1. Introduction

In order to exploit the superior soft tissue contrast exhibited in magnetic resonance (MR) images with respect to computerized tomography (CT), MR-linac or MR-60Co integrated systems have been introduced in clinical practice (Mutic and Dempsey 2014, Raaymakers et al 2017, Pathmanathan et al 2018), enabling online MR-guided radiotherapy. The main advantage of such hybrid systems is the option for online adaptive re-planning and re-optimization (Lim-Reinders et al 2017, Ruschin et al 2017, Winkel et al 2019a), if suggested by potential...
tumor or normal tissue changes. A direct visualization of the structures of interest at the treatment position with superior soft tissue contrast, also allows for robust patient re-positioning or, equivalently, plan adjustment to patient position (Ruschin et al 2017). By exploiting the capability of real-time monitoring of structures during treatment, other attractive features include intra-fractional motion detection and delivery gating (McPartlin et al 2016, Menten et al 2016, Pathmanathan et al 2018).

The Elekta Unity MR-linac (Elekta AB, Stockholm, Sweden), referred as Unity in the following for brevity, comprises an Achieva 1.5T MR scanner (Philips Healthcare, Best, The Netherlands) and a 7 MV flattening-filter-free linear accelerator with an Agility MLC configuration (Elekta AB, Stockholm, Sweden). Beam directions are always perpendicular to the static magnetic field, which is parallel to the treatment couch (Woodings et al 2018a). The main drawback of hybrid MR-linac systems is the inability to translate or rotate the treatment couch within the MR bore. Therefore, patient set-up errors are accounted for by virtual couch shifts, i.e. the plan is adapted to the MR-identified patient position. Virtual couch shifts (or, equivalently, isocenter shifts) are calculated following a rigid or deformable spatial registration between the pre-treatment on-couch MR scan and the original planning CT scan (Winkel et al 2019a).

In Monaco v. 5.40.00 (Elekta AB, Stockholm, Sweden) treatment planning system (TPS), following calculation of the virtual couch shifts, adaptive planning can be performed by various methods, divided into two main categories (Winkel et al 2019a); the ‘adapt-to-position’ (ATP) and the ‘adapt-to-shape’ (ATS). The former strategy does not allow for contour editing and the daily MR scan is only rigidly registered to the original planning CT scan. ATS is more robust as it involves structure re-contouring and deformable registration. Depending on the strategy followed, plan optimization can involve segment shifts-only, segment aperture morphing, weights optimization from segment, weights optimization from fluence, etc (Lim-Reinders et al 2017, Winkel et al 2019a). Plan optimization methods were recently reviewed and compared (Winkel et al 2019a).

In any case, the original plan created is rarely delivered and instead an adapted plan is prepared online with the patient at the treatment position (Winkel et al 2019a). However, the adapted plan should not be considered a priori the same with the original in terms of target conformity and normal tissue involvement. For cranial intensity-modulated radiation therapy (IMRT) cases, a recent study has detected increased maximum doses at critical organs for the adapted plan with respect to the original one (Ruschin et al 2017). Depending on the plan optimization method used, the adapted plans often failed to meet the clinical dose constraints criteria (Winkel et al 2019a, 2019b). The main challenge arises from the fact that moving an MLC aperture across the non-flattened irradiation field might result in a very different fluence pattern (Lim-Reinders et al 2017).

With respect to dose delivery quality assurance (QA) and dosimetry in a Unity system, several challenges exist which are associated with the presence of the high magnetic field strength. The Lorentz force impacts on secondary electron paths resulting in a reduction of the build-up distance while increasing the dose near the proximal side of an air cavity, the so-called electron return effect (Tseng et al 2017, Costa et al 2018). With respect to reference dosimetry, correction factors (up to 8%) have been introduced to take into account the effect of the magnetic field on ionization chambers responses, which are also orientation dependent (Pojtinger et al 2018). Furthermore, regarding dosimetry within plastic phantoms, the presence of air cavities around ionization chambers also affects detector response (Hackett et al 2016, Agnew et al 2017, O’Brien and Sawakuchi 2017). Even in a water phantom, a diamond detector was shown to exhibit increased angular dependence in the presence of 1.5T magnetic field (Woodings et al 2018b). Film dosimetry seems to be the least affected. Minor (just above uncertainty levels) differences (Delfs et al 2018) in radiation induced optical density (OD) change have been reported. In terms of dose from the red channel, a zero effect by the presence of 1.5T was measured in a recent study (Billas et al 2019). In addition, film response was not affected by the orientation of monomer crystals (parallel to the short film side) with respect to the magnetic field direction (Delfs et al 2018).

The potential of using 3D dosimetry in conjunction with MR dose read-out has been repeatedly demonstrated (Lee et al 2017, 2018, Roed et al 2017). The attractive feature of dose read-out at the treatment delivery position using the 1.5T MR scanner enables 3D dosimetry without the uncertainty related to the additional spatial registration step between measured and calculated dose distributions. Regarding dosimetry, the dose-response characteristics for a set of three 3D dosimeters (including polymer gels) were not significantly affected by the presence of a 1.5T magnetic field, suggesting that no correction factors need to be applied for dose measurements in Unity (Lee et al 2017). More importantly, polymer gels act as both the detector and the tissue equivalent material and, subsequently, do not perturb the photon fluence by the beam (Baldock et al 2010). In contrast to using plastic phantoms, with this characteristic, introduction of unintentional small air cavities is also avoided, which could be significant in the presence of 1.5T magnetic field. Furthermore, real-time dose accumulation monitoring in 3D can only be performed in MR-linacs using such dosimetric systems (Lee et al 2018). More specifically, Lee et al performed real-time relative 3D dosimetry using Fricke-type gels and T1-weighted pulse sequences, within homogeneous geometries. The implementation of the well-established polymer gel dosimetry formulations in combination with T2 relaxometry (Pappas et al 1999, Pappas 2009, Baldock et al 2010) has not yet been evaluated.
This work focuses on the development and implementation of a methodology for the overall dosimetric evaluation of the Unity system, specifically addressing related challenges. The novelty of this study lies on both the phantoms employed and the dosimetric protocols followed. In particular, phantom designs feature anthropomorphic geometry, as well as tissue and bone equivalent materials which exhibit realistic contrast in both MR and CT imaging modalities. Because of the latter characteristic, the phantoms used are suitable for performing clinically realistic virtual couch shifting and plan adaptation prior to treatment delivery, simulating the actual clinical procedure. Furthermore, each phantom incorporates either a radiochromic film or is filled with 3D polymer gel, i.e. 2D and 3D dosimeters, respectively. Specific methodology provisions ensured (i) presence of realistic (in size and shape) bone inhomogeneities, (ii) fiducials of adequate MR/CT contrast for facilitating registration of film measurements to MR planning coordinate system, (iii) absolute 2D dose measurements, (iv) relative 3D dose measurements and, (v) real-time dose accumulation monitoring based on T2-weighted (T2w) images. The methodology is implemented for treatment plan verification of an intracranial IMRT case, following virtual couch shifting. Dose measurements are compared against both the original and the adapted plans to highlight any potential discrepancies.

2. Materials and methods

2.1. Phantoms characteristics
The PseudoPatient™ head phantoms (RTsafe P.C., Athens, Greece) were employed for the majority of irradiations described in this study. Specifically, two identical head phantoms were used, each one encompassing a different dosimetric system (figure 1). These phantoms were produced by the vendor using the methodology described in Makris et al (2019). Briefly, a patient’s planning CT scan is used to 3D-print the patient’s phantom in terms of external contour and cranial bone anatomy. The finalized 3D-printed material is bone equivalent in the high energy photon fields used herein (Makris et al 2019). The resulting hollow phantom can accommodate any kind of dosimetric system (e.g. ionization chamber, film or even 3D polymer gel) and has been previously used for dosimetry in several studies (Pantelis et al 2018, Saenz et al 2018, Hillbrand et al 2019). Regarding 2D dosimetry, the rest of the phantom is filled with water, serving as soft tissue equivalent. In the 3D dosimetry case, the gel acts as both the detector and the soft tissue material.

The reason for selecting these specific head phantoms is that they comprise: (i) detailed anthropomorphic geometry in terms of size and shape, (ii) involving bone inhomogeneities, (iii) multi-detector dosimetry capabilities, including the option for 3D measurements and, most importantly, (iv) clinically realistic and sufficient contrast in both MR and CT imaging modalities. The latter characteristic enables the option to perform virtual couch shifting and adaptive treatment planning at the set-up position (following MR imaging and spatial co-registration to the planning CT scan coordinate system), simulating the actual clinical case (Raaymakers et al 2017, Pathmanathan et al 2018). It is particularly important for a QA protocol to involve this step of the workflow which is a unique characteristic related to the Unity design.

2.2. Treatment planning, adaptive planning and dose delivery
The anonymized actual patient’s planning CT scan (corresponding to this particular set of phantoms) were used to prepare a challenging plan for the treatment of a hypothetical large C-shape brain lesion in Monaco v5.40.00. This TPS version features dose calculations in 1.5T magnetic fields for Unity applications. Planning target volume (PTV) was 27.506 cm³ in a C-shape, partly surrounding the brainstem (contoured volume 11.805 cm³) which was regarded as the organ-at-risk (OAR), as shown in figure 2. A step-and-shoot IMRT technique was employed consisting of seven flattening-filter-free 7 MV beams at gantry angles 51°, 102°, 180°, 153°, 204°, 255° and 306°, using one isocenter. Collimator angle was always at the Unity’s fixed configuration of 270°. Dose delivery levels were selected to match the dosimeters’ calibration or optimal performance ranges (see sections below). Therefore, a dose of 8 Gy was prescribed to cover 95% of the PTV. A 3 x 3 x 3 mm³ dose calculation grid spacing was used with 1% statistical uncertainty per calculation of dose to medium. The specific calculation grid size, although it provides relatively limited spatial resolution, was a compromise between accuracy and feasible calculation times in clinical practice aiming to simulate a realistic treatment workflow. The plan is illustrated in figure 2.

The created structures were exported in dicom format and shared with the phantoms’ manufacturer (RTsafe P.C., Athens, Greece) in order to install one dosimetry insert per phantom and more specifically: (i) create a socket for film dosimetry at a coronal plane intersecting the PTV (figures 1 (a) and (c)) and (ii) fill the entire brain area of the second phantom with a 3D polymer gel dosimeter (figures 1 (b) and (d)). Specific details for each detector and dosimetry protocol are given in the following sections.

Each phantom was positioned on the Unity treatment couch in order to perform the necessary virtual couch shifts and adaptive planning. All on-couch pre-treatment MR images of the two head phantoms were acquired by implementing a 3D T2w turbo spin echo (TSE) pulse sequence with echo time (TE), repetition time (TR) and flip angle (FA) of 500 ms, 2000 ms and 90°, respectively. This is the routinely used MR protocol for patient...
Figure 1. (a) The first PseudoPatient™ phantom incorporating an insert for film dosimetry placed on the Unity treatment couch. (b) The second PseudoPatient™ phantom, a cubic container and four vials, all prior to being filled with polymer gel 3D dosimeters from the same batch. (c) A coronal slice of the pre-treatment MR scan of the first phantom on the film plane, depicting the custom MR compatible fiducials (red circles) used for spatial registration. (d) An indicative axial slice of post-treatment MR scan of the second phantom, with darker areas corresponding to higher dose levels.

Figure 2. Screenshots of the Monaco TPS illustrating the created original plan. Axial (a), coronal (b) and sagittal (c) views of the planning CT image stack with dose distribution represented by the colormap. The PTV, OAR and external surface are depicted by the red, green and brown contours, respectively. (d) A 3D reconstruction is also shown.
positioning. Acquisition duration did not exceed 2 min. Subsequently, images were spatially registered to original planning CT image stack (reconstructed voxel size of 0.74 × 0.74 × 1.25 mm³) using the mutual information-based rigid body registration algorithm incorporated in Monaco. The registration procedure was based on the clinically realistic MR contrast of the bone and soft tissue structures of the entire phantom geometry. Specifically for the PseudoPatient™ phantoms employed in this study, dose calculations in the patient CT image stack were recently shown to be equivalent to patient-derived TPS calculations and in-phantom measurements, at least in the brain area (Makris et al. 2019). Necessary virtual couch shifts were calculated by the TPS and an adapted treatment plan was prepared with the same dose calculation grid resolution of 3 × 3 × 3 mm³, in order to keep calculation times at a practical minimum, from a clinical point of view. Since contour editing was not of particular interest, the ATP strategy was followed, using the segment weight optimization mode. This is the suggested plan adaptation method, if minimal or no anatomical changes are expected (Winkel et al. 2019a). No changes were manually made to the created adapted plan. Dose delivery was subsequently performed. The same procedure was followed for both head phantoms.

### 2.3. 2D dosimetry

The phantom, incorporating a film dosimetry insert, has been described in detail in a previous publication (Makris et al. 2019). Briefly, an axial area of 7.5 × 2 cm² from the inferior side of the phantom is cleared of bone structures, allowing entrance of a film dosimetry holder in the brainstem and central brain area, as shown in figure 1(a). In the created socket, a film cassette of 6.5 × 15 cm², made of large RW3 slabs (provided by PTW, Freiburg, Germany) is installed. The dosimetry film is stacked in-between RW3 slabs.

In the original phantom design (Makris et al. 2019), four metal pins lanced through the film for film stabilization among the slabs, as well as facilitating registration of measured and calculated dose distributions. Although this technique has been proven successful, resulting in accurately registered distributions (Pappas et al. 2017, Saenz et al. 2018, Makris et al. 2019), it has been adapted to involve MR compatible materials only. Novel fiducial markers, suitable for Unity, were designed and constructed. Three small hollow cylindrical inserts (inner diameter and height of 4 mm and 7 mm, respectively, wall thickness of 1 mm) were made of acrylic (also known as polymethyl methacrylate or PMMA) which does not yield signal in MR imaging (Pappas et al. 2016, 2018). The created cavities were filled with copper sulfate solution and sealed, providing adequate MR contrast, as shown in figure 1(c). Both materials are MR (and CT) compatible and often used in MRI-dedicated phantoms (Pappas et al. 2016, 2018).

Holes were drilled in the RW3 slabs in order to incorporate the cylinder-shape fiducials, while same holes were cut on the film dosimetry piece. Similar to the approach previously followed using metal pins, the geometric center of the fiducials is placed on the film plane. Matching MR-identified centers of the fiducials with corresponding positions of the holes in the scanned film image offers the necessary set of reference points, defining the rigid transformation matrix in order to spatially register film measurements to TPS calculations. Appropriate routines have been developed in MATLAB R2015b (The MathWorks, Inc., Natick, MA) and validated in previous studies (Pappas et al. 2017, Saenz et al. 2018, Makris et al. 2019).

Regarding film dosimetry, GAFchromic EBT3 (Ashland Inc., Wayne, NJ) films were used implementing the single channel protocol (Devic et al. 2005). The utilized film batch was calibrated at the Secondary Standard Dosimetry Laboratory of the Greek Atomic Energy Commission using a 60Co beam in a dose range of 0.10–15 Gy. Calibration film pieces of dimensions 3 × 3 cm² were irradiated at a depth of 5 cm in a solid water slab phantom. Dose read-out was performed 24 h after irradiation. The same methodology regarding film scanning parameters and image processing was followed, as described in Makris et al. 2019), using the EPSON Perfection V850 Pro flat-bed color scanner and in-house developed MATLAB routines. A polynomial calibration curve was obtained for the red color and was used for conversion of the net OD values into dose values (Makris et al. 2019).

### 2.4. 3D dosimetry

The polymer gel formulation characterized in Papoutsaki et al. (2013) was used, also produced by RTsafe P.C. (Athens, Greece). Prior to implementing the gel-filled head phantom, the dosimetric characteristics of the specific gel batch were assessed under Unity irradiation conditions. In order to verify dose-response linearity, a set of four small vials (shown in figure 1(b)) were filled with gel samples from the same batch. The vials were placed in a custom-made water tank, using an in-house developed PMMA holder, so that the central sagittal plane of each vial lied at a depth of 5 cm. Each vial was irradiated by Unity using a simple static beam of 10 × 10 cm² field size and gantry angle of 90°, so that the central sagittal plane of the respective irradiated vial received a uniform dose of either 0, 4, 8 or 12 Gy. The delivered MU's were calculated by the TPS. After applying the MR pulse sequence (described below), used for the dose read-out, the measured R2 relaxation rates of each vial were extracted by averaging the values of a central circular ROI at the central sagittal plane of each irradiated vial. Although this check can also provide a dose-response curve, obtained results were not applied to head phantom
measurements for dose determination, in order to avoid any potential dependencies, temporal variations of the signal and volume effects (Baldock et al 2010).

In a second independent check, a cubic container, made of PMMA (internal dimensions: $10 \times 10 \times 12$ cm$^3$, wall thickness: 1 cm, shown in figure 1(b)), was also filled with gel from the same batch and irradiated by a single $3 \times 3$ cm$^2$ static field, centered at the cube and delivering 12 Gy at the depth of maximum dose. The endpoint of this check was to verify spatial integrity and maturity of the MR signal related to the radiation induced polymerization, immediately after irradiation (i.e. the dose read-out timing considered for the head phantom). In parallel, this check allowed for validation of the MR pulse sequence parameters (described below) to be used for dose read-out (quantitative T2 maps) of the head phantom. For this test, reconstructed voxel size was $1.94 \times 1.94 \times 2$ mm$^3$.

The second head phantom was filled with gel from the same batch, covering the entire brain volume while the rest of the phantom was filled with non-dosimetric gel. After placing it on the Unity treatment couch, virtual couch shifts and adaptive planning were performed by Monaco, again the segment weights optimization method, under the ATP strategy (Winkel et al 2019a).

Regarding MR dose read-out, the head phantom was also scanned immediately after plan delivery. A 3D multi-echo turbo spin echo (TSE) pulse sequence was used, which consisted of 20 echoes with the first one at 25 ms and echo time intervals of 40 ms (TR, FA) $= (2000$ ms, $90^\circ$). In order to minimize MR-related geometric distortion, the selected bandwidth was 1184 Hz/pixel (Hillbrand et al 2019). To result in acceptable MRI acquisition duration, the scan length on the superior–inferior direction (i.e. $z$-axis) was limited to 9.2 cm, which did not extend to cover the entire phantom (16 cm), disregarding areas of very low dose. As a result, MR scanning time did not exceed 35 min for a reconstructed voxel size of $1.53 \times 1.53 \times 2$ mm$^3$. Reconstruction of the measured $R_2$ relaxation rate maps in 3D was performed using in-house MATLAB routines which have been verified in previous works (Saenz et al 2018, Hillbrand et al 2019). An axial plane of the obtained $T_2$ ($=1/R_2$) distribution, intersecting the PTV, is shown in figure 1(d).

In addition, the gel-filled head phantom was continuously MR-scanned prior, during treatment delivery. Specifically, a 2D $T_2w$ dynamic pulse sequence was employed to capture sagittal MR images intersecting the PTV, with a temporal resolution of 6 s and specific parameters (TE, TR, FA, Bandwidth) $= (500$ ms, $2000$ ms, $90^\circ$, $100$ Hz/pix). The in-plane reconstructed pixel size was $1.72 \times 1.72$ mm$^2$. This side test served as a feasibility study demonstrating the potential of real-time 3D dose accumulation monitoring based on $T_2w$ images, for the first time.

2.5. Analysis and comparison
Throughout this study, all created plans and calculated dose distributions (adapted and original ones) were exported in dicom format, spatially registered to the original planning CT scan dicom coordinate system by Monaco, following the procedure described above. This also applies to the acquired MR images.

All 2D and 3D measurements, images and dose calculations were analyzed in MATLAB using in-house routines. A 2D Wiener noise removal filter was applied to smooth film measurements. Noise reduction in the gel-related distributions was performed by using a 3D Gaussian filter.

Film measurements were not normalized and comparison with TPS predictions was performed in absolute doses. On the other hand, polymer gel 3D results were normalized to the mean net $R_2$ value in a central homogeneous 180 mm$^2$ area inside the PTV. The same area was used for normalizing the TPS-calculated dose distributions, prior to comparing with 3D measurements. This normalization approach was used in all subsequent calculations such as plan quality indices, dose volume histograms (DVHs), etc (see below).

Quantitative comparison between measured and calculated datasets involved the 3D gamma index (GI) test (Low and Dempsey 2003), following the interpolation recommendations in Hussein et al (2017). Measurements always served as the reference dose distribution against which the TPS calculations were evaluated. Selection of passing criteria was partly associated with estimated experimental uncertainties of each dosimetric system and protocol (see section 4). Therefore, distance-to-agreement (DTA) and dose difference (DD) were always $\geq 2$ mm and $\geq 2\%$, respectively.

With respect to the original and adapted plans and corresponding 3D relative dose measurements, datasets were further processed to derive and compare plan quality and plan conformity metrics, such as RTOG conformity index (RCI), Paddick’s conformity index (PCI), homogeneity index (HI) and quality of coverage (Q) (Paddick 2000, Feuvret et al 2006). DVHs and relevant dose-volume metrics (such as the mean dose ($D_{mean}$), the minimum dose delivered to the 95% of the structure volume ($D_{95\%}$), the maximum ($D_{max}$) and minimum ($D_{min}$) doses) were also determined for the PTV.

3. Results

3.1. 2D dosimetry
Following the positioning of the film head phantom and pre-treatment MR scanning, the calculated virtual couch shifts (i.e. isocenter shifts) were $(x, y, z) = (−0.1$ mm, $−33.8$ mm, $32.6$ mm) in the dicom coordinate.
system. It is worth noting that if the necessary isocenter shift exceeds 50 mm, the system does not allow for adaptive planning and patient/phantom repositioning is required. Therefore, this case was close to the limits of the allowed magnitude of virtual couch shifting, and thus a rather extreme case.

Figure 3 presents the dosimetry results related to film measurements. Selected measured and calculated isodose lines are superimposed on the obtained 3D GI maps, using passing criteria of 3%/3 mm and 3%/2 mm. However, a dose cut-off threshold of 20% of the maximum calculated dose was applied prior to GI calculations in order to exclude points of relatively low dose and, consequently, increased uncertainty (Aldelaijan et al 2011, Pappas et al 2017). In-plane spatial agreement between all distributions can be easily acknowledged. From a dosimetric point of view, although the overall results imply a very good agreement between measurements and calculations for both the adapted and the original plans, a small high-dose area inside the PTV can be noticed where failing points occur systematically. GI passing rates are given in table 1, including the more stringent 2%/2 mm criteria. In all cases, passing rates lie well above 90%.

### 3.2. 3D dosimetry

In figure 4(a), the measured R2 relaxation rates are presented against the dose delivered to the four vials filled with gel from the same batch. A first order polynomial is fitted to check the gel’s dose-response linearity, under the same time and irradiation conditions and in the presence of 1.5T magnetic field. Gel response can be regarded as linear at least up to 12 Gy, even immediately after irradiation.

Results for the signal integrity check are given in figures 4(b) and (c). Central crossline and inline profiles are depicted and compared with TPS calculations for a static 3 × 3 cm3 field. All profiles are normalized to their local maxima values. 1D GI tests resulted in 100% passing rates using criteria of 3%/3 mm, for both profiles, without applying any dose cut-off threshold.

After positioning the second PseudoPatient™ head phantom on the Unity treatment couch, the calculated virtual couch shifts were almost identical for the x and y axes (dicom coordinate system) but on the y axis were even larger by 2.7 mm, almost reaching the limits of the allowed radial isocenter shifting.

Head phantom 3D measurements are shown in figure 5 and are compared with both the adapted and the original calculated dose distributions. Relative isodoses are superimposed on the reconstructed T2 maps for an indicative axial and a sagittal slice intersecting the PTV. MR contrast also qualitatively represents relative dose measurements.

Agreement between measurements and calculations is quantified in table 2, where the 3D GI passing rates are given for a bounding box surrounding and including the entire PTV. A dose cut-off threshold of 20% was used in this analysis.

The 3D relative dose information, provided by the gel, is further exploited in figure 6 where the measured DVHs for the PTV and the OAR are plotted alongside the corresponding TPS-calculated DVHs. It should be noted that the normalization method described in section 2.5, was used. Although DVHs are in excellent agreement for the PTV, minor discrepancies can be noticed for the OAR, which could be attributed to the increased uncertainties in the low dose area found in the OAR.

Furthermore, several relative dose-volume metrics for the PTV were calculated and are presented in table 3. Despite that almost all values agree within 2.5%, the maximum dose in the PTV as calculated for the original plan is approximately 3% higher than the one in the adapted plan and ~4% higher than measured. This is the only considerable difference detected for the PTV. However, maximum measured dose is determined by a single voxel and is, therefore, subject to statistical uncertainties. Table 3 also presents plan quality indices commonly employed in routine clinical practice. Excellent agreement is also verified by these values. The detected discrepancy in maximum dose in the PTV is also reflected to the HI. It should be noted, however, that the reference isovole considered for calculating the plan quality indices matches the 70% isodose (with respect to the normalization dose), as it better conforms to the PTV shape (see figure 6 and D95% values in table 3).

As a side study, real-time dose accumulation monitoring was performed based on T2w dynamic imaging. A video showing dose deposition in a central sagittal slice is provided as supplementary material (stacks.iop.org/PMB/64/225009/mmedia). Gradually darkening areas correspond to T2 values being reduced, (i.e. higher R2 relaxation rates) and, consequently, increasing deposited dose levels. To ensure reasonable time duration, video time is in fast forward mode.

### 4. Discussion

Benchmarking and periodic QA testing in Unity systems pose dosimetric challenges related to the presence of strong magnetic fields. In response to that, correction factors for ionization chamber readings (Pojtinger et al 2018), specially designed scanning water phantoms (Smit et al 2014) and MR compatible versions of commercially available diode arrays (Houweling et al 2016, De Vries et al 2018) have been proposed. All these studies utilize active detectors in homogeneous phantoms of generic size and shape (cubic or cylindrical).
Furthermore, adaptive planning (if included) cannot follow the clinical workflow due to the lack of realistic MR signal and contrast. In the present study, passive dosimetry has been employed for plan verification in an anthropomorphic phantom geometry, including the presence of high density inhomogeneities. Conveniently enough, the MR compatibility of the materials employed also comes with adequate MR signal and contrast, allowing for clinically realistic virtual couch shift calculations and adaptive planning, important steps in the Unity treatment workflow.

The presented methodology combines a well-established 2D film dosimetry protocol for absolute dose measurements with polymer gel, which allows for on-couch 3D relative dose read-out. All end-to-end tests performed were completed within approximately 30 min (additional 30 min of MR scanning were required for the dose read-out in 3D dosimetry), allowing the implementation of the methodology for routinely performed QA tests.

![Figure 3](image)

**Figure 3.** GI maps comparing film measurements with the adapted (a) and (b) and the original (c) and (d) plans. Passing criteria used were 3%/3 mm (a) and (c) or 3%/2 mm (b) and (d). Measured (dashed black lines) and calculated (solid black lines) isolines (in Gy) are superimposed to assist comparison. White corresponds to areas where film measurements are either not available or dose levels lie below the low dose cut-off threshold of 20% of the maximum calculated dose. An arbitrary on-film coordinate system is adopted.

**Table 1.** 3D GI passing rates comparing film measurements with the adapted and original plans using various passing criteria. A low dose cut-off threshold of 20% of the maximum calculated dose has been applied.

| Plan   | Passing criteria | Passing rate |
|--------|------------------|--------------|
|        | DD (%) | DTA (mm) | GI ≤ 1 (%) |
| Adapted|3        |3        |97.8 |
|        |3        |2        |95.1 |
|        |2        |2        |93.1 |
| Original|3        |3        |98.2 |
|        |3        |2        |96.3 |
|        |2        |2        |93.6 |

Abbreviations: DD: dose difference, DTA: distance to agreement, GI: gamma index.

Furthermore, adaptive planning (if included) cannot follow the clinical workflow due to the lack of realistic MR signal and contrast.

In the present study, passive dosimetry has been employed for plan verification in an anthropomorphic phantom geometry, including the presence of high density inhomogeneities. Conveniently enough, the MR compatibility of the materials employed also comes with adequate MR signal and contrast, allowing for clinically realistic virtual couch shift calculations and adaptive planning, important steps in the Unity treatment workflow. The presented methodology combines a well-established 2D film dosimetry protocol for absolute dose measurements with polymer gel, which allows for on-couch 3D relative dose read-out. All end-to-end tests performed were completed within approximately 30 min (additional 30 min of MR scanning were required for the dose read-out in 3D dosimetry), allowing the implementation of the methodology for routinely performed QA tests,
Figure 4. (a) Linearity check results from the four gel-filled vials. A first order polynomial is fitted (solid blue line) to the measured values. 95% confidence bounds (dashed blue lines) are also depicted. (b) and (c) Signal integrity check results from the gel-filled cubic container. A crossline (b) and an inline (c) profile normalized at the local maxima compared against corresponding TPS calculations. Error bars of ±1 mm (corresponding to approximately ±1/2 of the pixel size) are also drawn.

Figure 5. Axial (a) and (c) and sagittal (b) and (d) central slices of the reconstructed T2 maps derived from the post-irradiation MR scan of the gel-filled phantom. Corresponding relative dose measurements are also superimposed (dashed isolines) to assist comparison with relative dose TPS calculations (solid isolines) related to the adapted (a) and (b) and the original (c) and (d) treatment plans. Legend: blue, turquoise, green and yellow isolines correspond to the 40%, 60%, 80% and 100% isodose, respectively. The PTV, OAR and external surface are depicted by the red, green and brown contours, respectively.

Table 2. 3D GI passing rates comparing gel measurements with the adapted and original plans using various passing criteria, within a volume of interest of a bounding box surrounding and including the entire PTV. A low dose cut-off threshold of 20% has been applied.

| Plan     | Passing criteria |                   | Passing rate |
|----------|------------------|-------------------|--------------|
|          | DD (%) | DTA (mm) | GI ≤ 1 (%) |
| Adapted  | 3      | 3        | 99.5        |
|          | 3      | 2        | 97.1        |
|          | 2      | 2        | 95.5        |
| Original | 3      | 3        | 98.4        |
|          | 3      | 2        | 95.1        |
|          | 2      | 2        | 93.5        |

Abbreviations: DD: dose difference, DTA: distance to agreement, GI: gamma index.
since it is not time consuming. In addition, real-time dose accumulation monitoring based on T2w images was presented for the first time, as a feasibility proof-of-concept study.

After performing the necessary virtual couch shift (i.e. isocenter shift) of several centimeters, the adapted plans comprised off-axis IMRT non-flattened beams in order to irradiate a relatively large PTV. This yields challenges in TPS calculations, let alone the presence of 1.5T magnetic field, resulting to the electron return effect and crossline profile offsets (Woodings et al 2018a), combined with material inhomogeneities. For these reasons, we consider the benchmarking end-to-end test performed in this study particularly rigorous.

Overall results suggest that measurements agree well with both the adapted and original plans. A limited area with GI failing points was detected on the film plane (figure 3) but was not verified by the gel measurements. Nevertheless, the high passing rates achieved using both dosimeters, in combination with DVHs comparison and plan quality metrics results, clearly indicate that TPS predictions were verified within the dosimetric uncertainties involved, passing criteria considered and limitations of the methodology followed.

As part of the daily dose fraction delivery, an adapted plan should result in a dose distribution as close as possible to the original one. However, it was shown that this is not always the case, with adapted plans often being clinically unacceptable (Winkel et al 2019a, 2019b). For extracranial cases, the effect was strongly associated with the adaptation method used and treatment site considered (Winkel et al 2019a). Specifically for cranial radiotherapy, smaller discrepancies have been reported (Ruschin et al 2017), mainly related to increased maximum doses in OARs (up to 2.0%). In accordance with the latter study, results of the present work also indirectly suggest that the original and adapted cranial radiotherapy plans did not exhibit any considerable dosimetric differences. This is clearly illustrated in figure 6 and tables 1–3. However, this work involved only one treatment plan and geometry without investigating the effect of different plan adaptation methods and, therefore, results are not sufficient for any general conclusions.

---

**Table 3.** TPS-calculated and gel-measured dose-volume metrics for the PTV and plan quality indices. For the latter, the 70% isoline was considered as the reference isodose level.

| Index     | Adapted plan | Original plan | Gel     |
|-----------|--------------|---------------|---------|
| $D_{\text{mean}}$ (%) | 83.3         | 84.5          | 83.7    |
| $D_{95}$ (%)     | 63.1         | 63.9          | 65.5    |
| $D_{\text{max}}$ (%) | 111.7        | 112.6         | 108.6   |
| $D_{\text{min}}$ (%) | 45.5         | 48.4          | 44.5    |
| RCI       | 1.11         | 1.16          | 1.12    |
| PCI       | 0.68         | 0.67          | 0.69    |
| HI        | 1.60         | 1.61          | 1.55    |
| Q         | 0.65         | 0.69          | 0.63    |

Abbreviations: RCI: RTOG conformity index, PCI: Paddick’s conformity index, HI: homogeneity index, Q: quality of coverage.

---

Figure 6. DVHs for the PTV and OAR (brainstem) resulting from the normalized adapted and original plans, as well as the gel-measured 3D relative dose distribution.
Effort was made to employ dosimetry protocols and equipment with minimal levels of experimental uncertainties. Spatial accuracy in the 2D dosimetry methodology followed is governed by the spatial registration uncertainty between scanned film images and the planning CT scan image stack. This procedure was facilitated by the specially designed MR compatible fiducials, which defined the film plane and position within the dicom coordinate system. Induced uncertainty levels can only be estimated based on a previous work which employed the same concept, although metal pins were used (Pappas et al 2017). According to that study, registration uncertainty was 1.5 mm. Given the relatively larger fiducials (and, consequently, wider film holes) used herein, it can be assumed that the spatial uncertainty is even larger. Dosimetric uncertainty is dose dependent and mainly stems from the fitting procedure, necessary to obtain the parameters of the dose-response calibration curve (Alde-laijan et al 2011). At the 8 Gy dose level, the corresponding dosimetric uncertainty reaches 1.2% and increases with decreasing dose. Other scanner-related uncertainties do not exceed 0.4% (Pappas et al 2017). Regarding gel dosimetry, spatial uncertainty is related to the rigid registration step between on-couch MR scans and planning CT scan. In a multi-institutional study, MR/CT registration uncertainty was estimated at 1.8 mm for a cranial case (Ulin et al 2010). Since relative dosimetry was performed using gels, global systematic uncertainties are not relevant. However, type B (statistical) uncertainty is still relatively high as it is related to MR read-out. More specifically, based on the gel vials irradiation, T2 reproducibility was estimated at 2.3%.

A number of limitations are noteworthy and should be considered when relying on the presented results to evaluate the dosimetric performance of Unity. Regarding the PseudoPatient™ phantoms, despite being produced based on real patient CT scan, air cavities (e.g. sinuses) are flooded with water and, therefore, not modeled (Makris et al 2019). Thus, dose verification around air cavities (where the electron return effect would be of particular interest) could not be performed. In addition, uncertainty reduction in film dosimetry could be further achieved by employing the triple channel dosimetry protocol (Micke et al 2011). Dose calculations by Monaco were characterized by limited spatial resolution. The $3 \times 3 \times 3$ mm$^3$ dose calculation grid resolution might not be considered appropriate for a benchmarking test. Although it could be improved in a phantom study, this would come at the expense of calculation time and, consequently, burden the overall duration of each treatment session in clinical practice. On the other hand, one could argue that an end-to-end benchmarking test should involve all parameters and conditions employed in routine clinical practice and not those of the ideal case. Lastly, the real-time dose accumulation monitoring capabilities were not explored in-depth and were limited to a single MRI acquisition, as a proof-of-concept study.

Although calibration vials were available, gel measurements were normalized to avoid any potential dependencies, temporal variations and volume effects (Ballock et al 2010), yielding 3D relative dosimetry only. Although absolute dose determination in 3D is attractive, an in-depth assessment of the gel dose-response characteristics and dependencies was beyond the scope of this study. However, it should be mentioned that this work demonstrated that polymer gel is an excellent candidate for end-to-end QA in Unity, combining a series of attractive features; (i) magnetic field independency of dose-response, (ii) without inducing field perturbations, (iii) MR compatibility, (iv) realistic MR and CT signal allowing for realistic adaptive treatment planning, (iv) on-couch dose read-out at the treatment delivery position immediately after irradiation and (v) real-time dose accumulation monitoring capabilities.

In response to the increasing demand for methodologies and tools suitable for dosimetry in the Unity workflow (De Prez et al 2019), a PseudoPatient™ phantom in conjunction with any dosimetric system enables introduction of rigorous end-to-end QA protocols suitable for commissioning and benchmarking, periodic QA, confidence building and training purposes when introducing new techniques, as well as, in the case of suspected or reported problems in the existing practices. The phantom employed, features realistic MR/CT contrast and human-like bone structures in anthropomorphic geometry (Makris et al 2019), allowing the implementation of patient-specific pre-treatment QA protocols in the actual clinical workflow. Other potential fields of application involve dosimetry in challenging single isocenter multiple metastases cases (especially if combined with 3D dosimetry) and QA procedures in extracranial radiotherapy such as stereotactic body radiotherapy cases. Furthermore, with appropriate inserts, the phantom could also be used for QA in MR/CT imaging and spatial registration procedures. Specially designed tests focusing on the performance of virtual couch shifting and the evaluation of adaptive treatment planning methods are already being considered.

5. Conclusion

In this work, the PseudoPatient™ head phantoms were employed and combined with 2D and 3D dosimetry protocols to evaluate the overall dosimetric performance of the Elekta Unity MR-linac system for a cranial IMRT case.

The suitability of the employed phantoms for QA procedures in 1.5T Elekta Unity MR-linacs was demonstrated. Based on the realistic MR/CT contrast of the phantoms, the presented benchmarking test also included the virtual couch shifting and adaptive treatment planning, steps routinely performed in clinical practice.
Regarding dosimetry, several novel tools and methodologies were presented. MR compatible fiducial markers facilitating registration of the film measurements were constructed. MR scanning protocols for gel dose read-out were developed, while demonstrating that relative 3D dosimetry can provide acceptable results even immediately after irradiation.

Overall results suggest that TPS calculations (for both the adapted and original plans) were validated within the experimental uncertainties involved. Comparison did not solely rely on GI calculations. By exploiting the 3D relative dose information, agreement of measurements with TPS predictions was further verified by comparing DVHs, dose-volume metrics and plan quality indices. For this specific case and adaptation method used, no considerable discrepancies were detected between the adapted and original plans.

As a side study, real-time dose accumulation monitoring based on T2w images was presented for the first time.

Acknowledgments

Elekta is acknowledged for allowing access to one of the pre-clinical Unity systems where the irradiations of this study were performed.

RTsafe PC is acknowledged for providing the PseudoPatient™ phantoms and gel batch used throughout this work.

The Greek Atomic Energy Commission is gratefully acknowledged for the film calibration irradiations per-

ORCID iDs

E P Pappas https://orcid.org/0000-0003-4030-2241

References

Agnew J, O’Grady F, Young R, Duane S and Budgell G J 2017 Quantification of static magnetic field effects on radiotherapy ionization chambers Phys. Med. Biol. 62 1731–43

Aldelaijan S, Mohammeh H, Tomic N, Liang L-H, Delflois F, Sarfnehia A, Abdel-Rahman W, Seuntjens J and Devic S 2011 Radiochromic film dosimetry of HDR 192Ir source radiation fields Med. Phys. 38 6074–83

Baldock C, De Deere Y, Doran S, Ibbott G, Fisarke A, Lepage M, McAuley K B, Oldham M and Schreiner L J 2010 Polymer gel dosimetry Phys. Med. Biol. 55 R1–63

Billas I, Bouchard H, Oelfke U and Duane S 2019 The effect of magnetic field strength on the response of Gafchromic EBT3 film Phys. Med. Biol. 64 06NT03

Costa F, Doran S J, Hanson I M, Nill S, Billas I, Shipley D, Duane S, Adamovics J and Oelfke U 2018 Investigating the effect of a magnetic field on dose distributions at phantom-air interfaces using PRESAGE™ 3D dosimeter and Monte Carlo simulations Phys. Med. Biol. 63 05NT01

De Prez L, De Pooter J, Jansen B, Woodings S, Wolthus J, Van Asselen B, Van Soest T, Kok J and Raaymakers B 2019 Commissioning of a water calorimeter as a primary standard for absorbed dose to water in magnetic fields Phys. Med. Biol. 64 035013

De Vries J H W, Seravalli E, Houweling A C, Woodings S J, Van Rooij R, Wolthus J W H, Lagendijk J J W and Raaymakers B W 2018 Characterization of a prototype MR-compatible Delta4 QA system in a 1.5 tesla MR-linac Phys. Med. Biol. 63 02NT02

De Vries J H W, Seravalli E, Houweling A C, Woodings S J, Van Rooij R, Wolthus J W H, Lagendijk J J W and Raaymakers B W 2018 Characterization of a prototype MR-compatible Delta4 QA system in a 1.5 tesla MR-linac Phys. Med. Biol. 63 02NT02

Devic S, Seuntjens J, Sham E, Podgorsak E B, Schmidtlein C R, Kirov A S and Soares C G 2005 Precise radiochromic film dosimetry using a flat-bed document scanner Med. Phys. 32 2245–53

Feuvret L, Noël G, Mazeron J and Bey P 2006 Conformity index: a review Int. J. Radiat. Oncol. Biol. Phys. 64 333–42

Hackett S L, Van Asselen B, Wolthus J, Woodings S J, Lagendijk J J W and Raayakers B W 2016 Consequences of air around an ionization chamber: are existing solid phantoms suitable for reference dosimetry on an MR-linac? Med. Phys. 43 3961–8

Hillbrand M et al 2019 Gel dosimetry for three dimensional proton range measurements in anthropomorphic geometries Z. Med. Phys. 29 162–72

Houweling A C, De Vries J H W, Wolthus J, Woodings S, Kok J G M, Van Asselen B, Smit K, Bel A, Lagendijk J J W and Raayakers B W 2016 Performance of a cylindrical diode array for use in a 1.5T MR-linac Phys. Med. Biol. 61 N80–9

Hussein M, Clark C H and Nisbet A 2017 Challenges in calculation of the gamma index in radiotherapy—towards good practice Phys. Med. 36 1–11

Lee H J, Kadbi M, Bosco G and Ibbott G S 2018 Real-time volumetric relative dosimetry for magnetic resonance—image-guided radiation therapy (MR-IGRT) Phys. Med. Biol. 63 045021

Lee H J, Roed Y, Venkataraman S, Carroll M and Ibbott G S 2017 Investigation of magnetic field effects on the dose-response of 3D dosimeters for magnetic resonance—image guided radiation therapy applications Radiother. Oncol. 125 426–32

Lim-Reinders S, Keller B M, Al-Ward S, Sahgal A and Kim A 2017 Online adaptive radiation therapy Int. J. Radiat. Oncol. Biol. Phys. 99 9994–10003

Low D A and Dempsey J F 2003 Evaluation of the gamma dose distribution comparison method Med. Phys. 30 2455

Makris D N, Pappas E P, Zoros E, Papapolykoua N, Saeni D L, Kalatzakis G, Zourari K, Efstatopoulos E P, Maris T G and Pappas E 2019 Characterization of a novel 3D printed patient specific phantom for quality assurance in cranial stereotactic radiosurgery applications Phys. Med. Biol. 64 105009

McPartlin A J et al 2016 MRI-guided prostate adaptive radiotherapy—a systematic review Radiother. Oncol. 119 371–80
Menten M J, Fast M F, Nill S, Kamerling C F, McDonald F and Oelfke U 2016 Lung stereotactic body radiotherapy with an MR-linac—quantifying the impact of the magnetic field and real-time tumor tracking Radiother. Oncol. 119 461–6
Mick F, Lewis D F and Yu X 2011 Multichannel film dosimetry with nonuniformity correction Med. Phys. 38 2523–34
Mutic S and Dempsey J F 2014 The ViewRay system: magnetic resonance-guided and controlled radiotherapy Semin. Radiat. Oncol. 24 196–9
O’Brien D J and Sawakuchi G O 2017 Monte Carlo study of the chamber-phantom air gap effect in a magnetic field Med. Phys. 44 3830–8
Paddick I 2000 A simple scoring ratio to index the conformity of radiosurgical treatment plans. Technical note J. Neurosurg. 93 219–22
Pantisos A, Antypas C, Zoros E, Pantelatos P, Lekas L, Romanelli P, Zourari K and Houdakakis C J 2018 On the total system error of a robotic radiosurgery system: phantom measurements, clinical evaluation and long-term analysis Phys. Med. Biol. 63 165015
Papoutsaki M-V, Maris T G, Pappas E, Papadakis A E and Damilakis J 2013 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. 29 453–60
Pappas E 2009 On the role of polymer gels in the dosimetry of small photon fields used in radiotherapy J. Phys.: Conf. Ser. 164 012060
Pappas E, Seimenis I, Dellios D, Kollias G, Lampropoulos K I and Karaiskos P 2018 Assessment of sequence dependent geometric distortion in contrast-enhanced MR images employed in stereotactic radiosurgery treatment planning Phys. Med. Biol. 63 135006
Pappas E P, Seimenis I, Moutsatsos A, Georgiou E, Nomikos P and Karaiskos P 2016 Characterization of system-related geometric distortions in MR images employed in gamma knife radiosurgery applications Phys. Med. Biol. 61 6993–7011
Pappas E P, Zoros E, Moutsatsos A, Peppa V, Zourari K, Karaiskos P and Papagiannis P 2017 On the experimental validation of model-based dose calculation algorithms for 192Ir HDR brachytherapy treatment planning Phys. Med. Biol. 62 4160–82
Pappas E, Maris T, Angelopoulos A, Paparigopoulos M, Sakellios L, Sandilos P, Voyiatzi S and Vlachos I 1999 A new polymer gel for magnetic resonance imaging (MRI) Phys. Med. Biol. 44 2677–84
Pathmanathan A U et al 2018 Magnetic resonance imaging-guided adaptive radiation therapy: a ‘game changer’ for prostate treatment? Int. J. Radiat. Oncol. Biol. Phys. 100 361–73
Poettinger S, Dohm O S, Kapsch R P and Thorwarth D 2018 Ionization chamber correction factors for MR-linacs Phys. Med. Biol. 63 11NT03
Raaymakers B W et al 2017 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. 62 L41–50
Roed Y, Ding Y, Wen Z, Wang J, Pinsky L and Ibbott G 2017 The potential of polymer gel dosimeters for 3D MR-IGRT quality assurance J. Phys.: Conf. Ser. 847 012059
Ruschin M, Sahgal A, Tseng C L, Sonier M, Keller B and Lee Y 2017 Dosimetric impact of using a virtual couch shift for online correction of setup errors for brain patients on an integrated high-field magnetic resonance imaging linear accelerator Int. J. Radiat. Oncol. Biol. Phys. 98 699–708
Saenz D L, Li Y, Rasmussen K, Statidakis S, Pappas E and Papanikolaou N 2018 Dosimetric and localization accuracy of Elekta high definition dynamic radiosurgery Phys. Med. 54 146–51
Smit K, Sjöholm J, Kok J G M, Lagendijk J J W and Raaymakers B W 2014 Relative dosimetry in a 1.5 T magnetic field: an MR-linac compatible prototype scanning water phantom Phys. Med. Biol. 59 4099–109
Tseng C-L, Eppinga W, Saravalli E, Hackett S, Brand E, Ruschin M, Lee Y K, Atanafu E G and Sahgal A 2017 Dosimetric feasibility of the hybrid magnetic resonance imaging (MRI)-linac system (MRL) for brain metastases: the impact of the magnetic field Radiother. Oncol. 125 273–9
Ulin K, Urie M M and Cherlow J M 2010 Results of a multi-institutional benchmark test for cranial CT/MR image registration Int. J. Radiat. Oncol. Biol. Phys. 77 1584–9
Winkel D et al 2019a Adaptive radiotherapy: the Elekta unity MR-linac concept Clin. Transl. Radiat. Oncol. 18 54–9
Winkel D et al 2019b Evaluation of plan adaptation strategies for stereotactic radiotherapy of lymph node oligometastases using online magnetic resonance image guidance Phys. Imaging Radiat. Oncol. 9 58–64
Woodings S J et al 2018a Beam characterisation of the 1.5 T MRI-linac Phys. Med. Biol. 63 085015
Woodings S J, Wolthaus J W H, Van Asselen B, De Vries J H W, Kok J G M, Lagendijk J J W and Raaymakers B W 2018b Performance of a PTW 60019 microDiamond detector in a 1.5 T MRI-linac Phys. Med. Biol. 63 05NT04