and technology in stereotactic radiation therapy (Deutsche Gesellschaft für Radioonkologie/ Deutsche Gesellschaft für Medizinische Physik, DEGRO/DGMP) we assessed the harmonization of treatment plan quality for liver metastases SBRT in a multi-institutional multiplatform context on an international level. Besides previous knowledge gains from the harmonization of SBRT treatment planning for lung tumors through gross tumor volume (GTV)/clinical target volume (CTV)/internal tumor/target volume mean dose

prescription,[22] we implemented stricter planning goals for GTV and PTV dose coverage requirements based on recent findings on various dose-response relationships for liver metastases.[10,25,26] The aims of the present study were (1) harmonization of liver SBRT treatment plan quality for large lesions, (2) investigation on the possibility of using GTV median dose prescription in the liver, and (3) validation that the as low as reasonably achievable (ALARA) principle is better implemented after our initial planning benchmarks.[14,15]

Methods and Materials

Case selection and patient characteristics

After approval from the leading ethics committee of the medical faculty of the University of Kiel (reference number D 457/18), the SBRT liver databases of the lead institutions were screened for 3 suitable cases for this treatment planning benchmark study. The number of cases originated from previous studies and was found to be a balance between statistical power in analysis and workload of the participants.[14,15] For the search we defined the primary characteristics of the cases in accordance with prior pattern-of-care studies[7,10] to be (1) each a different primary histology commonly treated with SBRT with follow-up ≥ 1 year, to be (2) of varying tumor volume, but larger than a GTV of 25 cm^3 and to include (3) at least 1 case with (a) central location and/or (b) smaller whole liver volume < 1200 cm^3, and/or (c) close location to gastrointestinal organs (approximately 1.5 cm based on the study by Sterzing et al[11]). From the databases 12 cases were preselected based on the aforementioned characteristics from which, finally, 3 cases were selected for this benchmark based on study committee consensus decision.

Case 1 (criterion a) was a 52-year-old patient with liver metastases from rectal carcinoma. After several systemic and local treatments 1 oligo-persistent liver metastasis remained. It was located in segment IV close to the portal vein and hence unresectable and was treated with SBRT (GTV = 52 cm^3, whole liver = 2350 cm^3).

Case 2 (criterion b) was an 83-year-old patient with breast carcinoma. Liver metastases were first discovered 3 years after initial diagnosis and several asynchronous oligo-metastases were resected or ablated over the course of 12 years. After recurrence of yet another 2 adjacent oligo-metastases in segment IV/VIII the treatment was continued with SBRT due to the small remaining liver volume (GTV for both metastases = 69 cm^3, 1 merged CTV, whole liver = 1134 cm^3).

Case 3 (criterion c) was a 63-year-old patient with early stage non-small cell lung cancer that was initially treated with SBRT due to multiple comorbidities. Subsequently, an inoperable oligo-metastasis in liver segment V with close proximity to the intestine (1.5 cm) was treated with SBRT (GTV = 25 cm^3, whole liver = 1600 cm^3).

Treatment planning image data sets and contours were initially obtained from the treating institution after full case data anonymization. The planning computed tomography (CT) was acquired head first supine with vacuum bag at end expiration breath hold with 1.0-mm in-plane resolution and 1.5-mm slice thickness, and the planning magnetic resonance imaging was performed with the same resolution using a standard liver SBRT protocol as described previously.[26] For the purpose of this planning benchmark study the OAR contours were harmonized for all 3 cases and adapted/added where necessary according to DEGRO and international guidelines,[11,16] based on study committee consensus decision. The original GTV and CTV contours, for these 3 cases CTV = GTV + 5 mm isotropic in-liver expansion,[27] were not modified and an isotropic margin of 3 mm was added to the CTV to generate the PTV assuming an active motion compensation technique during treatment.[13,28] The final PTVs for this study were 126 cm^3 (case 1), 152 cm^3 (case 2), and 73 cm^3 (case 3) as illustrated in Figure 1.

Treatment techniques and treatment planning systems

The anonymized treatment planning CT images and the respective radiation therapy structure sets in Digital Imaging and Communications in Medicine standard format were distributed to all institutions participating in this trial. Treatment technique selection and treatment planning were performed with each institution's equipment using institutional-specific methods/techniques and society guidelines.[11,15,28] The use of a type B[23] dose calculation algorithm was recommended. All submitted treatment plans had to be clinically acceptable concerning OAR limits, judged by the residing radiation oncologist responsible for stereotactic radiation therapy. Plans were supposed to meet predefined dose prescription and clinical goals. If a goal was not met, a deviation was documented. To characterize the extent of the deviation, we predefined a set of thresholds to differentiate between "minor" and "major" deviations. In the following, the goals for "no deviation" and the thresholds for "minor deviation" (in brackets) are presented. "Major deviations" meant that thresholds for "minor deviations" were not met.

1. Based on previous studies,[7,15,17,22,25,29] the prescription dose for all 3 cases was defined as median GTV dose ($GTV_{D50\%}$) = 3 × 20 Gy at the 100% isodose line ($BED_{\alpha/\beta=10Gy}$ = 180 Gy_{10}). Further target planning objectives were (a) GTV near maximum (ie, $GTV_{D0.1cc}$) ≤ 107% = 64.2 Gy (≤110% = 66 Gy),[23,24] (b) GTV coverage at 54 Gy (ie, $GTV_{V90\%}$)[26] ≥ 98% (≥95%), and (c) PTV coverage at 42 Gy (ie, $PTV_{V70\%}$) ≥ 98% (≥95%).[7,25]
2. Based on commonly available OAR limits for liver SBRT in 3 fractions,[11,27,30] the major limitations for this study

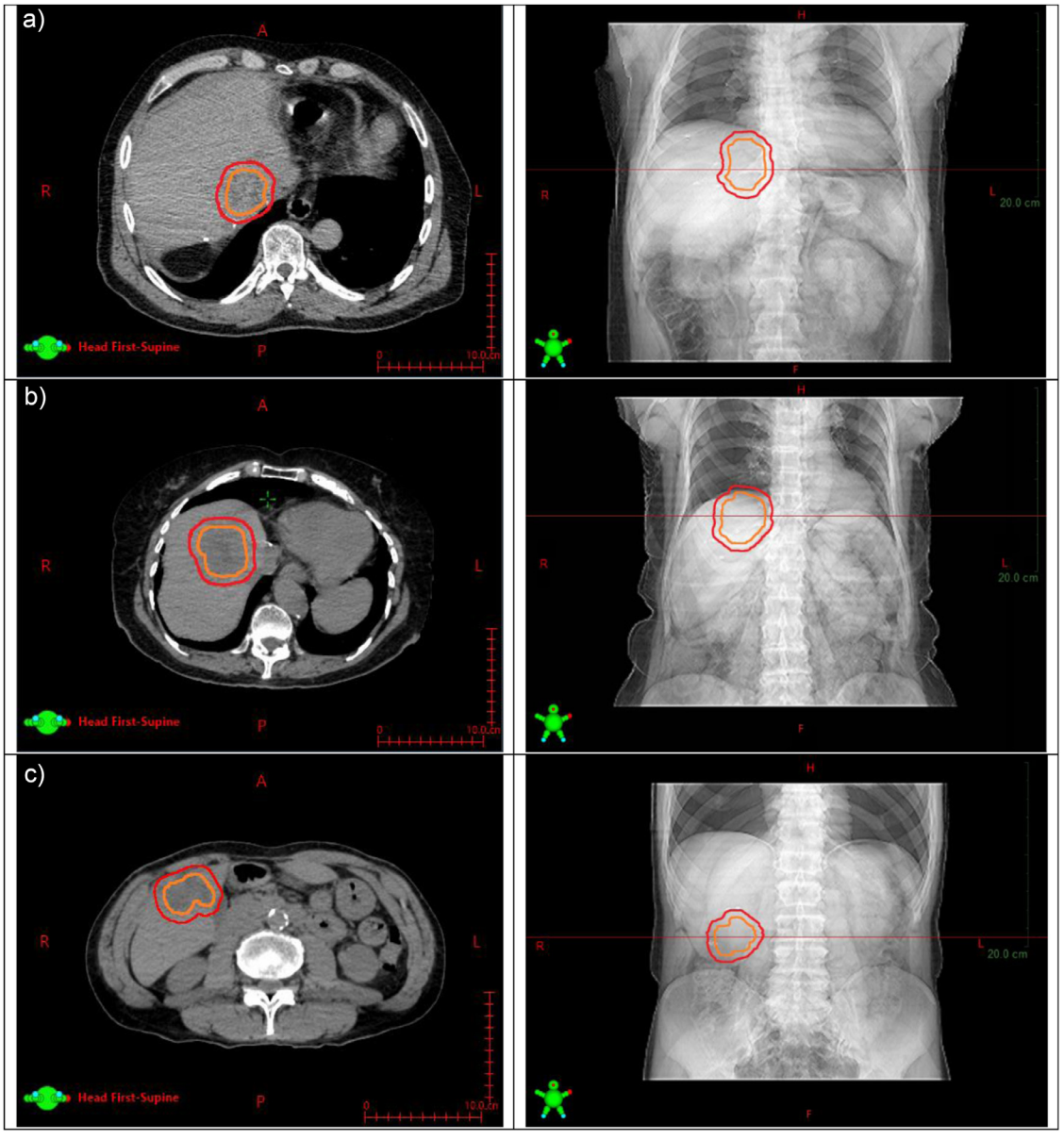

Fig. 1. Axial (left) and topogram (right) view of case 1 with planning target volume (PTV) = 126 cm^3 (a), case 2 with PTV = 152 cm^3 (b) and case 3 with PTV = 73 cm^3 (c). The red line illustrates the PTV and the orange line shows the gross tumor volume.

were (a) healthy liver minus $Liver_{V15Gy}$ ($Liver_{V17Gy}$) ≥ 700 cm^3, (b) gastrointestinal organs D_{max} (D_{1cc}) ≤ 24 Gy, (c) heart D_{max} (D_{1cc}) ≤ 30 Gy, and (d) esophagus/stomach D_{max} (D_{1cc}) ≤ 21 Gy. Further limitations are presented in Table E1 and all OARs were to be handled based on the ALARA concept.

Dosimetric evaluation

The resulting dose distributions were collected from all participating institutions in Digital Imaging and Communications in Medicine format and imported into a common TPS (Eclipse, version 15.6; Varian Medical Systems, Palo Alto, CA) for combined evaluation. A detailed questionnaire for each case (see Supplementary Materials) was also completed by the participants. The dosimetric evaluation was based on the International Commission on Radiation Units and Measurements (ICRU) report 91 on prescribing, recording, and reporting of stereotactic treatments with small photon beams.[23]

Primary dosimetric evaluation

Because each TPS will have minor differences in contour interpretation, not all submitted dose distributions may perfectly fulfill the dose prescription requirement of 60 Gy to

the 100% isodose line after the import in Eclipse. Hence, we first performed a minimal dose correction in Eclipse by renormalization of the $GTV_{D50\%}$ to 60 Gy for all treatment plans.

Renormalization to different prescription methods

To further assess different dose prescription methods, we also renormalized the dose to (a) $PTV_{D98\%}$ (near minimum) = 42 Gy, (b) PTV_{Dmin} (absolute minimum) = 42 Gy, (c) $PTV_{D2\%}$ (near maximum) = 64.2 Gy, and (d) PTV_{Dmax} (absolute maximum) = 64.2 Gy, resulting in overall 5 dose distributions for every submitted treatment plan. These are all common methods, although prescribing to a single voxel (minimum and maximum prescription) is in principle less robust than a volume-based prescription.

Target volume doses and indices evaluation

For each prescription method we evaluated GTV and PTV D_{min}, $D_{98\%}$, $D_{50\%}$, $D_{2\%}$, and D_{max} and GTV_{V54Gy} and PTV_{V42Gy}. Additionally, commonly used plan quality indices were evaluated based on previous studies and ICRU report 91 recommendations[14,15,23,31,32,33]:

1a. Homogeneity index PTV (HI_{PTV}) = ($PTV_{D2\%}$ - $PTV_{D98\%}$)/$PTV_{D50\%}$
1b. Homogeneity index GTV (HI_{GTV}) = ($GTV_{D2\%}$ - $GTV_{D98\%}$)/$GTV_{D50\%}$
2a. ICRU 91 conformity index (CI) = ($V_{PTV} \times V_{42Gy}$)/PTV_{V42Gy}^2
2b. External conformity index (CΔ) = (V_{42Gy} - PTV_{V42Gy})/V_{PTV}
3a. Gradient index PTV (GI_{PTV}) = V_{21Gy}/V_{42Gy}
3b. Gradient index GTV (GI_{GTV}) = V_{30Gy}/V_{60Gy}

Table 1 Key components from the dosimetric evaluation for $GTV_{D50\%}$ dose renormalization at 3 × 20 Gy

Mean ± standard deviation	All plans	3D-CRT	SF-IMRT	IMAT	RRS	HRT	PT
Plan D_{max} (Gy)	63.9 ± 1.6	63.2 ± 1.4	64.5 ± 1.2	64.0 ± 1.6	64.7 ± 1.3	62.5 ± 1.3	63.5 ± 2.0
GTV $D_{2\%}$ (Gy)	62.5 ± 1.1	62.5 ± 1.0	63.0 ± 0.9	62.5 ± 1.1	63.3 ± 0.7	61.6 ± 1.0	61.9 ± 1.0
GTV $D_{98\%}$ (Gy)	55.8 ± 1.9	55.8 ± 1.3	55.6 ± 1.4	55.9 ± 1.7	53.7 ± 1.9	56.9 ± 1.9	57.1 ± 1.8
PTV $D_{2\%}$ (Gy)	62.2 ± 1.0	62.2 ± 0.9	62.5 ± 0.8	62.3 ± 0.9	62.9 ± 0.7	61.4 ± 0.8	61.7 ± 1.1
PTV $D_{98\%}$ (Gy)	42.9 ± 3.3	42.0 ± 1.6	42.5 ± 1.7	43.4 ± 3.4	41.2 ± 3.6	42.6 ± 4.2	44.7 ± 4.8
PTV D_{mean} (Gy)	54.4 ± 1.5	55.5 ± 0.8	53.9 ± 1.9	54.4 ± 1.6	53.7 ± 0.9	55.2 ± 1.9	56.3 ± 1.0
HI_{PTV}	0.35 ± 0.07	0.36 ± 0.04	0.37 ± 0.05	0.34 ± 0.07	0.40 ± 0.07	0.33 ± 0.08	0.29 ± 0.1
HI_{GTV}	0.11 ± 0.04	0.10 ± 0.04	0.12 ± 0.03	0.11 ± 0.04	0.16 ± 0.04	0.08 ± 0.04	0.08 ± 0.03
CI	1.19 ± 0.18	1.22 ± 0.08	1.19 ± 0.18	1.19 ± 0.21	1.13 ± 0.06	1.30 ± 0.24	1.21 ± 0.08
CΔ	0.15 ± 0.19	0.17 ± 0.05	0.11 ± 0.08	0.15 ± 0.23	0.08 ± 0.05	0.24 ± 0.25	0.18 ± 0.10
GI_{PTV} case 1	3.86 ± 0.90	3.91 ± 0.20	4.39 ± 0.94	3.94 ± 0.39	3.19 ± 0.30	3.75 ± 0.20	2.68 ± 1.25
GI_{PTV} case 2	3.54 ± 0.59	3.78 ± 0.54	4.01 ± 0.58	3.71 ± 0.28	2.79 ± 0.19	3.54 ± 0.21	2.30 ± 0.82
GI_{PTV} case 3	3.20 ± 0.45	3.14 ± 0.33	3.43 ± 0.69	3.40 ± 0.25	2.80 ± 0.11	3.25 ± 0.32	2.19 ± 0.21
GI_{PTV} all cases	3.54 ± 0.72	3.88 ± 1.23	3.91 ± 0.75	3.68 ± 0.38	2.95 ± 0.28	3.51 ± 0.32	2.39 ± 0.79
GI_{GTV} case 1	4.31 ± 0.78	3.75 ± 0.14	4.83 ± 0.43	4.45 ± 0.61	4.48 ± 0.47	4.07 ± 0.84	2.73 ± 0.16
GI_{GTV} case 2	8.89 ± 4.63	7.27 ± 3.55	13.06 ± 3.68	8.96 ± 4.59	11.53 ± 5.87	5.99 ± 1.76	4.18 ± 0.50
GI_{GTV} case 1	11.40 ± 1.94	11.47 ± 0.55	12.36 ± 1.29	11.89 ± 2.13	10.58 ± 0.91	11.30 ± 0.85	7.98 ± 1.44
GI_{GTV} all cases	8.18 ± 4.15	7.64 ± 3.66	10.77 ± 4.28	9.21 ± 7.80	8.59 ± 4.52	7.12 ± 3.38	4.96 ± 2.47
Case 1 esophagus D_{max} (Gy)	18.8 ± 3.5	21.8 ± 1.1	19.6 ± 1.1	18.9 ± 3.1	16.9 ± 3.3	20.3 ± 3.3	14.4 ± 6.2
Case 2 heart D_{max} (Gy)	35.3 ± 5.3	33.9 ± 2.7	39.5 ± 5.0	36.2 ± 4.5	30.0 ± 4.1	38.3 ± 6.0	31.6 ± 8.2
Case 3 duodenum D_{max} (Gy)	24.0 ± 4.3	28.1 ± 3.9	24.0 ± 2.1	24.3 ± 3.8	20.9 ± 3.3	25.5 ± 2.6	18.0 ± 6.4
Case 1 liver V_{15Gy} (cm^3)	694.5 ± 149.5	800.3 ± 92.2	669.6 ± 211.8	714.3 ± 115.4	583.7 ± 72.2	809.0 ± 180.8	460.8 ± 178.1
Case 2 liver V_{15Gy} (cm^3)	489.5 ± 74.7	535.4 ± 59.2	566.8 ± 52.6	482.5 ± 66.1	476.0 ± 16.4	498.4 ± 71.5	346.1 ± 43.5
Case 3 liver V_{15Gy} (cm^3)	244.3 ± 28.0	249.9 ± 7.6	244.8 ±17.9	258.0 ± 18.6	232.4 ± 15.0	231.5 ± 8.3	171.5 ± 30.8
Number of plans	132	16	11	68	16	12	9

Abbreviations: 3D-CRT = 3-dimensional conformal radiation therapy; CΔ = external conformity index; CI = ICRU 91 conformity index; GI = gradient index; GTV = gross target volume; HI = homogeneity index; HRT = helical radiation therapy; IMAT = intensity modulated arc therapy; PT = proton therapy; PTV = planning target volume; RRS = robotic radiation surgery; SF-IMRT = static field intensity modulated radiation therapy.

where PTV_{VxGy} represents the PTV part and V_{XGy} denotes the total tissue volume that received at least X Gy. For delivery efficiency comparison, the monitor units and the estimated irradiation times were collected via the aforementioned questionnaire (Supplementary Materials).

OAR evaluation

Because the OAR limits will likely be violated when renormalizing based on different dose prescription methods than originally planned on, we only evaluated the OAR dosimetry based on the renormalized treatment plans of 60 Gy to the $GTV_{D50\%}$ in Eclipse. Based again on the suggestions of the ICRU report 91, we extracted D_{mean}, $D_{2\%}$, and D_{max} and a set of appropriate volume doses (V_{XGy}) and (D_{Ycc}) derived from guidelines and international protocols, as described previously and partly provided in Table 1.

Plan quality ranking

To compare the treatment plans qualitatively, independent of specific case properties and combined for various technologies, we applied the previously published and well-established relative plan ranking method.[14,15,18,32,33] In brief, with this ranking method the analyzed data are split into 4 categories (1 = excellent, 2 = above average, 3 = below average, 4 = poor) using the normal distribution (bell curve) of the best and the worst results for that data (first order ranking). For larger amounts of parameters in the data, each parameter was ranked separately. These separate ranks (1-4) were then summed for each plan and this sum was ranked again using the same normal distribution ranking method (second order ranking). The following data were ranked using these methods:

1. For the analysis of overall plan quality for the primary dose prescription ($GTV_{D50\%}$ = 60 Gy) we included all relevant and nonredundant plan parameters into the second order ranking method (ie, PTV_{V42Gy}, GTV_{V54Gy}, GTV_{D1cc}, CI, CΔ, GI_{PTV}, intestine/duodenum D_{1cc}/V_{15Gy}, esophagus/stomach/spinal-cord/aorta/inferior-vena-cava/skin D_{1cc}, and heart D_{1cc}/V_{24Gy}). We also created subscores for only the OAR (calculated like the overall plan quality rank, but without the target parameters) for all 3 cases combined and for only the OAR close to the PTV (the same as before without the nonclose OAR) for each case separately then combined into final subscore over the 3 cases. Here we also analyzed (a) individual planner and (b) technology.

2. For the analysis of deviations from GTV/PTV objectives and OAR limits for the primary dose prescription ($GTV_{D50\%}$ = 60 Gy) we gave penalty points for each minor (1 penalty point) or major (2 penalty points) deviation. The penalty points were summed for each plan and the first order ranking method was used to rank the plans based on the penalty sum (deviation score). Again, we analyzed (a) individual planner and (b) technology.

Because each of the 4 presented plan scores (ie, overall plan score, all OAR score, PTV close OAR score, and deviation score) may not fully reflect a specific quality of a treatment plan, we decided to also evaluate the average over all 4 scores and defined this average as final benchmark score for each participant.

Dose prescription evaluation

For the analysis of the renormalization to different prescription methods, we calculated the absolute deviation from the ideal GTV/PTV dose objectives as described in the previous section. For each plan, the sum of these absolute deviations in Gy was calculated and used to score each plan with each of the 5 prescription methods and to compare (a) prescription methods and (b) technology over all 3 cases.

Statistical analysis

Statistical differences and variances of the dose metrics and ranking scores between different planning techniques were analyzed by Kruskall-Wallis test and Levene's test, respectively. When statistical differences were found with $P < .05$, further post hoc pair-wise comparisons between planning techniques were performed applying Bonferroni's correction. All statistical analyses were conducted using R software version 3.5.1 (The R Foundation, Vienna, Austria) and were considered significant if the P value was <.05.

Results

Thirty-five institutions with experience in liver-SBRT participated in this study. Examples for dose distributions for case 1 can be found in Figure 2.

Treatment techniques and TPSs

For all 3 cases combined, 132 treatment plans (44 for case 1, 45 for case 2, and 43 for case 3) were generated using multiple techniques within various TPSs. Notably, some institutions provided more than 1 plan per case. For case 1, 1 institution provided 3 plans, 7 institutions provided 2 plans. For case 2, 1 institution provided 5 plans, 6 institutions provided 2 plans. For case 3, 1 institution provided 3 plans, 6 institutions provided 2 plans. Sixty-eight plans (52%) were generated with intensity modulated arc therapy (IMAT), 11 plans (8%) with static field intensity modulated radiation therapy (SF-IMRT), 16 (12%) with 3-dimensional conformal radiation therapy, 16 plans (12%) with robotic radiosurgery (RRS), 12 plans (9%) with helical radiation therapy (HRT), and 9 plans (7%) with proton therapy (PT) techniques. Of all plans, 53%, 23%, and 14% used a photon spectrum with a maximum energy of 6, 10, and 15 MeV; of those, 27, 6, and 18 plans used a flattening filter, and 48, 24, and 0 plans were flattening filter free, respectively.

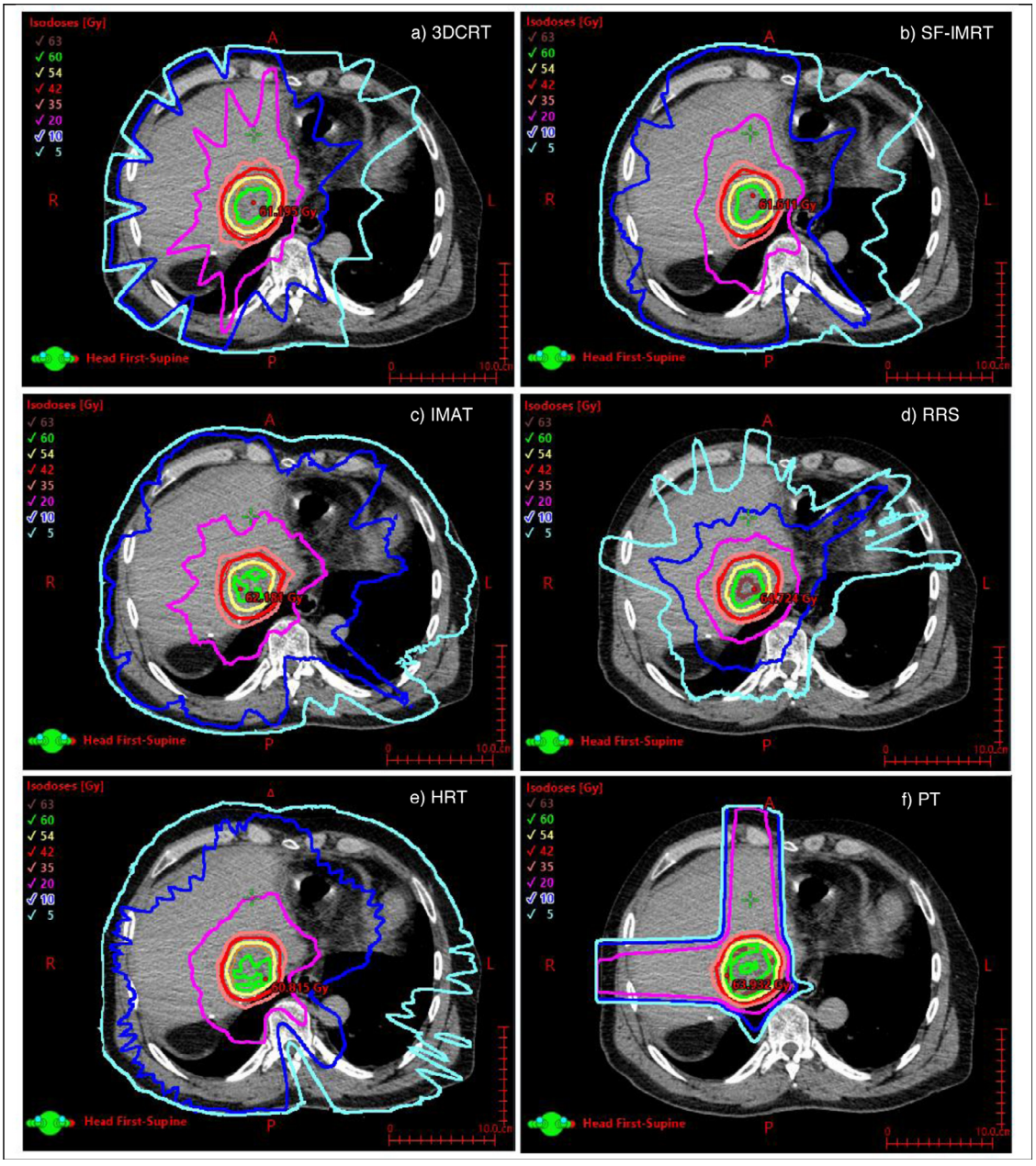

Fig. 2. Dose distributions for different treatment techniques for case 1 with three-dimensional conformal radiation therapy (3D-CRT) (a), static field intensity modulated radiation therapy (SF-IMRT) (b), intensity modulated arc therapy (IMAT) (c), robotic radiosurgery (RRS) (d), helical radiation therapy (HRT) (e), and proton therapy (PT) (f).

Dose calculation algorithms were mostly collapsed cone (29%), Monte Carlo (25%), and anisotropic analytical algorithms (25%; Varian Medical Systems), while superposition/convolution was used in 9% of the plans and Acuros (Varian Medical Systems) in 5%. Pencil beam algorithm was used for the proton plans (7%). The dose calculation grid was smaller than or equal to 2 mm for all plans and most used the planning CT resolution ($1 \times 1 \times 1.5$ mm^3). All plans were clinically accepted in each institution in terms of dose to critical organs before submission. Import of the submitted dose files was not straightforward in all cases, and some files had to be modified to be imported into the common planning system used for this study (Eclipse).

Primary dosimetric evaluation

Initial dose renormalization

The differences between the submitted plan doses from the questionnaire and the Eclipse doses for the mandatory prescription of 3×20 Gy to the GTV $D_{50\%}$ were small. For cases 1, 2, and 3 we found differences of 0.1 ± 0.2 Gy (range,

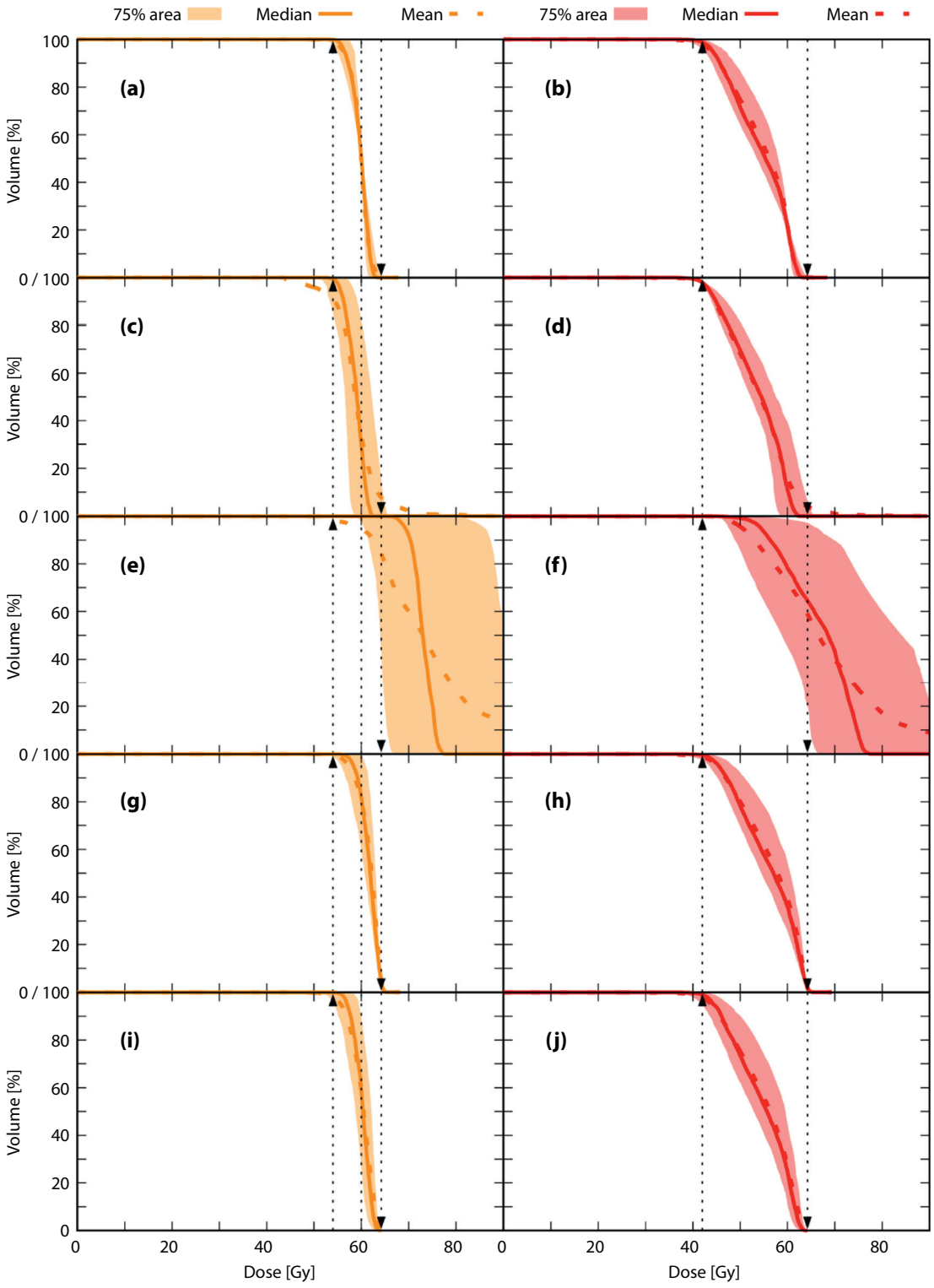

Fig. 3. Dose-volume histograms for different prescription methods for all plans of all cases. Clinical goals are marked with arrows. For gross target volume (GTV; orange) and planning target volume (PTV; red) the median and mean curve and the area of the central 75% of data is shown. The subplots refer to the following prescriptions: reference prescription of 60 Gy to $GTV_{D50\%}$ (a, b), prescription of 42 Gy to $PTV_{D98\%}$ (c, d), 42 Gy to PTV_{Dmin} (e, f), 64.2 Gy to $PTV_{D2\%}$ (g, h), and 64.2 Gy to PTV_{Dmax} (i, j).

−0.6% to 1.3%), 0.0 ± 0.2 Gy (range, −1.4% to 1.0%), and 0.0 ± 0.1 Gy (range, −0.5% to 1.1%), respectively. This may point to very similar volume and dose interpretation of the varying TPSs. Hence, the dose renormalization in Eclipse of 3 × 20 Gy to the GTV $D_{50\%}$ for all cases was regarded as dosimetrically negligible.

Target volume doses and indices

Despite strict requirements in prescription and several other parameters, the dose and homogeneity inside the GTV and PTV varied between individual planners, even between similar technologies. However, the interplanner intersystem treatment plan harmonization in this study was successful overall (see dose-volume histograms for all cases in Fig. 3), with mean ± standard deviation plan D_{max}, $PTV_{D98\%}$, $GTV_{D98\%}$, and PTV_{mean} of 63.9 ± 1.6 Gy (goal ≤ 64.2 Gy), 42.9 ± 3.3 Gy (goal ≥ 42.0 Gy), 55.8 ± 1.9 Gy (goal ≥ 54.0 Gy), and 54.4 ± 1.5 Gy (no goal given), respectively. The resulting HI_{PTV} and HI_{GTV} were 0.35 ± 0.07 and 0.11 ± 0.04, and the CI and CΔ were 1.19 ± 0.18 and 0.15 ± 0.19,

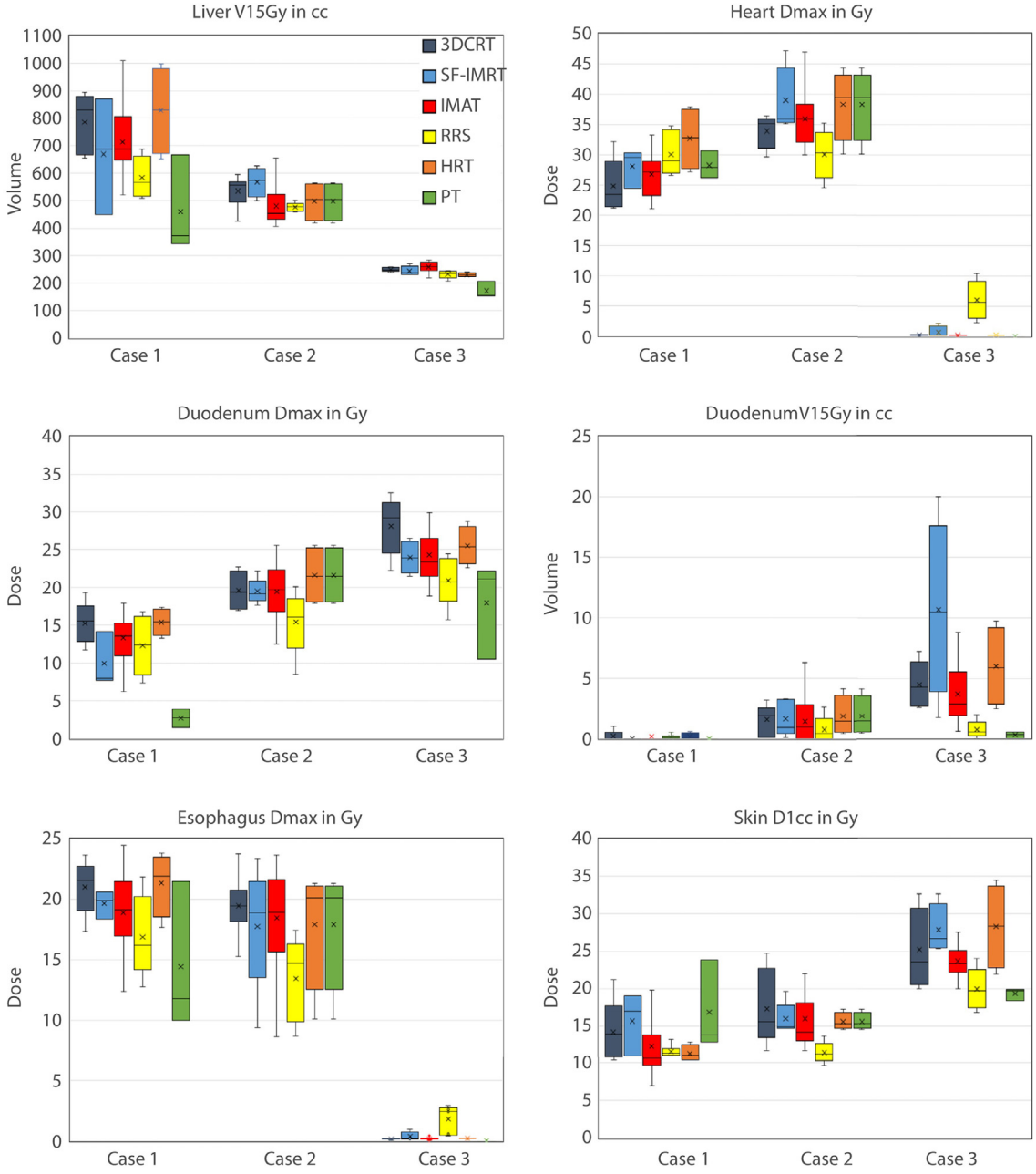

Fig. 4. Organs-at-risk dosimetry as boxplots for various organs at risk for the 3 benchmark cases and for different techniques: 3-dimensional conformal radiation therapy, static field intensity modulated radiation therapy, intensity modulated arc therapy, robotic radiosurgery, helical radiation therapy, and proton therapy.

respectively. On the other hand, the dose gradients outside the PTV varied to a much greater degree with GI_{PTV} of 3.54 ± 0.72 (see Table 1).

OAR doses

The closest OARs were the esophagus (case 1), heart (case 2), and duodenum (case 3). Their corresponding absolute maximum doses were 18.8 ± 3.5 Gy (goal ≤ 21 Gy), 35.3 ± 5.3 Gy (goal ≤ 30.0 Gy), and 24.0 ± 4.3 Gy (goal ≤ 24 Gy) for the esophagus, heart, and duodenum, respectively (see Table 1). The V_{15Gy} of the liver was 694.5 ± 149.5 cm^3 (case 1, goal ≤ 1650 cm^3), 489.5 ± 74.7 cm^3 (case 2, goal ≤ 434 cm^3) and 244.3 ± 28.0 cm^3 (case 3, goal ≤ 900 cm^3). Further details are presented in Figure 4.

Deviations from clinical goals

Overall, only 20 treatment plans (18 for case 1 using various techniques and 1 PT plan each for case 1 and 3) had no deviations from the clinical goals. Out of the total 132 we found 55 treatment plans with at least 1 major deviation from clinical goals. Concerning specific deviations from GTV and PTV objectives we found 2/0 and 6/0 minor/major

Table 2 Ranking evaluation for the GTV $D_{50\%}$ prescription grouped by system for main and subscores*

Ranking evaluation grouped by system All plan metrics, all systems	3D-CRT	SF-IMRT	IMAT	RRS	HRT	PT	Kruskal-Wallis and post hoc Tukey's honest test All systems	$P < .001$
Minimum	2.0	2.0	1.0	1.0	2.0	1.0	PT versus 3D-CRT	$P < .001$
Maximum	4.0	4.0	4.0	3.0	4.0	2.0	PT versus SF-IMRT	$P = .002$
Mean	3.1	2.7	2.6	2.3	2.9	1.1	PT versus IMAT	$P < .001$
Standard deviation	1.0	0.8	1.5	0.8	0.8	0.3	PT versus RSS	$P = .047$
Median	3.0	3.0	2.0	2.0	3.0	1.0	PT versus HRT	$P < .001$
All OARs, all systems							All systems	$P < .001$
Minimum	2.0	2.0	1.0	1.0	2.0	1.0	PT versus 3D-CRT	$P < .001$
Maximum	3.0	4.0	4.0	4.0	4.0	1.0	PT versus SF-IMRT	$P < .001$
Mean	2.7	3.0	2.6	2.4	2.9	1.0	PT versus IMAT	$P < .001$
Standard deviation	0.9	0.9	1.4	0.8	0.8	0.3	PT versus RSS	$P = .008$
Median	3.0	3.0	3.0	2.0	3.0	1.0	PT versus HRT	$P < .001$
All OARs, PT excluded							PT excluded	$P = .01$
Minimum	2.0	1.0	1.0	1.0	1.0	NA	RSS versus 3D-CRT	$P = .199$
Maximum	3.0	4.0	4.0	3.0	4.0	NA	RSS versus SF-IMRT	$P = .077$
Mean	2.5	2.7	2.4	1.8	2.9	NA	RSS versus IMAT	$P = .185$
Standard deviation	0.8	0.8	1.4	0.6	0.8	NA	RSS versus HRT	$P = .007$
Median	2.0	3.0	2.0	2.0	3.0	NA	Other systems	$P > .19$
OAR close to PTV, PT excluded							PT excluded	$P < .001$
Minimum	1.0	1.0	1.0	1.0	1.0	NA	RSS versus 3D-CRT	$P < .001$
Maximum	4.0	4.0	4.0	3.0	4.0	NA	RSS versus SF-IMRT	$P = .002$
Mean	2.9	2.9	2.3	1.4	2.8	NA	RSS versus IMAT	$P = .01$
Standard deviation	1.0	0.8	1.4	0.5	0.8	NA	RSS versus HRT	$P = .001$
Median	3.0	3.0	2.0	1.0	3.0	NA	Other systems	$P > .25$
Protocol deviations, all systems							All systems	$P = .027$
Minimum	2.0	1.0	1.0	1.0	2.0	1.0	PT versus HRT	$P = .035$
Maximum	4.0	4.0	4.0	3.0	4.0	3.0	Other systems	$P > .12$
Mean	2.3	2.5	2.1	1.9	2.6	1.4		
Standard deviation	0.8	0.8	1.3	0.7	0.8	0.4		
Median	2.0	2.0	2.0	2.0	2.0	1.0		

Abbreviations: 3D-CRT = 3-dimensional conformal radiation therapy; HRT = helical radiation therapy; IMAT = intensity modulated arc therapy; NA = not applicable; OAR = organ at risk; PT = proton therapy; PTV = planning target volume; RRS = robotic radiation surgery; SF-IMRT = static field intensity modulated radiation therapy.

* Values for 3D-CRT, SF-IMRT, IMAT, RRS, HRT, and PT are based on a score of 1 to 4, as described in Methods and Materials. The Kruskal-Wallis test results of technique differences include an omnibus for all systems and a post hoc Tukey's honest test adjusted for pairwise comparison.

Table 3 Absolute deviations from planning objectives and the ranking evaluation of the absolute differences

	Absolute deviations from planning objectives (Gy)					Ranking evaluation of the absolute deviations				
Prescription method	**GTV $D_{50\%}$**	**PTV $D_{98\%}$**	**PTV $D_{2\%}$**	**PTV D_{min}**	**PTV D_{max}**	**GTV $D_{50\%}$**	**PTV $D_{98\%}$**	**PTV $D_{2\%}$**	**PTV D_{min}**	**PTV D_{max}**
IMAT	5.26	7.71	10.0	66.12	6.04	2.03	1.78	2.12	1.81	2.10
SF-IMRT	4.27	6.83	9.7	35.71	4.58	2.00	1.81	2.27	1.27	1.78
HRT	7.46	12.54	12.87	107.52	9.34	2.50	2.25	2.67	2.33	2.75
RRS	4.43	9.45	7.24	60.84	5.21	1.87	1.81	1.50	2.06	1.87
PT	8.97	17.99	15.68	116.63	10.44	3.22	2.89	3.00	2.56	3.11
3D-CRT	4.57	7.19	9.07	80.30	6.41	2.25	1.87	2.06	2.50	2.31
	P values									
Comparing all systems	.003*	.011*	<.001*	.003*	.001*	.002*	.037*	< .001*	< .001*	.002*
IMAT versus SF-IMRT	1.000	.992	1.000	.746	.922	1.000	.998	.990	.382	.807
IMAT versus HRT	.317	.362	.239	.227	.248	.493	.757	.338	.444	.285
IMAT versus RRS	.838	.999	.223	.995	.911	.984	1.000	.127	.878	.936
IMAT versus PT	.010*	.012*	.007*	.375	.012*	.002*	.014*	.032*	.168	.020*
IMAT versus 3D-CRT	.998	.935	.994	.045	.969	.899	.993	1.000	.037*	.937
SF-IMRT versus HRT	.577	.840	.674	.069	.155	.747	.985	.909	.046*	.108
SF-IMRT versus RRS	.976	.992	.485	.653	1.000	.998	1.000	.204	.160	.999
SF-IMRT versus PT	.067	.170	.106	.117	.011*	.029*	.209	.386	.014*	.009*
SF-IMRT versus 3D-CRT	.999	.999	.992	.018*	.738	.968	1.000	.991	.003*	.525
HRT versus RRS	.110	.402	.008*	.707	.116	.380	.902	.007*	.982	.153
HRT versus PT	.811	.798	.832	1.000	.872	.464	.554	.921	.991	.903
HRT versus 3D-CRT	.744	.940	.228	1.000	.832	.985	.982	.501	.987	.904
RRS versus PT	.003*	.024*	<.001*	.787	.006*	.003*	.071	<.001*	.799	.012*
RRS versus 3D-CRT	.801	.899	.777	.437	.713	.750	.999	.436	.695	.679
3D-CRT versus PT	.102	.238	.012*	1.000	.195	.117	.155	.086	1.000	.309

Abbreviations: 3D-CRT = 3-dimensional conformal radiation therapy; GTV = gross target volume; HRT = helical radiation therapy; IMAT = intensity modulated arc therapy; PT = proton therapy; PTV = planning target volume; RRS = robotic radiation surgery; SF-IMRT = static field intensity modulated radiation therapy.

* Significant differences.

(case 1), 4/3 and 9/8 minor/major (case 2), and 6/2 and 12/3 minor/major (case 3) deviations. For the OAR dose, we found 33/3 minor/major (case 1, major = stomach and skin D_{1cc}), 73/39 minor/major (case 2, major mainly = liver V_{17Gy} and heart D_{1cc}), and 80/49 minor/major (case 3, major mainly = duodenum $V_{15/18Gy}$ and skin D_{1cc}) deviations. Based on a case-by-case evaluation, we found that only 23.1%, 18.2%, and 7.1% of the plans with these deviations traded GTV/PTV dose coverage for meeting the OAR limits, which was the individual decision of the residing radiation oncologist.

Best practice guidelines

Based on the individual relative plan scoring and ranking system used for this study (averaged subscores for combined and selective plan metrics and deviations from clinical goals as described in Methods and Materials), we selected 5 individual planners with the best scores for IMAT, SF-IMRT, RRS, HRT, and PT to present their best practice approach for liver SBRT (Supplementary Materials). The ALARA concept was in general followed in this study; however, outliers from this concept were also noted for all 3 cases (Fig. 4).

Treatment techniques

PT significantly outperformed all other planning techniques, showing the best averaged and selective subscores for overall plan quality, dosimetry of close OAR and all OAR, and summed deviations from clinical goals (Table 2, $P < .001$). Excluding PT, RRS significantly outperformed the other techniques in terms of subscore for close OARs ($P < .01$ for paired tests, Table 2). However, regarding all OARs combined, PT outperformed only HRT ($P < .01$, Table 2). RRS showed the best dose conformity over all systems and the highest GTV and PTV HI (meaning the least homogeneous dose). Further details for the primary evaluation are presented in Tables 2 and 3.

Treatment delivery times per fraction (without patient setup), were lowest for flattening filter-free IMAT (mean,

2.0 min; range, 1.7-2.8 min), followed by 3-dimensional conformal radiation therapy (mean, 4.4 min; range, 3.0-7.0 min), IMAT (all energies, mean, 6.3 min; range, 1.7-15.0 min), SF-IMRT (mean, 6.8 min; range, 5.0-7.0 min), and PT (mean, 9.6 min; range, 7.7-10.0 min), tailed by HRT (mean, 36.0 min; range, 26.6-60.9) and RRS (mean, 59.0 min; range, 26.0-91.0 min).

Renormalization to different prescription methods

Five different renormalization methods were performed as described in Methods and Materials (PTV D_{min}, $D_{98\%}$, $D_{2\%}$, D_{max}, GTV $D_{50\%}$), resulting in a total of 660 treatments plans that were evaluated based on GTV and PTV metrics to study the effect of dose renormalization to different prescription methods. In terms of absolute deviation from the planning objectives for GTV and PTV we found that the GTV $D_{50\%}$ prescription had the smallest mean differences across all parameters (5.5 ± 3.9 Gy), followed by the PTV D_{max}, $D_{98\%}$, and $D_{2\%}$ prescription (6.5 ± 4.2 Gy, 8.9 ± 9.6 Gy, and 10.7 ± 6.9 Gy) and, finally, the PTV D_{min} prescription (71.9 ± 101.3 Gy). Dose-volume histogram graphs for all cases are presented in Figure 3. Using the adjusted Kruskal-Wallis test, the pairwise comparisons for the deviations were significant ($P < .001$) for all combinations but for the PTV D_{max} prescription compared with GTV $D_{50\%}$ ($P = .36$) and PTV $D_{98\%}$ ($P = .74$) prescriptions, while the GTV $D_{50\%}$ and PTV $D_{98\%}$ had again significant differences in deviations ($P = .02$).

In terms of comparing different systems with each other for the deviations from the GTV and PTV planning objectives and based on the various dose prescription methods we found that all systems significantly differed from each other ($P < .04$ for all systems and prescription, Table 3). That did not change when the group ranking method was applied to the absolute deviations. Aside from the PTV D_{min} prescription, which had the largest deviations for all systems, we found that the GTV $D_{50\%}$ and the PTV $D_{98\%}$ prescription showed the least significant differences between the systems in the pair-wise comparisons (see Table 3). Here, only PT showed significant differences in comparison with IMAT ($P = .010$-0.012) and RRS ($P = .003$-0.024). For the group ranking evaluation of the absolute deviation, PT kept significant differences with IMAT ($P = .014$) for the PTV $D_{98\%}$ prescription while for the GTV $D_{50\%}$ prescription PT was significantly different from IMAT ($P = .002$), SF-IMRT ($P = .029$), and RRS ($P = .003$)

RATING score

Recently, radiation therapy treatment planning study guidelines (RATING) were published along with a scoring metric for study quality assessment.[34] Based on self-assessment of our study and the evaluation of 2 independent experts we achieved a mean RATING score of 98% (RATING fraction 201 out of 205 points, Table E2), which was validated during the authors' review process of the manuscript.

Discussion

To our knowledge, with 132 submitted plans from 35 institutions, this is one of the largest treatment planning studies to date, demonstrating the large interest in interinstitutional collaboration and exchange of information. The present study demonstrates the possibility to generate very similar SBRT treatment plans with various TPSs for a variety of treatment delivery systems including all common techniques. The multiparametric specification of target dose (GTV $D_{50\%}$, GTV $D_{0.1cc}$, $GTV_{V90\%}$, $PTV_{V70\%}$) leads to more harmonized plans than in our previous non-small cell lung cancer planning study[15] with the objective of a PTV encompassing dose and the prescription isodose line. More details are presented in the Supplementary Materials.

Prospective and retrospective clinical studies can suffer from a large variability of target doses, even if the dose prescription is nominally the same for all patients, especially if different planning and delivery techniques are used in a multicenter setting.[21,35] This is due to unspecific or a too limited set of planning objectives and may lead to inconsistent dose-to-outcome correlations, if not at least a sufficient set of dose parameters is reported for analysis. Because of that, we recommend multiparametric target dose objectives for prospective clinical trials, which can lead to a harmonized patient plan collective, as demonstrated in this study. Another important aspect for clinical studies is the recommendation or at least documentation of the dose calculation algorithm type. Here type B algorithms were proposed, which was fulfilled from all participants (only anisotropic analytical algorithms is an intermediate algorithm, according to the report by Seuntjens et al[24]), except the proton facilities, where the pencil beam algorithm is the clinical standard. Because of that and the small involvement of tissue inhomogeneities in the liver cases, we do not expect an influence of the dose calculation algorithm on our results.

The OAR limits were not met in all plans and for the 3 cases, only 23.1%, 18.2%, and 7.1% of the plans with deviations traded GTV/PTV dose coverage for meeting the OAR limits. Because of the recommendation to submit only clinically acceptable plans, this may reflect different clinical practice between the institutions. The “best practice” guideline in the Supplementary Materials might be helpful in balancing the target and OAR goals given the differences in approach we saw.

The ranking of treatment techniques must be seen under the limitation of the assumption of an active motion management during the beam delivery using no ITV and a CTV-to-PTV margin of only 3 mm. The accuracy of different treatment delivery and motion monitoring techniques was not considered. This would only be possible through individual margin definition, which counteracts our method of plan evaluation. In particular, the fact that PT plans outperformed all photon techniques in all cases regarding target and OAR goals does not necessarily mean a better treatment in liver SBRT. Active motion management techniques in PT are still under investigation.[36,37]

In the current investigation the prescription to the GTV $D_{50\%}$ showed the smallest amount of deviations from the planning objectives and, even more importantly, between different delivery techniques. It is the main limitation of the study that only for this single dose prescription the optimization was done and thus the result is likely biased, especially for the PTV D_{min} prescription. To minimize the bias, all other dose prescriptions used for renormalization of all plans were part of the set of dose objectives, hence a plan was not necessarily changed through renormalization, if all objectives were met. The alternative would have included obtaining different optimized and prescribed plans for all cases from the planners. However, this would have resulted in 3 × 5 plans per planner and we considered the workload involved too excessive for study participants in this scenario.

Conclusions

This study shows the feasibility of harmonizing liver SBRT treatment plans across different TPSs and delivery techniques when a sufficient set of clinical goals is given. The method of GTV $D_{50\%}$ prescription can be performed in all systems, improving overall consistency. The ALARA principle was followed for most OARs, but in many plans dose limits in OARs close to the target were exceeded to meet the target dose. Besides the comparison between different treatment techniques and platforms, advice for planning strategies is provided in the Supplementary Materials.

References

1. Guckenberger M, Baus WW, Blanck O, et al. Definition and quality requirements for stereotactic radiotherapy: consensus statement from the DEGRO/DGMP Working Group Stereotactic Radiotherapy and Radiosurgery. *Strahlenther Onkol* 2020;196:417–420.
2. Klement RJ, Sonke JJ, Allgäuer M, et al. Correlating dose variables with local tumor control in stereotactic body radiation therapy for early-stage non-small cell lung cancer: A modeling study on 1500 individual treatments. *Int J Radiat Oncol Biol Phys* 2020;107:579–586.
3. Mazzola R, Ruggieri R, Figlia V, et al. Stereotactic body radiotherapy of central lung malignancies using a simultaneous integrated protection approach: A prospective observational study. *Strahlenther Onkol* 2019;195:719–724.
4. Hörner-Rieber J, Bernhardt D, Blanck O, et al. Long-term follow-up and patterns of recurrence of patients with oligometastatic NSCLC treated with pulmonary SBRT. *Clin Lung Cancer* 2019;20:e667. -e677.
5. Chiang CL, Chan MKH, Yeung CSY, et al. Combined stereotactic body radiotherapy and trans-arterial chemoembolization as initial treatment in BCLC stage B-C hepatocellular carcinoma. *Strahlenther Onkol* 2019;195:254–264.
6. Gkika E, Strouthos I, Kirste S, et al. Repeated SBRT for in- and out-of-field recurrences in the liver. *Strahlenther Onkol* 2019;195:246–253.
7. Andratschke N, Alheid H, Allgäuer M, et al. The SBRT database initiative of the German Society for Radiation Oncology (DEGRO): patterns of care and outcome analysis of stereotactic body radiotherapy (SBRT) for liver oligometastases in 474 patients with 623 metastases. *BMC Cancer* 2018;18:283.
8. Guckenberger M, Sweeney RA, Hawkins M, et al. Dose-intensified hypofractionated stereotactic body radiation therapy for painful spinal metastases: Results of a phase 2 study. *Cancer* 2018;124:2001–2009.
9. Méndez Romero A, Schillemans W, van Os R, et al. The Dutch-Belgian registry of stereotactic body radiation therapy for liver metastases: Clinical outcomes of 515 patients and 668 metastases. *Int J Radiat Oncol Biol Phys* 2021;109:1377–1386 S0360-3016(20)34567-3.
10. Mahadevan A, Blanck O, Lanciano R, et al. Stereotactic body radiotherapy (SBRT) for liver metastasis - clinical outcomes from the international multi-institutional RSSearch® Patient Registry. *Radiat Oncol* 2018;13:26.
11. Sterzing F, Brunner TB, Ernst I, et al. Stereotactic body radiotherapy for liver tumors: principles and practical guidelines of the DEGRO Working Group on Stereotactic Radiotherapy. *Strahlenther Onkol* 2014;190:872–881.
12. Brunner TB, Blanck O, Lewitzki V, et al. Stereotactic body radiotherapy dose and its impact on local control and overall survival of patients for locally advanced intrahepatic and extrahepatic cholangiocarcinoma. *Radiother Oncol* 2019;132:42–47.
13. Boda-Heggemann J, Jahnke A, Chan MKH, et al. In-vivo treatment accuracy analysis of active motion-compensated liver SBRT through registration of plan dose to post-therapeutic MRI-morphologic alterations. *Radiother Oncol* 2019;134:158–165.
14. Moustakis C, Chan MKH, Kim J, et al. Treatment planning for spinal radiosurgery: A competitive multiplatform benchmark challenge. *Strahlenther Onkol* 2018;194:843–854.
15. Moustakis C, Blanck O, Ebrahimi Tazehmahalleh F, et al. Planning benchmark study for SBRT of early stage NSCLC: Results of the DEGRO Working Group Stereotactic Radiotherapy. *Strahlenther Onkol* 2017;193:780–790.
16. Habraken SJM, Sharfo AWM, Buijsen J, et al. The TRENDY multi-center randomized trial on hepatocellular carcinoma - Trial QA including automated treatment planning and benchmark-case results. *Radiother Oncol* 2017;125:507–513.
17. Esposito M, Maggi G, Marino C, et al. Multicentre treatment planning inter-comparison in a national context: The liver stereotactic ablative radiotherapy case. *Phys Med* 2016;32:277–283.
18. Giglioli FR, Garibaldi C, Blanck O, et al. Dosimetric multicenter planning comparison studies for stereotactic body radiation therapy: Methodology and future perspectives. *Int J Radiat Oncol Biol Phys* 2020;106:403–412.
19. Villaggi E, Hernandez V, Fusella M, et al. Plan quality improvement by DVH sharing and planner's experience: Results of a SBRT multicentric planning study on prostate. *Phys Med* 2019;62:73–82.
20. Esposito M, Masi L, Zani M, et al. SBRT planning for spinal metastasis: Indications from a large multicentric study. *Strahlenther Onkol* 2019;195:226–235.
21. Giglioli FR, Strigari L, Ragona R, et al. Lung stereotactic ablative body radiotherapy: A large scale multi-institutional planning comparison for interpreting results of multi-institutional studies. *Phys Med* 2016;32:600–606.
22. Wilke L, Moustakis C, Blanck O, et al. Improving inter-institutional and inter-technology consistency of pulmonary SBRT by dose prescription to the mean ITV dose. *Strahlther Onkol* 2021;197:836–846.
23. Seuntjens J, Lartigau EF, Cora S, et al. ICRU report 91. Prescribing, recording, and reporting of stereotactic treatments with small photon beams. *J ICRU* 2014;14:1–160.
24. Wilke L, Andratschke N, Blanck O, et al. ICRU report 91 on prescribing, recording, and reporting of stereotactic treatments with small photon beams: Statement from the DEGRO/DGMP working group stereotactic radiotherapy and radiosurgery. *Strahlenther Onkol* 2019;195:193–198.
25. Klement RJ, Guckenberger M, Alheid H, et al. Stereotactic body radiotherapy for oligo-metastatic liver disease - Influence of pre-treatment chemotherapy and histology on local tumor control. *Radiother Oncol* 2017;123:227–233.
26. Stera S, Balermpas P, Chan MKH, et al. Breathing-motion-compensated robotic guided stereotactic body radiation therapy: Patterns of failure analysis. *Strahlenther Onkol* 2018;194:143–155.
27. Grimm J, LaCouture T, Croce R, et al. Dose tolerance limits and dose volume histogram evaluation for stereotactic body radiotherapy. *J Appl Clin Med Phys* 2011;12:3368.

28. Schmitt D, Blanck O, Gauer T, et al. Technological quality requirements for stereotactic radiotherapy: Expert review group consensus from the DGMP Working Group for Physics and Technology in Stereotactic Radiotherapy. *Strahlenther Onkol* 2020;196:421–443.
29. Klement RJ, Abbasi-Senger N, Adebahr S, et al. The impact of local control on overall survival after stereotactic body radiotherapy for liver and lung metastases from colorectal cancer: A combined analysis of 388 patients with 500 metastases. *BMC Cancer* 2019;19:173.
30. Guckenberger M, Andratschke N, Alheit H, et al. Definition of stereotactic body radiotherapy: principles and practice for the treatment of stage I non-small cell lung cancer. *Strahlenther Onkol* 2014;190:26–33.
31. Paddick I, Lippitz B. A simple dose gradient measurement tool to complement the conformity index. *J Neurosurg* 2006;105(Supplement):194–201.
32. Blanck O, Wang L, Baus W, et al. Inverse treatment planning for spinal robotic radiosurgery: an international multi-institutional benchmark trial. *J Appl Clin Med Phys* 2016;17:313–330.
33. Wagner TH, Bova FJ, Friedman WA, et al. A simple and reliable index for scoring rival stereotactic radiosurgery plans. *Int J Radiat Oncol Biol Phys* 2003;57:1141–1149.
34. Rønn Hansen C, Crijns W, Hussein M, et al. RAdiotherapy Treatment plannINg study Guidelines (RATING): A framework for setting up and reporting on scientific treatment planning studies. *Radiother Oncol* 2020;153:67–78.
35. Guckenberger M, Andratschke N, Dieckmann K, et al. ESTRO ACROP consensus guideline on implementation and practice of stereotactic body radiotherapy for peripherally located early stage non-small cell lung cancer. *Radiother Oncol* 2017;124:11–17.
36. Paganetti H, Beltran CJ, Both S, et al. Roadmap: Proton therapy physics and biology [e-pub ahead of print]. *Phys Med Biol* 2020 Feb 26;66(5). https://doi.org/10.1088/1361-6560/abcd16.
37. Ribeiro CO, Visser S, Korevaar EW, et al. Towards the clinical implementation of intensity-modulated proton therapy for thoracic indications with moderate motion: Robust optimised plan evaluation by means of patient and machine specific information. *Radiother Oncol* 2021;2 S0167-8140(21)00014-1.