Evaluation of an anthropomorphic ion chamber and 3D gel dosimetry head phantom at a 0.35 T MR-linac using separate 1.5 T MR-scanners for gel readout

Lukas Nierer[a], Florian Kamp[a,b], Michael Reiner[a], Stefanie Corradini[a], Moritz Rabe[a], Olaf Dietrich[c], Katia Parodi[d], Claus Belka[a,e], Christopher Kurz[a], Guillaume Landry[a,*]

[a] Department of Radiation Oncology, University Hospital, LMU Munich, Marchioninistr. 15, 81377 Munich, Germany
[b] Department of Radiation Oncology, University Hospital Cologne, Kerpener Str. 62, 50937 Cologne, Germany
[c] Department of Radiology, University Hospital, LMU Munich, Marchioninistr. 15, 81377 Munich, Germany
[d] Department of Medical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München, 85748 Garching, Germany
[e] German Cancer Consortium (DKTK), partner site Munich, Munich, Germany

Received 10 September 2021; accepted 31 January 2022

Abstract

Purpose: *To date, no universally accepted technique for the evaluation of the overall dosimetric performance of hybrid integrated magnetic resonance imaging (MR) – linear accelerators (linacs) is available. We report on the suitability and reliability of a novel phantom with modular inserts for combined polymer gel (PG) and ionisation chamber (IC) measurements at a 0.35 T MR-linac.*

Methods: *Three 3D-printed, modular head phantoms, based on real patient anatomy, were used for repeated (2 times) PG irradiations of cranial treatment plans on a 0.35 T MR-linac. The PG readout was performed on two 1.5 T diagnostic MR-scanners to reduce scanning time. The PG dose volumes were normalised to the IC dose (normalised dose N1) and to the median planning target volume dose (normalised dose N2). Linearity of the PG dose response was validated and dose profiles, centres of mass (COM) of the 95% isodoses and dose volume histograms (DVH) were compared between planned and measured dose distributions and a 3D gamma analysis was performed.*

Results: *Dose linearity of the PG was good ($R^2 > 0.99$ for all linear fit functions). High agreement was found between planned and measured dose volumes in the dose profiles and DVHs. The largest dose deviation was found in the intermediate dose region (mean dose deviation 0.2 Gy; 5.6%). A mean COM offset of 1.2 mm indicated high spatial accuracy. Mean 3D gamma passing rates (2%, 2 mm) of 83.3% for N1 and 91.6% for N2 dose distributions were determined. When comparing repeated PG measurements to each other, a mean gamma passing rate of 95.7% was found.*

Conclusion: *The new modular phantom was found practical for use at a 0.35 T MR-linac. In contrast to the high dose region, larger mean deviations were found in the mid dose range. The PG measurements showed high reproducibility. The MR-linac performed well in a non-adaptive setting in terms of spatial and dosimetric accuracy.*

Keywords: Polymer gel dosimetry, MR-guided radiotherapy, Multimodality phantom, Dosimetric accuracy

*Corresponding author: Lukas Nierer, Department of Radiation Oncology, University Hospital, LMU Munich, Marchioninistr. 15, 81377 Munich, Germany
E-mail: lukas.nierer@med.uni-muenchen.de (L. Nierer).

Z Med Phys **32** (2022) 312–325
https://doi.org/10.1016/j.zemedi.2022.01.006
www.elsevier.com/locate/zemedi

Introduction

Hybrid integrated magnetic resonance imaging – linear accelerators (MR-linacs) have been in clinical use for about four years now and the global number of such machines grows continuously [1]. MR-linacs show high potential for improving radiotherapy (RT) cancer treatment. While, to this date, literature is still lacking extensive phase 3 study evidence for improved patient survival or reduced side effects, it is already clear that some entities, previously considered unsuitable for RT (e.g. inoperable pancreatic cancer or multiple abdominal lymph nodes), can now be treated with MR-linacs [2,3]. This is due to the online plan adaption and tumour tracking capabilities, based on the daily 3D setup MR image of the patient in treatment position, and cine MR imaging during the treatment application, respectively. In contrast to conventional image-guided RT (IGRT), online MR-guided RT (oMRgRT) offers superior soft-tissue contrast, which allows for the direct visualisation of the tumour in many cases [4]. The clinical application of MR-linacs is described widely in literature [5–12]. Information about the technical design of the 6 MV, flattening filter free, 0.35 T MR-linac (ViewRay Inc., Oakwood Village, OH, USA) is provided by Klüter [13].

Apart from the clinical development of MR-linac use, several quality assurance (QA) related challenges arise [14]. Checks of the MR and the RT systems make up a large part of the QA regimen. It is not sufficient to solely test each system independently; checks applicable to both systems simultaneously are desirable. An example of such integrated tests would be the isocentricity check, which ensures that the isocentres of the MR, RT and, if used, laser system match within tolerance [15]. It is desirable to verify the integrity of the oMRgRT treatment chain, beginning with image acquisition for offline treatment planning, generation of the baseline treatment plan, patient / phantom setup, online imaging and treatment application. If such a test shall be performed at an MR-linac, the employed phantom should provide good CT and MR contrast, resemble patient anatomy, provide volumetric absolute dose information to maximise information content, and have tissue-like material properties. In the case of photon beams, dosimetric tissue equivalency means that the photon interaction properties of the phantom materials should be similar to those of the tissue materials, which they mimic, for a given photon beam quality. Polymer gel (PG) based 3D dosimetry in combination with an anthropomorphic phantom meets these requirements [16]. PG dosimetry relies on the change of relaxation rates due to radiation-induced polymerisation, which can be measured with an MRI scanner. A linear correlation between absorbed dose and change in relaxation rate can be seen in certain, PG-dependent dose ranges. Detailed information on the characteristics and applications of PG dosimetry can be found in literature [17–25]. PGs are suitable for 3D dosimetry, and the PG serves as soft tissue equivalent material and dosimeter at the same time [26]. Furthermore, the influence of the magnetic field on PG dose measurements is small [27].

Dorsch et al. developed a new phantom utilising PG dosimetry, in order to investigate RT- and MR-isocentre alignment and spatial distortion of the MR [28]. Pappas et al. used commercially available head phantoms, similar to those used in this study (but unable to house both IC and PG), and investigated the dosimetric performance of a 1.5 T MR-linac [29]. Elter et al. developed a phantom with deformable anthropomorphic structures and validated an adaptive workflow with small PG filled volumes at a 0.35 T MR-linac [30]. Hoffmans at al. developed a deformable pelvis phantom and investigated an adaptive workflow at a 0.35 T MR-linac utilising 2D film dosimetry [31]. Axford et al. employed a pelvic phantom for validating an adaptive scenario at a 1.5 T MR-linac with dosimetric film and alanine [32]. Steinmann et al. developed a head and neck phantom and performed reproducibility tests at 1.5 T and 0.35 T MR-linacs using radiochromic film and thermoluminescence dosimetry [33]. Bernchou et al. investigated the spatial deviation between planned and delivered dose in many adapted fractions at two 1.5 T MR-linacs with a self-developed phantom and radiochromic film [34]. Mittauer et al. used a deformable anthropomorphic phantom to quantify the image deformation accuracy at a 0.35 T MR-linac and used thermoluminescence dosimetry for point dose accumulation [35]. Stark et al. reported on the stability of their gafchromic film-based, clinical QA end-to-end tests over a period of one year [36].

In this study, three patient-specific multimodal 3D printed head phantoms were evaluated for their suitability for dose measurements at a 0.35 T MR-linac, with dosimetric gel readout at separate 1.5 T MR scanners. These phantoms are similar to those used by Pappas et al. but with the added option to mount an IC insert [29]. We used realistic treatment plans, resembling a challenging off-axis setup with an organ at risk (OAR) adjacent to the target volume. The recently available phantom design [37] allowed the insertion of a PG volume or an ionisation chamber (IC), thus enabling measurement-based dose normalisation. The aim of the study was to evaluate the suitability and reproducibility of the novel phantom for use at a 0.35 T MR-linac.

Materials and methods

Preparation, treatment planning and plan validation

Anthropomorphic head phantoms

Three anthropomorphic head phantoms, resembling real patient anatomies, were used for this study. The phantoms were 3D printed, based on a patient CT scan and are commercially available (PseudoPatient^{TM} Prime, RTsafe P.C., Athens, Greece). Further information about an earlier version, housing only PG, of the 3D printed phantom design is provided by Makris et al [38]. The phantoms used in this study consist of a

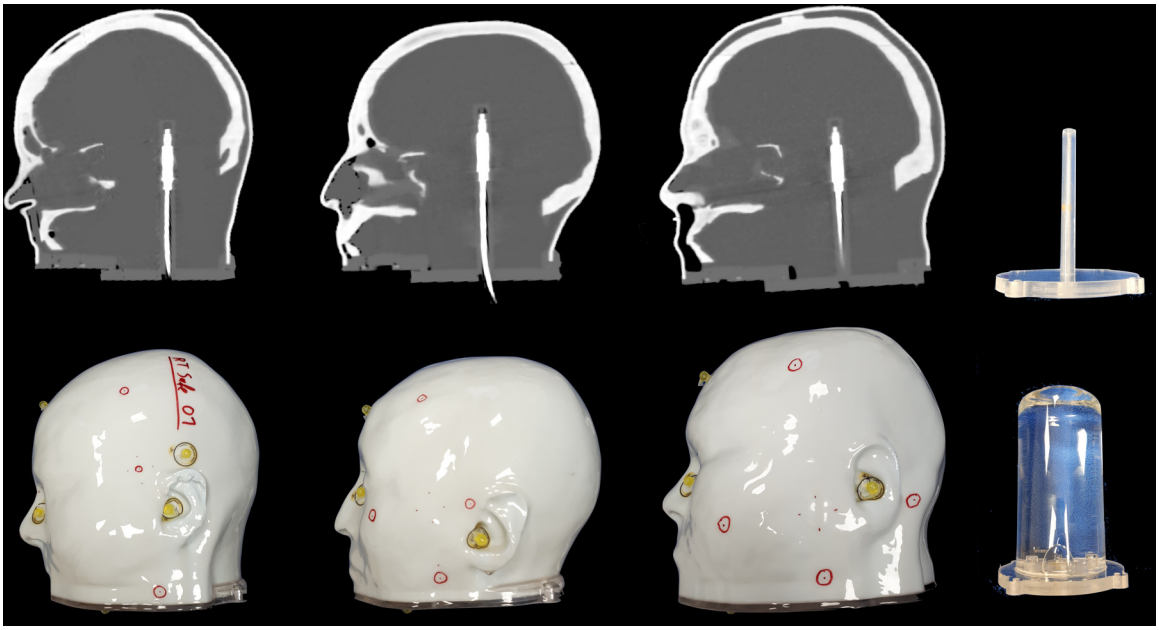

Figure 1. (upper half) Sagittal CT slices of the planning CTs (IC insert with IC in place) of the three head phantoms (numbered 1, 2, 3 from left to right) and photograph of the IC insert. Air density override was performed in the TPS for the CT images. (lower half) Corresponding photographs of the three different head phantoms and an example of the PG filled cylindrical insert prior to irradiation.

PMMA base and a bone-mimicking material used for printing (Hounsfield unit ≈ 1300 at 120 kV peak voltage; physical density ≈ 1.7 g/cm^3) [38]. The modular design of the phantoms allows for the placement of one insert for IC measurements and one PG insert, which consists of a PG-filled glass cylinder (Ø 80 mm, length 165 mm) mounted on a PMMA plate. The residual volume inside the hollow phantom, not occupied by the inserts, was filled with pure water for soft tissue equivalency (in terms of CT contrast and dosimetric properties). Figure 1 shows sagittal CT slices and photographs of the three phantoms and examples of the IC and PG inserts. The head phantoms dimensions range between 13 × 17 × 17 cm^3 (LR, SI, AP directions; head 2) and 15 × 19 × 21 cm^3 (head 3).

CT acquisition

First, computed tomography (CT) scans were acquired for all three phantoms on an Aquilion LB CT scanner (Canon Medical Systems Corp., Otawara, Japan), with the IC and with the PG inserts. The phantoms were positioned on a foam head wedge. Prior to the first CT scan of each phantom, the wall-mounted calibrated in-room lasers at the CT scanner were used to mark the phantoms to reproducibly set up the phantoms without any roll / tilt deviation. Additionally, CT markers were attached to the phantoms to define the CT reference point. For the IC insert, the cavity for the IC was first filled with water, then the IC (Semiflex 3D MR 31021, PTW, Freiburg, Germany) was inserted to remove all air between IC and IC insert wall, since any air near the IC would later affect dosimetry in the presence of the magnetic field of the 0.35 T MR-linac. The IC was taped to the phantom to prevent it from moving. The exact position of the IC insert was marked on the IC insert and the phantoms. In case of the PG insert CT scans, a dummy water-filled insert was used, which was not used for dose measurements afterwards. All CTs were acquired with a resolution of 1 mm (isotropic voxels) using a stereotactic head scanning protocol.

Treatment planning

The CTs of each head phantom were then transferred to the MR-linac treatment planning system (TPS of the ViewRay system; version 5.4) and registered to each other via rigid image registration (translation). Several regions of interest were segmented for each phantom: a brainstem-like structure in close proximity to a C-shaped planning target volume (PTV) inside the PG volume, another OAR structure in the region of the eyes (for a more realistic dose optimisation), a structure encompassing the PG volume (override with relative electron density of 1.03; see [20]), the 0.07 cm^3 sensitive volume of the IC and a 5 mm expansion thereof (to have a larger sampling volume in the PG for dose normalisation in order to increase the robustness and the signal-to-noise-ratio of the reference dose value) and the skin / external contour. Furthermore, a ring structure around the PTV was contoured (used to better control dose optimisation and to achieve good dose conformality around the PTV), and a region for electron density override (air bubbles inside the water-filled phantoms). Although schematically comparable, the exact shapes and

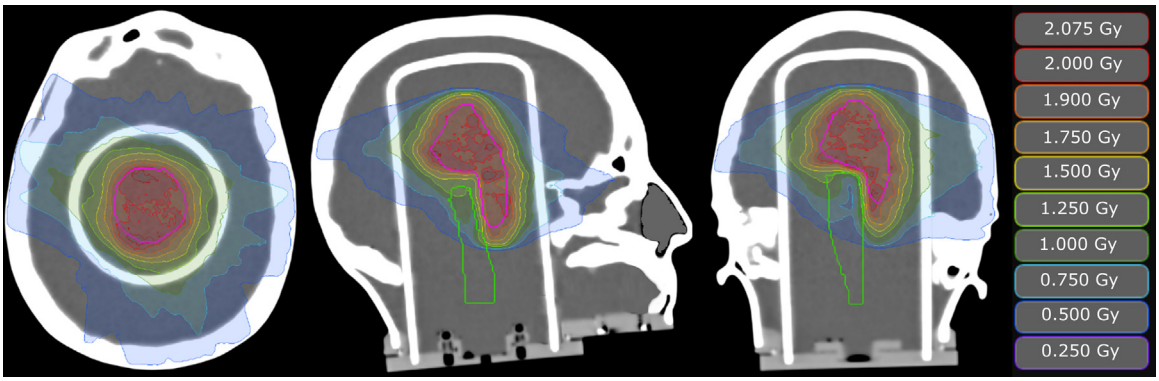

Figure 2. Axial, sagittal and coronal slices and isodoses of the treatment plan of phantom 2, calculated on the PG insert CT with electron density overrides of air-bubbles. The brainstem-like OAR is indicated in green and the PTV in pink.

Table 1
Treatment plan characteristics of the step-and-shoot intensity modulated radiotherapy (IMRT) plans for each phantom.

Phantom No.	1	2	3
PTV volume [cc]	30.2	37.3	35.2
Number of beams	19	21	21
Number of segments	60	56	39
Off-axis distance* [cm]	4.3	4.7	6.8
Dose / fraction [Gy]	2	2	2

* The off-axis distance is the distance between the RT isocentre and the centre of mass of the PTV.

locations of the brainstem and PTV structures were varied between the phantoms. A single fraction dose of 2 Gy was prescribed to the 100% isodose and we aimed at a 95% isodose coverage of the PTV. The IC sensitive volume was in a region of homogeneous dose. Table 1 shows treatment plan characteristics of the plans for each phantom. The CT reference point indicated by the CT markers was used as isocentre. Typical of oMRgRT treatments, the distance between the isocentre and the centre of mass (COM) of the PTV, called off-axis distance, were relatively large (range 4.3 cm – 6.8 cm). This resulted in a more realistic but also more demanding off-axis setup, which is more error-prone compared to isocentric treatments, because small rotational patient positioning errors might result in larger spatial offsets. The dose was calculated with a Monte Carlo statistical dose uncertainty of 0.2% and a dose grid resolution of 1 mm (isotropic voxels) and accounting for the effect of the magnetic field. The plans were initially optimised on the IC insert CT with electron density override of air bubbles. Afterwards, the final plans were re-calculated on the corresponding PG insert CTs with electron density override of air bubbles and PG volumes. Figure 2 depicts an exemplary dose distribution. To fully exploit the sensitive dose range of the PG, the final 2 Gy plans were then up-scaled by a factor of 4 to 8 Gy per fraction. The 2 Gy plans were used for the IC dose measurements and for diode array plan validation, while the 8 Gy plans were used for PG irradiation as per manufacturer specifications.

Validation of treatment plans with diode detector array

The three treatment plans underwent patient-specific plan QA similarly to clinical plans. Plans were measured with a diode detector array (ArcCheck-MR; Sun Nuclear Corporation, Melbourne, FL, USA). Measured and planned doses were compared and gamma passing rates (2%, 2 mm, absolute dose, dose threshold = 10% of maximum dose) evaluated [39]. To verify the correct setup of the array, a dummy plan was measured before the actual plan validation. This dummy plan is regularly measured in the clinical QA process, thus providing gamma passing rate reference data. The results of the dummy plan measurement allowed us to make sure that the phantom was set up without relevant (rotational and translational) deviations from the (isocentric) reference position, which is essential for reproducible plan QA measurements. After the dummy plan measurement and between the three plans no changes were made to the setup to guarantee comparability between plans. According to our internal standard of procedure (SOP), a plan is considered clinically acceptable if 90% of pixels pass the 2%, 2 mm criterion (absolute dose, dose threshold = 10% of maximum dose).

Phantom irradiation and dose readout

Phantom positioning accuracy and IC irradiations

Prior to the IC and PG irradiations, the linac output was measured in a water phantom to correct for daily variations of the dose output (daily output correction factor was determined to be 1.006 before IC and PG irradiations). This is equivalent to a re-adjustment of the linac monitor chamber gain values (which would only be adjusted in case the output deviation exceeds the limit defined in QA specifications). The head phantoms were filled with water, air bubbles were removed, the phantoms were set up on the treatment table according

Table 2
Overview of PG inserts, PG batch numbers used to fill them, phantom numbers in which the PG inserts were used, treatment plans that were irradiated and MRI scanners which were used for readout.

PG Insert No.	1	2	3	4	5	6
PG Batch No.	1	2	1	2	3	2
Phantom No.	1	1	2	2	3	3
Treatment Plan No.	1	1	2	2	3	3
Readout MRI Scanner No.	2	1	1	2	1	2

to the room lasers, the MR receiver coils were placed around the phantoms and a setup MR scan (1.5×1.5 mm^2 in-plane resolution, 3.0 mm slice thickness) was performed. Then the setup deviation was evaluated by registering the current MR image to the planning CT (manual rigid registration; translational correction only). This step was performed to verify, that the setup via room lasers results in high phantom positioning accuracy. For the subsequent IC and PG measurements, setup was performed according to the room lasers and no receiver coils were placed around the phantoms and no setup MRs were acquired. This was done to avoid confounding dosimetric effects related to variation of the position of the receiver coils. After filling of the IC insert cavity with water and insertion of the IC (as described above), the water temperature inside the IC insert was measured to correct for temperature and the plans were delivered to the phantoms. After correction of the measured dose for daily output variations, the measured dose values were compared to the mean calculated dose values of the 5 mm expansion of the IC sensitive volume of each treatment plan. The expansion of the IC sensitive volume in the TPS was used to reduce the statistical uncertainty of the TPS inside the small sensitive volume (0.07 cm3) of the IC. The obtained absolute dose values were also used later to normalise the relative dose of the PG measurement.

PG characteristics and irradiations

PG filled inserts for the 3D dose measurements were purchased from RTSafe (Athens, Greece). Further characterisation of the N-vinylpyrrolidone based PG (VIP gel) can be found in literature [40]. A total of 6 such PG filled inserts were used for the measurements, which were filled with PG from three different production batches. The first batch filled one insert, the second batch filled two of the inserts, and the third batch filled three of the inserts (also see Table 2). All three PG batches were produced on the same day under the same conditions. Two dose measurements were performed with each head phantom, switching the PG inserts in between. To evaluate the integrity and reproducibility of the PG production, PG handling, phantom setup, treatment application, dose readout, image registration and dose evaluation process, all single three head phantoms were equipped with PG inserts filled with PG from different batches. This means that for each head phantom, the same treatment plan was applied twice using different batches. All PG inserts were irradiated consecutively on the same day. The time interval between PG production of all three batches and irradiation was about two days, while the time interval between irradiation and dose readout was one day. Special attention was given to the temperature history of the PG. All PG inserts were delivered in an ice-cooled package and brought to vault temperature equilibrium (19.0 °C) in a water bath prior to irradiation. The phantoms were also brought to room temperature and filled with water of the same temperature to prevent any temperature drift during irradiation.

PG readout at MRI

Dose readout was performed one day after irradiation (to allow for stabilisation in the polymerisation process of the PG), simultaneously on two separate diagnostic MR scanners to keep scanning time manageable (scanner 1: 1.5 T MAGNETOM Aera and scanner 2: 1.5 T MAGNETOM Avanto, Siemens Healthineers AG, Erlangen, Germany). A single-shot turbo spin echo sequence was used for readout (TR = 4200 ms; 4 echoes with echo time (TE) intervals of 398 ms and the first echo starting at 38 ms (38 ms, 436 ms, 834 ms, and 1232 ms; these effective TEs refer to the echo times at the central k-space lines in phase-encoding direction with phase-encoding gradient = 0); echo spacing = 4.7 ms; echo train length = 72 (using a partial-Fourier factor of 9/16); 2.5 mm slice thickness; 128×256 matrix size (interpolated to 256×512) resulting in ~0.7 mm interpolated in-plane pixel size; 71 slices; FOV 175×350 mmy; number of averages = 8; measurement time ~40 min.), which was specifically provided for this purpose by the phantom manufacturer [41], and the standard clinical distortion correction was applied. Each readout was performed with the PG insert inside the phantom. The phantom water filling was renewed after each scan to ensure temperature stability. One readout scan took about 40 minutes and increased the water temperature by an average of 1.7 °C. The PG temperature was monitored throughout, ensuring that the PG never exceeded 21.0 °C. Table 2 gives an overview over the PG inserts, PG batches, phantoms used, irradiated plans and readout MRI scanners.

3D dose evaluation

After reconstructing the MR-images of all 4 echoes, the spin–spin relaxation time (T_2) was calculated voxel-wise by applying an exponential fit to the four data-points at the effective TEs (38 ms, 436 ms, 834 ms, and 1232 ms). This step was performed by the phantom manufacturer. Due to the inverse relationship between T_2 and absorbed dose, the inverse of the T_2 map ($1/T_2 = R_2$) is linearly dependent on the absorbed dose [19]. No PG dose calibration was performed, and we relied instead on normalisation of the PG signal. The T_2 maps were then manually registered to the planning CT (rigid registration; translation + rotation), using the geometrical bone-mimicking structures of the phantoms. In this step, the T_2 maps were re-sampled to the 1 mm isotropic grid of

the CTs. Digital Imaging and Communications in Medicine (DICOM) data conversion was performed with Plastimatch [42], image registration was done with MITK [43], and all other data processing was implemented in Matlab (The Mathworks Inc., Natick, USA).

Two different dose normalisation approaches were followed. First, the mean R_2 value in the expanded IC sensitive volume was normalised to the measured IC dose value, resulting in the normalised dose distribution N1. Second, following the procedure outlined in Pappas et al. [29], the median R_2 value of the PTV was normalised to the TPS PTV median dose, resulting in the normalised dose distribution N2. To account for the inherent baseline R_2 signal of the PG, which could, in theory, also be measured in an un-irradiated PG sample, both normalisation steps were performed after baseline subtraction, where the R_2 value corresponding to the 10^{th} percentile of the R_2 values of all voxels inside the PG volume was subtracted from the R_2 values of all voxels inside the PG volume. The 10^{th} percentile was chosen based on the distributions of R_2 values, which show peaks at this position for each phantom. These peaks correlate to parts of the PG, which are distant from the beam paths and receive very limited scatter dose. This low-dose signal can be defined as baseline signal in good approximation. To validate this, the 10^{th} percentile R_2 values were also compared to the respective average R_2 values outside the radiation fields for all six cylinders. The mean of the absolute deviation values across all six samples was 0.5%.

To evaluate the quality of the registration and the linearity of the PG response, density plots were generated for each PG measurement, showing R_2 vs. the TPS doses of each voxel inside the PG structure. A linear fit was then performed (method of least squares) on the data. Please keep in mind that this fit was not used for conversion of R_2 to dose; only to evaluate linearity.

To reduce the noise inherent to the dose read-out of the PG, a $3 \times 3 \times 3$ mm^3 median filter was applied to the PG dose. 1D profiles were then extracted and compared between the measured PG dose and the TPS dose. This was done for all PG measurements in all three cardinal directions for the N1 and N2 normalisation approaches. To quantify potential systematic deviations between measured PG doses (N1 and N2) and the TPS reference doses, mean differences were calculated in the mid (40% - 60% of prescribed dose) and high (> 80% of prescribed dose) dose ranges across all phantoms and samples. To evaluate the phantom setup, irradiation and image registration accuracy, the absolute offset between the COMs of binary masks corresponding to the 95% isodose volume was calculated (N2 and TPS dose). To estimate the overall performance of the MR-linac system at a clinically interpretable level, cumulative dose volume histograms (DVH) were generated and compared (N2 and TPS dose) for the brainstem and PTV. Additionally, a 3D gamma comparison (global 3%, 3 mm and 2%, 2 mm criteria) was made between the PG doses and the TPS doses to evaluate overall dose similarity. Finally, sample pairs of each phantom were compared to each other to evaluate the reproducibility of the PG dose measurements. A gamma comparison was performed within the PG insert and a dose threshold of 10% of the maximum dose was applied [44,45].

Table 3
Gamma passing rates (2%, 2 mm; absolute dose; threshold = 10% of max. dose) of the diode array plan verification and percent deviations between the IC obtained absolute doses and the mean dose of a 5 mm expansion of the sensitive volume structure of the TPS for all three phantoms.

Phantom No.	1	2	3
Diode array gamma passing rate [%]	98.4	93.7	97.2
Deviation IC measured vs. planned dose [%]	0.4	0.2	-0.1

Results

Diode array plan validation and ionisation chamber measurements

Gamma passing rates of the clinical QA dummy plan were in the 65th percentile of all QA dummy plan measurements ($n = 42$). The results of the diode array plan verification and the IC measurements are given in Table 3 for all three phantoms. All IC measurements differed less than 0.5% from the TPS reference. The compliance of the diode array measurements with the planned dose was also high. However, the diode array passing rate of phantom 2 was lower than those of phantoms 1 and 3. As described previously, a setup test was made prior to the IC and PG irradiations in order to determine the setup uncertainty via in-room lasers. After phantom setup via lasers, acquisition of 3D MR scans and registration with the reference CTs of the treatment plans, all translational correction values for all directions separately were < 0.5 mm for all three phantoms. Since these offset values were below the couch motor drive accuracy, we considered them negligible and did not correct for them.

Polymer gel dose linearity

Dose to R_2 density plots of all PG measurements are shown in Figure 3. All density plots exhibit a similar shape and show a narrow distribution of values. The coefficient of determination (R^2) was above 0.99 for all linear fit functions.

Dose profiles and dose differences

Dose profiles in the anterior-posterior (AP), left-right (LR) and superior-inferior (SI) directions and their locations are shown in Figure 4. Both measured PG samples show good reproducibility. The N2 and especially the N1 profiles show slightly higher values compared to the TPS in most parts. Looking at the profiles, the largest deviations can be found in the mid dose range at around 4.0 Gy. Especially in this

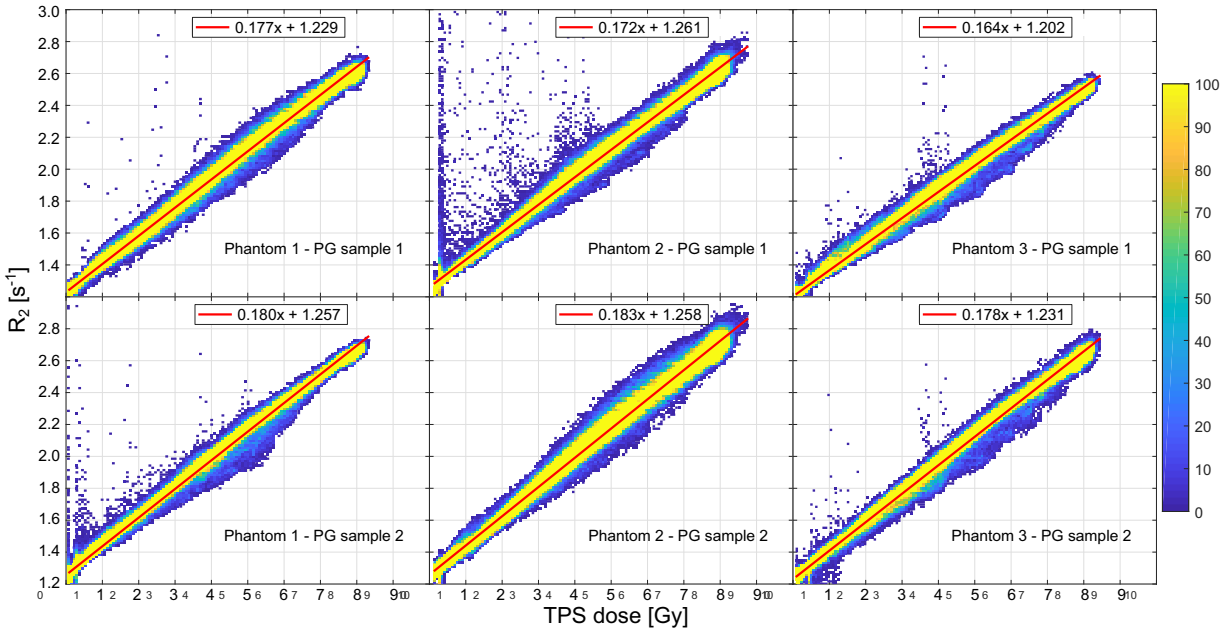

Figure 3. Density plots of R_2 vs. TPS dose of all voxels within the PG structure for the three phantoms and two measurements. Linear fits to the data are shown in red. The colour bar on the right indicates the number of counts per bin of the density plots. The R_2 values are not corrected for baseline signal.

Table 4
Mean absolute (abs.) and relative (rel.) dose differences between measured PG doses (N1 and N2) and the TPS reference doses across all phantoms for two dose regions: 40% - 60% of prescribed dose (PD) and > 80% of prescribed dose.

	N1 abs.	N1 rel.	N2 abs.	N2 rel.
> 80% of PD	0.21 Gy	2.9%	0.07 Gy	1.0%
40% - 60% of PD	0.23 Gy	5.6%	0.15 Gy	3.7%

dose region, the measured (N1 and N2) dose profiles show systematically higher values compared to the TPS. Overall, the measured profiles coincide well with the reference (TPS). Table 4 shows the mean dose deviations between PG measurements and the TPS doses for the mid (40% - 60% of prescribed dose) and high (> 80% of prescribed dose) dose ranges. Higher absolute and relative mean deviations are found in the mid dose range.

Centres of mass of prescribed dose volumes

The COM offsets between the measured and planned 95% isodose volumes are given in Table 5. The mean and median offsets across all samples were 1.2 mm and 1.0 mm, respectively.

Dose volume histograms

Figure 5 A-C show the cumulative DVHs of all measured PG samples (N2). The samples 1 and 2 of each phantom

Table 5
COM offsets between the binarised 95% isodose volumes of the PG measurement and the planned dose in anterior-posterior (AP), left-right (LR) and superior-inferior (SI) directions and the total offset for all three phantoms, which are defined as the Euclidean length of the offset vectors. Values are given for both measurements (measurement 1 / measurement 2).

Phantom No.	1	2	3
AP [mm]	0.8 / 0.1	1.0 / 1.5	-0.4 / 0.2
LR [mm]	-0.1 / -0.2	0.1 / -0.7	0.3 / 0.2
SI [mm]	0.8 / 0.2	1.6 / 1.3	0.8 / 0.5
Total [mm]	1.1 / 0.3	1.9 / 2.1	0.9 / 0.5

coincide well. Furthermore, the measured doses agree closely with the planned doses in phantom 1 and 3. A larger deviation can be seen for the brainstem-like structure in phantom 2. After correcting the shift of the COM offset evaluation of phantom 2, the DVH of the brainstem-like structure improves notably (Figure 5 D).

3D gamma analysis

An example gamma analysis result is shown in Figure 6 for the first sample of phantom 2 (without COM shift correction). Table 6 shows the results of the 3D gamma comparisons. Passing rates for N1 were systematically lower compared to N2 when comparing measured and planned doses. Phantom 2 showed overall slightly reduced passing rates in comparison to the other phantoms. When comparing the two measured

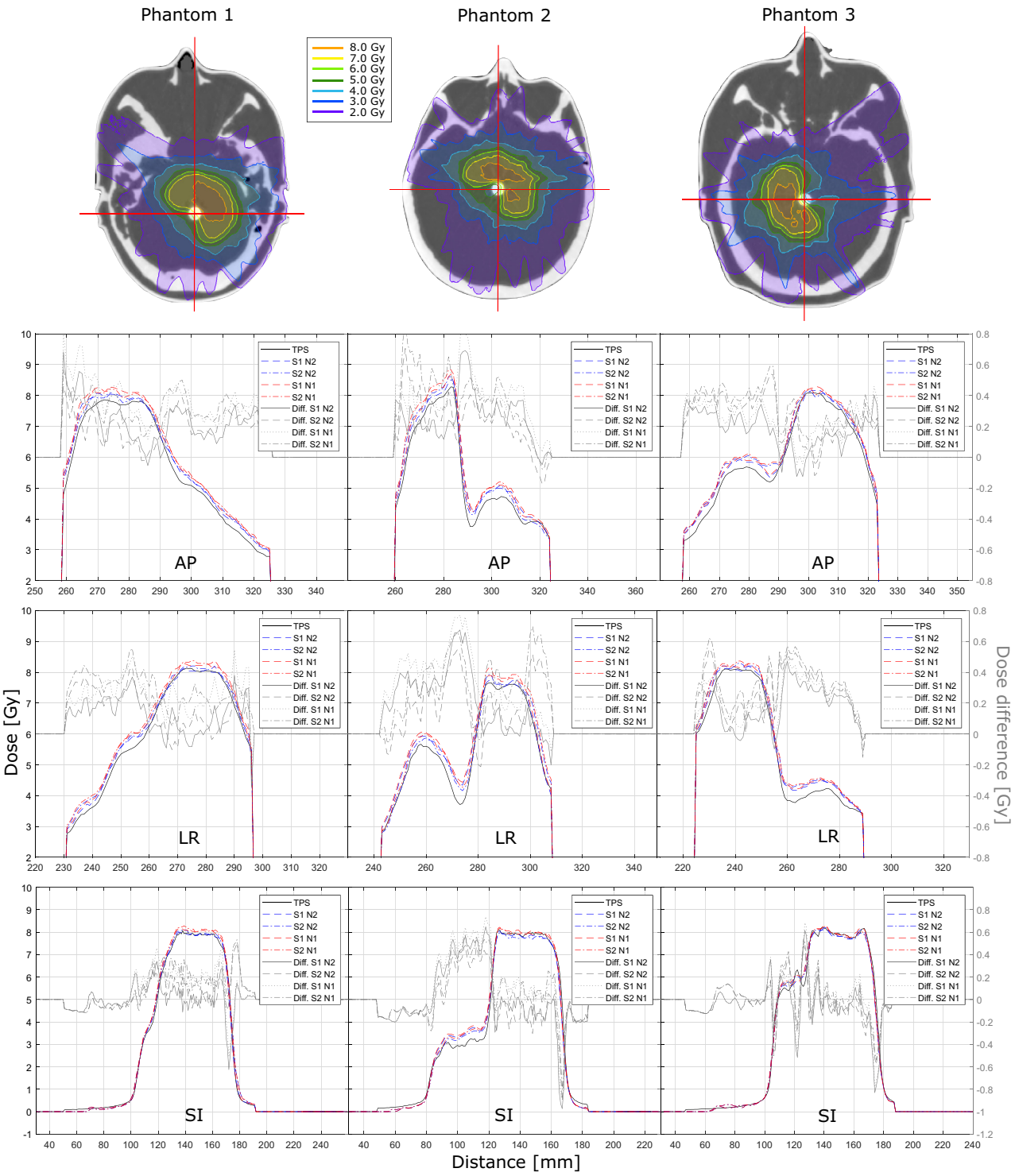

Figure 4. Dose profiles in anterior-posterior (AP), left-right (LR) and superior-inferior (SI) direction of all three head phantoms after application of a $3 \times 3 \times 3$ mm^3 median filter: TPS (black), normalised dose 2 (N2, planned PTV dose) of the PG samples S1 and S2 (dashed blue) and normalised dose 1 (N1, measured IC dose) of samples S1 and S2 (dashed red). The corresponding dose difference profiles are shown in gray. The abrupt fall-off on the edges is due to the binary masking of dose volumes within the PG structure, which was contoured in the TPS. The red crosshairs on the axial slices in the top row indicate the location of the profiles, which were taken through the centre of the ionisation chamber insert in the central slice of the overlap region between the PTV and the brainstem structure. The overlay isodose lines show the TPS dose, calculated on the head phantoms with the polymer gel inserts with density overrides of air bubbles and the polymer gel volumes.

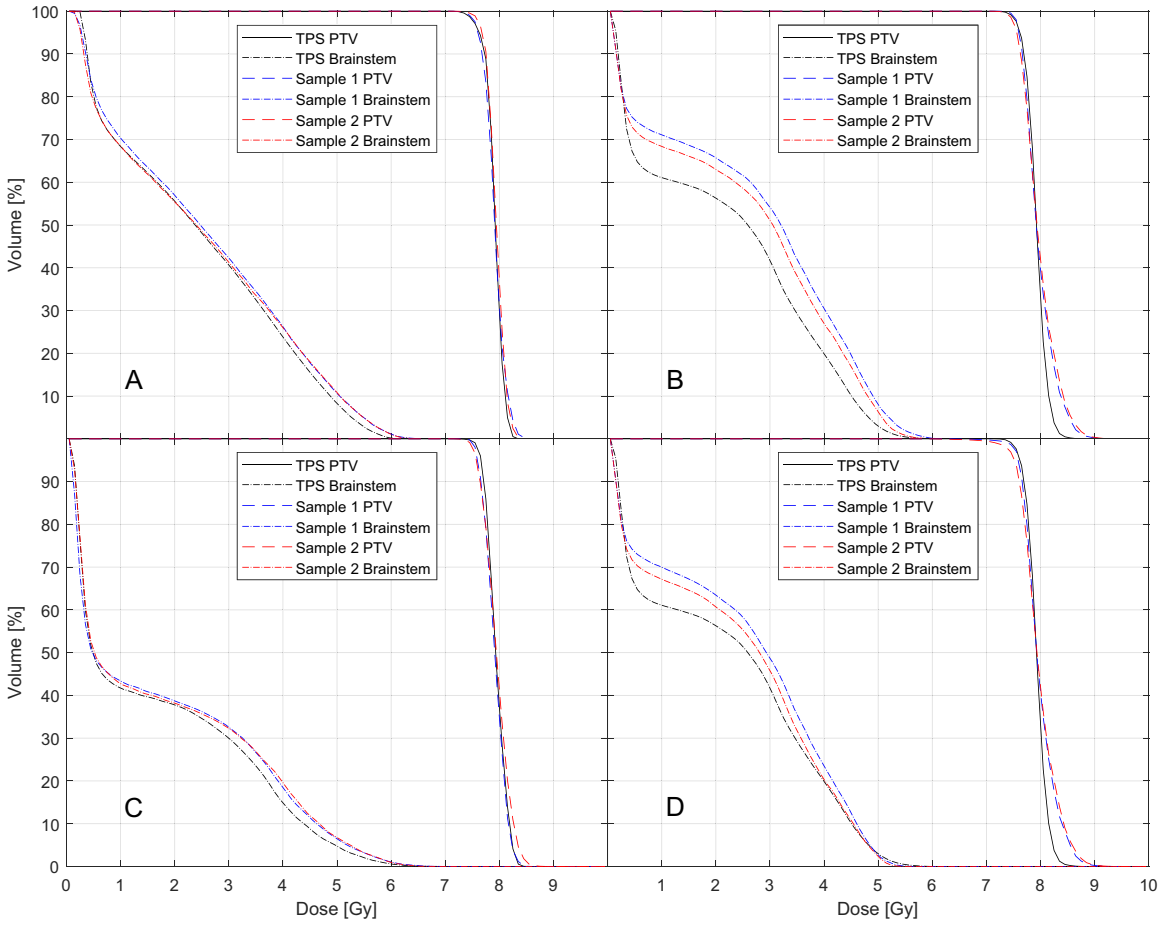

Figure 5. Cumulative dose volume histograms of the measured N2 PG doses of both samples S1 and S2 and the planned dose for the brainstem-like and the PTV structures of (A) phantom 1, (B) phantom 2, (C) phantom 3 and (D) the DVH of phantom 2 after correction of the COM offsets from Table 5.

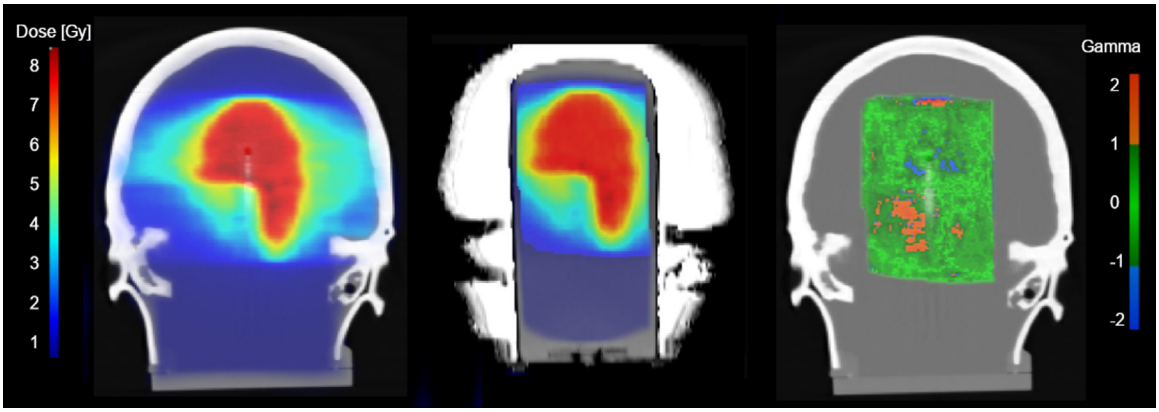

Figure 6. Exemplary visualisation of the 3D gamma analysis in a single coronal slice of phantom 2, sample 1. (left) TPS dose, (middle) PG measured N1, (right) gamma map.

Table 6
Percent gamma passing rates for different dose comparisons and gamma criteria. PG measured doses were compared to the planned doses and the PG measured doses of the two samples (s.) of each phantom (ph.) were compared to each other. N1 = normalised dose 1 (IC measurement); N2 = normalised dose 2 (planned PTV dose); e.g. Ph. 2 s 1 = phantom 2, sample 1.

Dose 1	Dose 2	2%, 2 mm N1	3%, 3 mm N1	2%, 2 mm N2	3%, 3 mm N2
Ph. 1 s. 1	TPS	79.9	95.8	93.5	99.2
Ph. 1 s. 2	TPS	79.3	97.4	93.3	99.1
Ph. 2 s. 1	TPS	83.0	94.2	90.5	97.6
Ph. 2 s. 2	TPS	76.2	92.1	85.5	97.0
Ph. 3 s. 1	TPS	90.4	98.4	93.2	98.9
Ph. 3 s. 2	TPS	90.7	98.6	93.5	99.0
	Mean:	83.3	96.1	91.6	98.5
Ph. 1 s. 1	Ph. 1 s. 2	96.5	99.7	96.4	99.7
Ph. 2 s. 1	Ph. 2 s. 2	95.3	99.6	95.9	99.7
Ph. 3 s. 1	Ph. 3 s. 2	95.2	99.2	95.2	99.2
	Mean:	95.7	99.5	95.8	99.5

samples against each other, all gamma passing rates were > 95.0%, even for the strict 2%, 2 mm criterion.

Discussion

Plan validation with diode array and ionisation chamber

The diode array plan validation showed acceptable results for all three plans, which means all three plans would have passed our institutional SOP criterion of 90% of pixels passing. However, the plan of phantom 2 clearly showed a worse gamma passing rate compared to the treatment plans of phantoms 1 and 3. Thus we would expect plan 2 to perform worse in the PG evaluation compared to the other plans. The absolute dose IC values complied well with the TPS for all phantoms, especially when considering the inhomogeneous anthropomorphic phantom geometry and the complexity of the treatment plans. Plan scaling from 2 Gy / fraction to 8 Gy / fraction was performed to limit irradiation time for the IC measurements and still exploit the full linear dose response range of the PG. Considering the step and shoot IMRT technique of the 0.35 T MR-linac, the only difference between the 2 Gy and 8 Gy up-scaled plans is the number of monitor units per multi-leaf collimator segment. This ensures dosimetric linearity between these two plans.

Polymer gel measurements

MRI systems with a magnetic field strength of 1.5 T (as opposed to the MR-linac system with 0.35 T) have been chosen for PG readout, because the signal-to-noise ratio (SNR) of one specific MRI sequence is directly proportional to the magnetic field strength B_0, and to the square root of the scanning time. Thus, for the same SNR substantially more averages (number of excitations) would be necessary at a 0.35 T MRI compared to a 1.5 T MRI, resulting in a prolonged scanning time by a factor of approximately 18. However, we did not explicitly test if a lower SNR and/or spatial resolution would still be sufficient for the PG 3D dose evaluation and if this could result in acceptable scanning times on the 0.35 T MRI. The MR readout sequence was chosen because of its good SNR gain. In theory, the varying T2-weighting in k-space of this sequence (by sequential read out of the k-space lines in phase encoding direction) could slightly influence the image resolution (slight blurring along phase-encoding direction in regions of very high spatial frequencies), but no such effect was visually perceivable on the MRI images. Furthermore, since no very high spatial frequencies are contained within the dose distribution (e.g. limited steepness of dose gradient due to scatter dose), no relevant influence of this effect can be expected on the results of the 3D dose evaluation.

When looking at the density plots (Figure 3), we can conclude that no serious systematic errors were made during the PG measurements. In fact, most data points lie within a small band around the linear fit functions. A few outliers can be seen in Figure 3, especially for phantom 2, sample 1. These outliers stem from voxels of the R_2 map, which lie (at least partially) inside the glass wall of the PG containing cylinder. This is the case because the PG contour, which was created during treatment planning, extends to the glass wall. Only a very small portion (< 0.3%) of all voxels within the PG volume are affected and thus no relevant disturbance introduced by these outliers can be expected regarding the dose evaluation. Overall, the narrow distribution of data points around the linear fit functions provides evidence that phantom positioning as well as image registration between the measured R_2-maps and the planned reference doses were good. Furthermore, the high R^2 values of all linear fit functions indicate an acceptable dose response linearity of all three PG batches. A slight non-linearity in dose response is inherent in polymer gel dosimetry. The effect of non-linearities in dose response of PG can be reduced when performing an additional PG calibration for each batch, however this is not the procedure recommended by the phantom manufacturer [29]. Since results are still good without extra dose calibration in the given dose

range, we consider the approach described here reliable and sufficient.

When evaluating the dose profiles (Figure 4), we can conclude that both normalisation approaches (N1 and N2) show good agreement between measured and planned doses. Larger deviations can be seen in the mid dose range at around 4.0 Gy. The same tendency can be seen when looking at the mean deviations between measured PG doses and the TPS reference doses (Table 4). The systemically larger absolute and relative deviations in the mid dose regions compared to the high dose regions indicate a slight systematic error, which might be a result of the normalisation approach. This systematic error could be caused by a slight non-linearity / saturation effect in the dose response of the PG, which was already described above. Since the normalisation is obtained in the high dose region, this would result in an increase in the measured normalised dose if one assumes a linear response of the PG. Although the density plots (Figure 3) and the high R^2 values do not indicate substantial non-linearities, a small saturation effect might remain unnoticed. A conventional polymer gel calibration procedure using uniformly irradiated gel samples could possibly further reduce these deviations, since the influence of the potentially slightly non-linear response of the PG could be corrected for. One weakness of this study is, that no direct comparison was made between the approach described in this paper and a conventional polymer gel calibration. This could be investigated in further studies. Overall, high similarity between the profiles of the two samples of each phantom was found, which indicates reproducible and robust PG fabrication, phantom setup, plan application, dose readout, image registration and dose evaluation.

Geometric offset errors and DVHs

Some small position offsets can be seen for some profiles and are attributed to the phantom setup and image registration uncertainties. Given a treatment table drive motor accuracy of 0.5 mm and the setup uncertainty with lasers of 0.5 mm, a setup uncertainty of 0.7 mm can be assumed. Image registration uncertainty can be assumed to be as much as 1.5 mm when considering the resolution of the CT and dose readout MR image. This results in an overall uncertainty of about 1.7 mm, which correlates well with the mean offset between the COMs of the measured and planned PTV-covering doses of 1.2 mm. Considering the in-plane resolution (0.7 mm) and slice thickness (2.5 mm) of the readout MR scans, most COM offsets are accurate to within this resolution. Both samples of phantom 2 showed relatively large total offsets of 1.9 mm and 2.1 mm, which may explain the poorer brainstem DVH agreement for this phantom. It can be assumed that two factors play a major role for this behaviour. When bearing in mind the inferior performance of the treatment plan of phantom 2 during diode array plan validation, it was expected that phantom 2 would also perform less well in the PG evaluation. However, the brainstem discrepancy cannot be solely explained by poorer plan QA performance. Since the dose gradient is rather steep at the boundary of this structure, setup deviations can also have a major impact on the DVH of the same structure. To demonstrate this, the COM offsets were applied to the measured dose distribution and a new DVH was generated (Figure 5 D), which clearly shows improved results (compared to the DVH of the same phantom 2; Figure 5 B) for the brainstem-like structure, compared to the original DVH of phantom 2. Thus, relatively small shifts can have a notable influence on the DVH of structures close to or within steep dose gradients, as expected. Non-negligible dose deviations can still be seen for the brainstem-like structure (Figure 5 D), even after correction for the COM offset. This is especially the case in the lower dose region. A reason for this might be the inferior performance of the treatment plan validation for phantom 2 using the diode array, which was mentioned above.

3D gamma comparison

The 3D gamma comparison showed good results, even for the N1 dose distributions with a mean of 83.3% for the strict 2%, 2 mm criterion and 96.1% for a 3%, 3 mm criterion. The similarity between the two samples of each phantom was very high with all gamma passing rates exceeding 95.0% for a 2%, 2 mm criterion. This is in accordance with the dose profiles and DVHs and provides high level evidence not only of the overall integrity of the MR-linac system but also of the good reproducibility of the PG measurements and the suitability of both normalisation approaches.

A limitation of the study is that no plan adaptation was performed on the phantoms. Nevertheless, from a QA point of view, when using a full re-optimisation oMRgRT approach, the only difference between the adapted and non-adapted plans is the propagation of CT numbers of the planning CT to the current MR via deformable image registration. Although this does indeed affect the overall uncertainty due to image registration uncertainties, the dosimetric performance of the linac is unchanged. There are no technical or conceptual limitations, which would limit this approach only to non-adaptive workflows.

Summary

The scope of this study was to validate the suitability of the novel phantom for use at a 0.35 T MR-linac and quantify the reproducibility and reliability, which was achieved by repeating the PG measurement of each phantom twice. In contrast to Pappas et al., who used two comparable head phantoms, one for PG and one for gafchromic film measurements, we used PG and IC inserts with the same phantom, eliminating the influence of potential disparities in phantom geometry and allowing us to normalise the relative 3D dose obtained with PG to the absolute IC reference dose. This technique is only possible with dedicated phantoms with exchangeable inserts.

Furthermore, we used a $1 \times 1 \times 1$ mm^3 dose calculation grid for improved dosimetric accuracy.

Dose readout of the PG inserts was performed at diagnostic MR scanners, 24 h post irradiation, which allowed the PG to further polymerise and give a more stable MR signal. Dose readout directly in position at the MR-linac after irradiation would be desirable but we do not consider this ideal due to the required extensive scanning time at 0.35 T and the time-dependent stabilisation process of the PG. Furthermore, the quality of MR components of current low-field MR-linacs is generally not approaching that of diagnostic-grade scanners, due to the requirements of linac coupling, which may lead to suboptimal dose read-out.

Compared to 2D diode array measurements, the 3D dose volume obtained with PG measurements provides more information and allows for a more in-depth insight into the MR-linac system. The commercial availability of the phantoms, in combination with the possibility to re-order PG filled inserts and the MR imaging capabilities, make the phantoms used here also interesting for routine QA (e.g. end-to-end test) at an MR-linac.

Conclusion

A novel hybrid anthropomorphic phantom was evaluated for use at a 0.35 T MR-linac. Good agreement was obtained in the high dose region used for normalisation, however a mean error of 0.2 Gy (5.6%) was observed in the intermediate dose region (3.2 Gy – 4.8 Gy) for normalisation N1 (IC measured dose). The modular design of the 3D printed anthropomorphic phantoms allowed us to normalise the R_2-map to the IC absolute dose reference. The obtained 3D dose allows for extensive dose analysis including OARs and PTV DVH. The dosimetric and spatial accuracy of the MR-linac system was quantified for three similar cases in the cranial region. Reproducibility and stability of PG dose measurements was validated by repeated measurements of each phantom using PG from different production batches and readout on two separate 1.5 T diagnostic MR scanners, which also helped to keep scanning times manageable. Generally, high spatial accuracy was found for the complex, non-adapted cases. Small setup deviations were found to potentially compromise the DVH of OAR close to or within steep dose gradients.

Declarations

Funding

The head phantoms and PG inserts were funded by FöFoLe commission of the Medical Faculty of the LMU Munich [grant number 994, F. Kamp]. The IC insert was funded by the DFG cluster of excellence MAP. The Department of Radiation Oncology of the University Hospital of LMU Munich has research agreements with Brainlab, Elekta and ViewRay. This project was outside the scope of these agreements. ViewRay was not involved and had no influence on the study design, the collection, analysis and interpretation of data, on the writing of the manuscript, or the decision to submit the manuscript for publication.

Declaration of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Consent for publication

Not applicable

Availability of data and supporting materials

Not applicable

Code availability

Not applicable

Authors' contributions

Conceptualisation: LN, CK, GL; methodology: LN, FK, MR, SC, MRA, KP, CB, CK, GL; preparation, treatment planning, plan validation and IC measurements: LN; PG measurements: LN, CK, GL; dose readout MR acquisition: OD, LN, CK, GL; data analysis: LN; manuscript draft: LN; manuscript review and editing and supervision: CK, GL. All authors read and approved the manuscript.

Acknowledgements

We thank Emmanouil Zoros and colleagues from RT-Safe for providing the PG filled cylinders and for generating the T_2 maps. Mark Podesta is thanked for sharing the gamma evaluation code.

References

[1] Raaymakers BW, Jürgenliemk-Schulz IM, Bol GH, Glitzner M, Kotte ANTJ, van Asselen B, et al. First patients treated with a 1.5 T MRI-Linac: clinical proof of concept of a high-precision, high-field MRI guided radiotherapy treatment. Phys Med Biol 2017;62:L41–50, http://dx.doi.org/10.1088/1361-6560/aa9517.

[2] Rudra S, Jiang N, Rosenberg SA, Olsen JR, Roach MC, Wan L, et al. Using adaptive magnetic resonance image-guided radiation therapy for treatment of inoperable pancreatic cancer. Cancer medicine 2019;8:2123–32, http://dx.doi.org/10.1002/cam4.2100.

[3] Placidi L, Romano A, Chiloiro G, Cusumano D, Boldrini L, Cellini F, et al. On-line adaptive MR guided radiotherapy for locally advanced pancreatic cancer: Clinical and dosimetric considerations. Technical innovations & patient support in radiation oncology 2020;15:15–21, http://dx.doi.org/10.1016/j.tipsro.2020.06.001.

[4] Corradini S, Alongi F, Andratschke N, Belka C, Boldrini L, Cellini F, et al. MR-guidance in clinical reality: current treatment challenges and future perspectives. Radiat Oncol 2019;14:92, http://dx.doi.org/10.1186/s13014-019-1308-y.
[5] Acharya S, Fischer-Valuck BW, Kashani R, Parikh P, Yang D, Zhao T, et al. Online Magnetic Resonance Image Guided Adaptive Radiation Therapy: First Clinical Applications. Int J Radiat Oncol Biol Phys 2016;94:394–403, http://dx.doi.org/10.1016/j.ijrobp.2015.10.015.
[6] Fischer-Valuck BW, Henke L, Green O, Kashani R, Acharya S, Bradley JD, et al. Two-and-a-half-year clinical experience with the world's first magnetic resonance image guided radiation therapy system. Adv Radiat Oncol 2017;2:485–93, http://dx.doi.org/10.1016/j.adro.2017.05.006.
[7] Paulson ES, Ahunbay E, Chen X, Mickevicius NJ, Chen GP, Schultz C, et al. 4D-MRI driven MR-guided online adaptive radiotherapy for abdominal stereotactic body radiation therapy on a high field MR-Linac: Implementation and initial clinical experience. Clinical and translational radiation oncology 2020;23:72–9, http://dx.doi.org/10.1016/j.ctro.2020.05.002.
[8] Tetar SU, Bruynzeel AM, Lagerwaard FJ, Slotman BJ, Bohoudi O, Palacios MA. Clinical implementation of magnetic resonance imaging guided adaptive radiotherapy for localized prostate cancer. Physics and Imaging in Radiation Oncology 2019;9:69–76, http://dx.doi.org/10.1016/j.phro.2019.02.002.
[9] Henke LE, Kashani R, Hilliard J, DeWees TA, Curcuru A, Przybysz D, et al. In Silico Trial of MR-Guided Midtreatment Adaptive Planning for Hypofractionated Stereotactic Radiation Therapy in Centrally Located Thoracic Tumors. Int J Radiat Oncol Biol Phys 2018;102:987–95, http://dx.doi.org/10.1016/j.ijrobp.2018.06.022.
[10] Henke LE, Olsen JR, Contreras JA, Curcuru A, DeWees TA, Green OL, et al. Stereotactic MR-Guided Online Adaptive Radiation Therapy (SMART) for Ultracentral Thorax Malignancies: Results of a Phase 1 Trial. Adv Radiat Oncol 2019;4:201–9, http://dx.doi.org/10.1016/j.adro.2018.10.003.
[11] Mittauer KE, Hill PM, Geurts MW, Costa A-M, de, Kimple RJ, Bassetti MF, Bayouth JE. STAT-ART: The Promise and Practice of a Rapid Palliative Single Session of MR-Guided Online Adaptive Radiotherapy (ART). Frontiers in oncology 2019;9:1013, http://dx.doi.org/10.3389/fonc.2019.01013.
[12] Palacios MA, Bohoudi O, Bruynzeel AM, van Sörsen de Koste JR, Cobussen P, Slotman BJ, et al. Role of Daily Plan Adaptation in MR-Guided Stereotactic Ablative Radiation Therapy for Adrenal Metastases. International Journal of Radiation Oncology*Biology*Physics 2018;102:426–33, http://dx.doi.org/10.1016/j.ijrobp.2018.06.002.
[13] Klüter S. Technical design and concept of a 0.35 T MR-Linac. Clinical and translational radiation oncology 2019;18:98–101, http://dx.doi.org/10.1016/j.ctro.2019.04.007.
[14] Kurz C, Buizza G, Landry G, Kamp F, Rabe M, Paganelli C, et al. Medical physics challenges in clinical MR-guided radiotherapy. Radiat Oncol 2020;15:93, http://dx.doi.org/10.1186/s13014-020-01524-4.
[15] Dorsch S, Mann P, Elter A, Runz A, Klüter S, Karger CP. Polymer gel-based measurements of the isocenter accuracy in an MR-LINAC. J. Phys.: Conf. Ser 2019;1305:12007, http://dx.doi.org/10.1088/1742-6596/1305/1/012007.
[16] Baldock C, Karger CP, Zaidi H. Gel dosimetry provides the optimal end-to-end quality assurance dosimetry for MR-linacs. Med Phys 2020;47:3259–62, http://dx.doi.org/10.1002/mp.14239.
[17] Niepel KB, Stanislawski M, Wuerl M, Doerringer F, Pinto M, Dietrich O, et al. Animal tissue-based quantitative comparison of dual-energy CT to SPR conversion methods using high-resolution gel dosimetry. Phys Med Biol 2020, http://dx.doi.org/10.1088/1361-6560/abbd14.
[18] Hillbrand M, Landry G, Ebert S, Dedes G, Pappas E, Kalaitzakis G, et al. Gel dosimetry for three dimensional proton range measurements in anthropomorphic geometries. Z Med Phys 2019;29:162–72, http://dx.doi.org/10.1016/j.zemedi.2018.08.002.
[19] Papoutsaki M-V, Maris TG, Pappas E, Papadakis AE, Damilakis J. Dosimetric characteristics of a new polymer gel and their dependence on post-preparation and post-irradiation time: effect on X-ray beam profile measurements. Phys Med 2013;29:453–60, http://dx.doi.org/10.1016/j.ejmp.2013.01.003.
[20] Deene Yde. Essential characteristics of polymer gel dosimeters. J. Phys.: Conf. Ser 2004;3:34–57, http://dx.doi.org/10.1088/1742-6596/3/1/006.
[21] Crescenti RA, Scheib SG, Schneider U, Gianolini S. Introducing gel dosimetry in a clinical environment: customization of polymer gel composition and magnetic resonance imaging parameters used for 3D dose verifications in radiosurgery and intensity modulated radiotherapy. Med. Phys 2007;34:1286–97, http://dx.doi.org/10.1118/1.2712042.
[22] Baldock C, Deene Y, de, Doran S, Ibbott G, Jirasek A, Lepage M, et al. Polymer gel dosimetry. Phys Med Biol 2010;55:R1–63, http://dx.doi.org/10.1088/0031-9155/55/5/R01.
[23] Maynard E, Hilts M, Heath E, Jirasek A. Evaluation of accuracy and precision in polymer gel dosimetry. Med Phys 2017;44:736–46, http://dx.doi.org/10.1002/mp.12080.
[24] Baldock C, Murry P, Kron T. Uncertainty analysis in polymer gel dosimetry. Phys Med Biol 1999;44:N243–6, http://dx.doi.org/10.1088/0031-9155/44/11/402.
[25] Deene Y, de, Vandecasteele J. On the reliability of 3D gel dosimetry. J. Phys.: Conf. Ser 2013;444:12015, http://dx.doi.org/10.1088/1742-6596/444/1/012015.
[26] Keall P, Baldock C. A theoretical study of the radiological properties and water equivalence of Fricke and polymer gels used for radiation dosimetry. Australas Phys Eng Sci Med 1999;22:85–91.
[27] Lee HJ, Roed Y, Venkataraman S, Carroll M, Ibbott GS. Investigation of magnetic field effects on the dose-response of 3D dosimeters for magnetic resonance - image guided radiation therapy applications. Radiother Oncol 2017;125:426–32, http://dx.doi.org/10.1016/j.radonc.2017.08.027.
[28] Dorsch S, Mann P, Elter A, Runz A, Spindeldreier CK, Klüter S, Karger CP. Measurement of isocenter alignment accuracy and image distortion of an 0.35 T MR-Linac system. Phys Med Biol 2019;64:205011, http://dx.doi.org/10.1088/1361-6560/ab4540.
[29] Pappas E, Kalaitzakis G, Boursianis T, Zoros E, Zourari K, Pappas EP, et al. Dosimetric performance of the Elekta Unity MR-linac system: 2D and 3D dosimetry in anthropomorphic inhomogeneous geometry. Phys Med Biol 2019;64:225009, http://dx.doi.org/10.1088/1361-6560/ab52ce.
[30] Elter A, Dorsch S, Mann P, Runz A, Johnen W, Spindeldreier CK, et al. End-to-end test of an online adaptive treatment procedure in MR-guided radiotherapy using a phantom with anthropomorphic structures. Phys. Med. Biol 2019;64:225003, http://dx.doi.org/10.1088/1361-6560/ab4d8e.
[31] Hoffmans D, Niebuhr N, Bohoudi O, Pfaffenberger A, Palacios M. An end-to-end test for MR-guided online adaptive radiotherapy. Phys Med Biol 2020;65:125012, http://dx.doi.org/10.1088/1361-6560/ab8955.
[32] Axford A, Dikaios N, Roberts DA, Clark CH, Evans PM. An end-to-end assessment on the accuracy of adaptive radiotherapy in an MR-linac. Phys Med Biol 2021;66:55021, http://dx.doi.org/10.1088/1361-6560/abe053.
[33] Steinmann A, Alvarez P, Lee H, Court L, Stafford R, Sawakuchi G, et al. MRIgRT head and neck anthropomorphic QA phantom: Design, development, reproducibility, and feasibility study. Med Phys 2020;47:604–13, http://dx.doi.org/10.1002/mp.13951.
[34] Bernchou U, Christiansen RL, Bertelsen A, Tilly D, Riis HL, Jensen HR, et al. End-to-end validation of the geometric dose delivery performance of MR linac adaptive radiotherapy. Phys Med Biol 2021;66:45034, http://dx.doi.org/10.1088/1361-6560/abd3ed.
[35] Mittauer KE, Hill PM, Bassetti MF, Bayouth JE. Validation of an MR-guided online adaptive radiotherapy (MRgoART) program: Deformation accuracy in a heterogeneous, deformable, anthropomorphic phantom. Radiother Oncol 2020;146:97–109, http://dx.doi.org/10.1016/j.radonc.2020.02.012.
[36] Stark LS, Andratschke N, Baumgartl M, Bogowicz M, Chamberlain M, Dal Bello R, et al. Dosimetric and geometric

end-to-end accuracy of a magnetic resonance guided linear accelerator. Physics and Imaging in Radiation Oncology 2020;16:109–12, http://dx.doi.org/10.1016/j.phro.2020.09.013.

[37] Sebastian Neppl. Measurement-based range evaluation for quality assurance of CBCT-based dose calculations in adaptive proton therapy. Med. Phys. in press. doi:10.1002/mp.14995.

[38] Makris DN, Pappas EP, Zoros E, Papanikolaou N, Saenz DL, Kalaitzakis G, et al. Characterization of a novel 3D printed patient specific phantom for quality assurance in cranial stereotactic radiosurgery applications. Phys Med Biol 2019;64:105009, http://dx.doi.org/10.1088/1361-6560/ab1758.

[39] Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. Med Phys 1998;25:656–61, http://dx.doi.org/10.1118/1.598248.

[40] Pappas E, Maris T, Angelopoulos A, Paparigopoulou M, Sakelliou L, Sandilos P, et al. A new polymer gel for magnetic resonance imaging (MRI) radiation dosimetry. Phys Med Biol 1999;44:2677–84, http://dx.doi.org/10.1088/0031-9155/44/10/320.

[41] Deene Y, de, Baldock C. Optimization of multiple spin-echo sequences for 3D polymer gel dosimetry. Phys Med Biol 2002;47:3117–41, http://dx.doi.org/10.1088/0031-9155/47/17/306.

[42] Gregory C, Sharp, Rui Li, John Wolfgang, George TY, Chen, Marta Peroni, Maria Francesca Spadea, Shinichro Mori, Junan Zhang, James Shackleford, Nagarajan, Kandasamy. Plastimatch - An Open Source Software Suite for Radiotherapy Image Processing. Proceedings of the XVIth ICCR 2010.

[43] Wolf I, Vetter M, Wegner I, Nolden M, Bottger T, Hastenteufel M, et al. The medical imaging interaction toolkit (MITK): a toolkit facilitating the creation of interactive software by extending VTK and ITK. In: Galloway JL, editor. Medical Imaging 2004; Saturday 14 February 2004;. San Diego: CA: SPIE;; 2004. p. 16, http://dx.doi.org/10.1117/12.535112.

[44] Ezzell GA, Burmeister JW, Dogan N, LoSasso TJ, Mechalakos JG, Mihailidis D, et al. IMRT commissioning: multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. Med Phys 2009;36:5359–73, http://dx.doi.org/10.1118/1.3238104.

[45] Song J-H, Kim M-J, Park S-H, Lee S-R, Lee M-Y, Lee DS, Suh TS. Gamma analysis dependence on specified low-dose thresholds for VMAT QA. J Appl Clin Med Phys 2015;16:263–72, http://dx.doi.org/10.1120/jacmp.v16i6.5696.