Physica Medica 71 (2020) 71–81

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Original paper

PETER PHAN: An MRI phantom for the optimisation of radiomic studies of the female pelvis

Linda Bianchini[a,*], Francesca Botta[b], Daniela Origgi[b], Stefania Rizzo[c], Manuel Mariani[d], Paul Summers[e], Pablo García-Polo[f,1], Marta Cremonesi[g], Alessandro Lascialfari[a,d]

[a] Department of Physics and INSTM RU, Università degli Studi di Milano, Italy
[b] Medical Physics Unit, IEO, European Institute of Oncology IRCCS, Milan, Italy
[c] Clinica di Radiologia EOC, Istituto di Imaging della Svizzera Italiana, Sede Ospedale Regionale di Lugano, Switzerland
[d] Department of Physics and INSTM RU, Università degli Studi di Pavia, Italy
[e] Division of Radiology, IEO, European Institute of Oncology IRCCS, Milan, Italy
[f] Southern Europe Global Research Organization, GE Healthcare, Madrid, Spain
[g] Radiation Research Unit, IEO, European Institute of Oncology IRCCS, Milan, Italy

ARTICLE INFO

Keywords:
Radiomics
Magnetic resonance imaging (MRI)
Texture
Radiomic phantom
Repeatability

ABSTRACT

Purpose: To develop a phantom for methodological radiomic investigation on Magnetic Resonance (MR) images of female patients affected by pelvic cancer.
Methods: A pelvis-shaped container was filled with a $MnCl_2$ solution reproducing the relaxation times (T_1, T_2) of muscle surrounding pelvic malignancies. Inserts simulating multi-textured lesions were embedded in the phantom. The relaxation times of muscle and tumour were measured on an MR scanner on healthy volunteers and patients; T_1 and T_2 of $MnCl_2$ solutions were evaluated with a relaxometer to find the concentrations providing a match to *in vivo* relaxation times. Radiomic features were extracted from the phantom inserts and the patients' lesions. Their repeatability was assessed by multiple measurements.
Results: Muscle T_1 and T_2 were 1128 (806–1378) and 51 (40–65) ms, respectively. The phantom reproduced *in vivo* values within 13% (T_1) and 12% (T_2). T_1 and T_2 of tumour tissue were 1637 (1396–2121) and 94 (79–101) ms, respectively. The phantom insert best mimicking the tumour agreed within 7% (T_1) and 24% (T_2) with *in vivo* values. Out of 1034 features, 75% (95%) had interclass correlation coefficient greater than 0.9 on T_1 (T_2)-weighted images, reducing to 33% (25%) if the phantom was repositioned. The most repeatable features on phantom showed values in agreement with the features extracted from patients' lesions.
Conclusions: We developed an MR phantom with inserts mimicking both relaxation times and texture of pelvic tumours. As exemplified with repeatability assessment, such phantom is useful to investigate features robustness and optimise the radiomic workflow on pelvic MR images.

1. Introduction

Cancer remains one of the main causes of death worldwide [1]. With the anatomy and physiology of a tumour being the result of multiple gene expression patterns, its origin and growth is a complex process, involving both spatial and temporal interactions at scales that range from the molecular to the whole-organ. This non-linear, interacting process can lead to heterogeneity within tumour tissue, and thus complicate clinical workflows, from diagnosis to treatment selection and response assessment.

Given the link of anatomy with the underlying gene expression pattern, quantitative imaging has been proposed as a means to assess functional and progression mechanisms of the imaged tumour [2]. To go beyond the visual interpretation of the images in pursuing this goal, the practice of high-throughput extraction of large numbers of quantitative features from radiological images, referred to as "radiomics", has emerged [3]. Dedicated software solutions allow the calculation of "radiomic features", each a mathematical algorithm applied to the intensity values in each voxel in the image. These features allow a wide range of image properties to be extracted, many of which - such as those related to texture - may reflect tumour heterogeneity. In clinical oncology, radiomic features are investigated for their correlation with the

* Corresponding author at: Physics Department, Università degli Studi di Milano, Via Celoria 16 – 20133 Milan, Italy.
E-mail address: linda.bianchini@unimi.it (L. Bianchini).
[1] Pablo García-Polo is a GE Healthcare employee.

https://doi.org/10.1016/j.ejmp.2020.02.003
Received 14 October 2019; Received in revised form 29 January 2020; Accepted 4 February 2020

patients' clinical history, including response to therapy, survival and genomic profile. The final goal is to integrate the most informative features into predictive models (or signature) useful for prognosis, patient stratification, choice of treatment strategy, early assessment of response, or recurrence probability. In this sense, radiomics could represent a supportive tool for clinical decision, towards the achievement of personalized medicine [4,5].

To bring radiomics into clinical practice, the proposed predictive models must be reliable and strongly validated as well as the results (the features) need to be highly stable and reproducible. Developing a robust model requires a broad training and validation data set, with image data from large number of patients, for any specific pathology [6], setting radiomics in the framework of "big data" analyses. Sharing data and images of patients in multicentric studies could represent the answer to meet this specific demand, but introduces the complication of variability due to differences in scanners and imaging protocols. Consequently, methodological studies are needed to define suitably reproducible procedures for images acquisition and processing, as well as to optimise computational and data analysis approaches, that must possess an undisputable and robust statistical meaning.

In this regard, Magnetic Resonance Imaging (MRI) has been less extensively investigated than Computed Tomography (CT) or Positron Emission Tomography (PET) for use in radiomics. The main reason lays in the complexity of MRI, where the numerous variables involved are optimised in light of the clinical needs and equipment capacities, resulting in wide variety of MRI protocols having nominally the same appearance but subtle differences among them. Moreover, the lack of a standard intensity scale for MR images necessitates the use of additional pre-processing, such as intensity normalisation and bias field correction [7].

Early clinical demonstrations of radiomics in MRI have been seen in prostate [8] and breast cancer [9], as well as glioblastoma multiforme [10] and head-and-neck squamous cell carcinoma [11]. Promising results have also been seen in female pelvic tumours, where preoperative radiomic features were seen to be predictive of lymph-vascular space invasion [12] and of lymph node metastasis [13] in cervical cancer patients. The prediction of recurrence in advanced cervical cancer patients treated with concurrent chemoradiotherapy has also been demonstrated [14]. Similarly, MRI based radiomics of patients with rectal cancer has shown potential both to predict the tumour response to therapy [15,16] and to non-invasively evaluate tumour differentiation and other biological properties [17]. In our Institute, several studies on MR images of female patients with pelvic tumours have been carried out [18,19], opening the way to possible radiomic application in this field, as previously performed on CT images [20].

In the literature, however, the methods employed in every step of the radiomic analysis, including image acquisition, segmentation, features extraction and statistical analysis, vary widely. This variability makes comparison between studies challenging and hinders generalisation of results between centers. A further complication factor is that some radiomic features have been shown to be dependent on the MR image acquisition parameters [21], a particularly important consideration as validation of radiomic findings via multicentric and retrospective studies is critical to clinical acceptance.

To understand the sensitivity of radiomic models to differences between scanners and scanning protocols, a comprehensive investigation of the influence of different image acquisition settings (involving variations in MR sequences parameters, scanner, vendor and/or centre) on radiomic features is warranted. Establishing a well-defined, standard procedure for supporting radiomic analyses is also advisable in view of prospective clinical studies [22].

Some of the needed methodological work can be fulfilled through studies on phantoms, as these provide a controlled experimental setting and the possibility for repeated measurements. The phantoms currently available for quality control testing of MR scanners, however, are not suitable for radiomic investigation. In fact, they tend to lack the heterogeneity that serves to differentiate radiomic features between tissues. Phantoms designed for routine quality assurance usually consist of a container filled with homogeneous aqueous solutions in which are embedded objects of specific shape and variable size (commonly plastic homogeneous inserts) to assess individual aspects of scanner performance in terms of geometry, distortion and resolution [23]. In particular situations, other solutions have been investigated, including jelly-like materials to provide tissue-equivalent phantoms [24], phantom created with fruits to simulate different texture [25] and anthropomorphic and multimodal phantoms that approximate *in vivo* anatomy and imaging appearance [26].

Some efforts have also pursued digital phantoms, mainly for brain applications. Yang, Ford et al. [27,28] made use of a digital MR phantom to confirm that the radiomic features are influenced by the image acquisition process and reconstruction algorithm. They also underlined the need for a physical phantom for experimental validation of their preliminary results.

Although the available MRI phantoms are increasingly specialized across a widening range of applications, to date there is still none optimised for radiomic purposes.

Physical phantoms specifically dedicated to the evaluation of radiomic features reliability and robustness in CT [29] and PET [30] have recently been reported. To our knowledge, a physical MRI phantom dedicated to radiomic analysis has not been designed yet.

The aim of this paper is to introduce a new phantom, called PETER PHAN (PElvis TExtuRe PHANtom), that simulates the female pelvis for radiomic purposes and allows methodological studies in support of the clinical use of radiomics. As a secondary aim, to show how the phantom can be exploited, we provide an assessment of the radiomic features repeatability on T_1-weighted (T_1-W) and T_2-weigthed (T_2-W) images of the phantom, acquired with MR sequences included in the gynaecological imaging protocol of our Institute.

2. Materials and methods

The characteristics desired of a phantom suitable for supporting clinical MRI-based radiomic studies of tumours located in female pelvis include: being pelvis-shaped, constructed of MR-safe materials, providing relaxation times similar to human tissues and containing regions of different textures, similar to those found *in vivo* for the tumoural tissue. As described in the following section, we addressed these design criteria through a series of experimental steps regarding:

- measurements of *in vivo* longitudinal relaxation time (T_1) and transverse relaxation time (T_2) of pelvis tissues on both patients and healthy volunteers;
- choices of materials and formulations to reproduce the relaxation times measured *in vivo*;
- investigation of means to incorporate different textures into the phantom;
- validation of the assembled phantom through comparison of relaxation times and texture with the *in vivo* reference values.

After describing each step for the phantom assembling, at the end of this Section an example of the phantom usage in radiomic analysis is proposed.

2.1. *In vivo* T_1 *and* T_2 *mapping*

19 healthy volunteers (22–53 yrs, mean 35) and 5 patients with a pelvic tumour (29–57 yrs, mean 47) were imaged with dedicated MR sequences for *in vivo* assessment of T_1 and T_2 relaxation times of healthy tissue (muscle) and pelvis tumour. Written informed consent to undergo the examination, and to the use of anonymised clinical and imaging data for scientific and/or educational purposes, was obtained from all subjects involved. Clinical characteristics of patients are listed in

Table 1
Characteristics of the patient cohort.

Patient	Age (years)	Pathology	Tumour site
1	57	Vulvar cancer	Inguinal lymph nodes
2	47	Carcinoma of cervix	Uterine cervix
3	50	Adenocarcinoma of the rectum	Rectum
4	29	Carcinoma of cervix	Uterine cervix
5	54	Submucous myoma of the uterus	Endometrium

Table 1. Inclusion criterion for patients was the presence of a lesion visible on MR images, with a minimum tumour diameter 18 mm.

Images acquisition was carried out using a 1.5 T MR scanner (Optima MR450W, General Electric Healthcare, Waukesha, USA) with a 24 channels anterior body array coil and a 24 channels spine array coil.

Maps of T_1 can be obtained with various approaches, including Inversion Recovery (IR) pulse sequence [31] and novel techniques like Magnetic Resonance Fingerprinting [32], although the last one is far from being clinically validated and confirmed by studies in a sizeable number of different centers. In this study *in vivo* T_1 mapping made use of the variable flip angle (VFA) spoiled gradient recalled echo (SPGR) technique [33]. Three axial scans were acquired, each with a different flip angle (5, 10 or 20°), keeping the other parameters fixed: repetition time (TR) 6.7 ms, echo time (TE) 4.2 ms, slice thickness 3 mm, slice spacing 3 mm, field of view (FoV) 220x220 mm^2.

T_2 mapping was based on a dedicated 2D multi-echo spin-echo pulse sequence which ran under research mode and could read up to 16 echoes. The sequence parameters were: TR 5000 ms, min TE 7.5 ms, max TE 119.8 ms, echo spacing 7.5 ms, slice thickness 5 mm, slice spacing 6.5 mm, FoV 220x220 mm^2.

After the acquisition, the DICOM images were anonymised and exported for processing. The VFA data were processed with the NOVIFAST [34] algorithm (version 1.0.0.3, University of Antwerp, Belgium) to calculate T_1 maps. T_2 maps were generated by fitting a mono-exponential curve to the 16-echo spin-echo data on a voxel-by-voxel basis, using a MATLAB (version R2018b, The MathWorks, Inc., USA) script developed in-house.

Regions of interest (ROIs) in the gluteus maximus muscle were drawn manually on the T_1 and T_2 maps of all subjects. In addition, an expert radiologist drew a ROI for the tumour on the axial T_2-W images obtained as part of the routine clinical protocols. The tumour ROI was then transferred onto the relaxation time maps. For each ROI, the subject-wise mean values of T_1 and T_2 were extracted.

2.2. Relaxometry of $MnCl_2$ solutions

Conventional MR phantoms are based on solutions of paramagnetic contrast agents, which are easily soluble in water. The solutions obtained are stable in time and mimic the typical relaxation times of human tissues well [35]. The paramagnetic ion manganese chloride ($MnCl_2$) was chosen for this study due to its availability and low cost. A Fourier Transform spectrometer (Apollo, Tecmag, Houston, USA) for pulsed Nuclear Magnetic Resonance (NMR) was used to measure T_1 and T_2 of water solutions with increasing $MnCl_2$ concentration (range 0.1–1.2 mM, in steps of 0.1 mM). The measurements were performed at room temperature (25°) with an operating proton resonance frequency of 63.86 MHz, corresponding to a 1.5 T static magnetic field, the same used for T_1 and T_2 mapping *in vivo*. T_1 (longitudinal relaxation time) measurements were performed with a Saturation Recovery (SR) Spin–Echo (SE) sequence, composed of three saturation $\pi/2$ pulses – aimed at saturating the absorption line - followed by a Spin-Echo sequence [36] for reading the signal. T_2 (transverse relaxation time) was evaluated with a Carr-Purcell-Meiboom-Gill (CPMG) sequence [37,38]. The time between $\pi/2$ pulse and the first π pulse, referred to as τ, was chosen as short as possible to reduce the effects of diffusion and a high number of echoes was used in order to follow the entire relaxation process. Sequences diagrams and parameters are reported in Fig. 1 and Table 2.

Table 2
Parameters set on the NMR spectrometer for the SR - SE and CPMG sequences.

SR - SE				CPMG			
$(\pi/2)_{sat}$	τ_{sat}	$\pi/2$	τ_{echo}	π	$\pi/2$	τ	π
5.3 µs	200 ms	5.1 µs	1 ms	10.2 µs	5.1 µs	50 µs	10.2 µs

The NMR raw data were analysed with the in-house QtNMR software and with OriginPro (version 8.1, OriginLab Corporation, Northampton, USA) software. The signal from the SR-SE experiment, performed on each sample with different $MnCl_2$ concentration, describes the temporal evolution of the longitudinal nuclear magnetization $M_z(t)$ as

$$M_z(t) = M_\infty\left(1 - e^{-\frac{t}{T_1}}\right) \tag{1}$$

where M_∞ is the magnetization at thermal equilibrium.

From this equation, we defined:

$$e^{-\frac{t}{T_1}} = \frac{M_\infty - M_z(t)}{M_\infty} \equiv Y \tag{2}$$

that was fitted with a mono-exponential function to evaluate the spin–lattice relaxation time T_1.

Similarly, the signal form the CPMG

$$M_{xy}(t) = M_0 e^{-\frac{t}{T_2}} \tag{3}$$

describes the temporal evolution for the transverse magnetization $M_{xy}(t)$. A mono-exponential fit function was used to evaluate T_2 from the experimental data. Combining the errors due to electronics and software in the data collection phase with the errors on model fitting, the maximum total *a priori* error was estimated to be 10% of the values measured.

The relaxation times obtained from the samples were plotted as a function of the $MnCl_2$ concentration in order to build two calibration curves, one for T_1 and one for T_2. The two set of experimental points were fitted with the function

$$T_i(C) = \frac{a}{C + b} \tag{4}$$

with $i = (1, 2)$, C as the concentration and (a, b) as fitting parameters. The optimal fit for the experimental data was obtained with a MATLAB script developed in-house.

The concentration - relaxation time curves were used for the

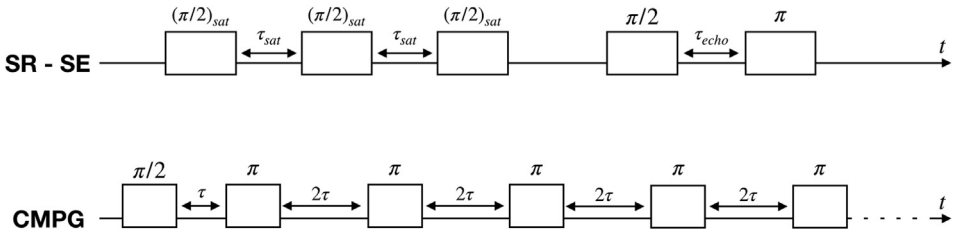

Fig. 1. Diagrams of pulse sequences used to measure T_1 and T_2 of $MnCl_2$ solutions.

selection of concentrations that produce relaxation times matching the values measured for tissues in the previous phase of the work.

2.3. Phantom design and assembly

To approximate the shape of the pelvis, the NEMA IEC Body Phantom Set™ (Spectrum Corporation, Durham, USA), provided by the Nuclear Medicine department of our Institute, was used as an abdomen-shaped container made of MR compatible materials. This phantom is routinely used for quality assurance on radionuclide imaging devices and is constructed entirely of solid plastics. It consists of an external plastic container, that serves as a background compartment, and an empty internal cylinder exploited to assemble PETER PHAN.

Two versions of the phantom were prepared: version one (V1) in which the T_1 relaxation time of tissues were reproduced, and version two (V2), with matching of T_2 relaxation times. In each version, the background compartment of the phantom was filled with a solution of $MnCl_2$ to reproduce the relevant (T_1 or T_2) relaxation time of muscle. To reproduce the lesion, which represents the target for radiomic studies, four cylindrical inserts (48 mm diameter, 72 mm height) were placed inside the phantom. The role of the inserts was twofold: to mimic the MR signal of a tumour and to provide a selection of textures for identifying a possible simulant for tumours. These paired demands were fulfilled by filling the inserts with a mixture of polystyrene spheres and agar gel, using the recipe introduced for the PSAG phantoms [39] as a starting point. The four inserts for each phantom were prepared with a 0.1% water solution of agar (Sigma-Aldrich, St. Louis, USA) which was heated to 100° and poured over the spheres, adding 1 ml of 0.1% NaN_3 per 1 L of agar as antiseptic. The first and second inserts (INS1-INS2) were filled with the medium-sized (3–4 mm) spheres, INS3 contained only small (1 mm) spheres and INS4 was prepared with a mixture of spheres of small, medium and large (7–8 mm) diameter spheres.

The difference among the inserts of phantom V1 and V2 was in the amount of spheres filling the inserts. The number of spheres per unit volume affected the T_1 and T_2 values of the spheres and gel mixture, so it was varied experimentally until the contrast between the T_1 (or T_2) values of the inserts and the background compartment solution (the last reproducing the muscle T_1 or T_2) matched the contrast between tumour tissue and muscle observed in patients' images. This procedure was chosen due to the fact that a direct measurement of T_1 and T_2 of the mixture in the NMR spectrometer was not feasible. The four inserts were fixed to the main NEMA cylinder in a central position inside the phantom, via a purpose-built plastic support.

To test the long-term stability of the phantom inserts, six months after their production PETER PHAN V1 and V2 were imaged on the same MR scanner with the same imaging parameters of T_1-W and T_2-W sequences described in the following paragraph. We note that between evaluations, the inserts were kept refrigerated at 4 °C. The visual inspection showed no deterioration of the object structures, confirmed by the comparison of the phantom images at the time of the main experiment and after six months. Additional evaluations on the phantom stability for longer periods of time are ongoing.

2.4. Validation

T_1 and T_2 mapping acquisitions were performed on PETER PHAN V1 and V2 respectively, in order to verify that the phantom adequately reproduced the range of relaxation times of muscle and tumour tissue. The relaxation time mapping acquisitions were performed with the same MRI sequences and settings as used above for *in vivo* T_1 and T_2 mapping.

Additionally, PETER PHAN V1 and V2 were imaged with the routine T_1-W and T_2-W sequences used for clinical pelvis diagnostic imaging. More in detail, the acquisition of PETER PHAN V1 was performed with a series of axial T_1-W images, with TR 354 ms, TE 8.8 ms, slice thickness 5 mm, spacing 5 mm, 512 × 512 matrix, whereas PETER PHAN V2 was acquired with an axial T_2-W sequence, with TR 3723 ms, TE 107 ms, slice thickness 5 mm, spacing 6 mm, matrix 512 × 512. Such images were visually compared with the clinical images of patients listed in Table 1, acquired on the same scanner with the following parameters: TR 541 ms, TE 8.1 ms, slice thickness 5.5 mm, spacing 5 mm, matrix 512 × 512 for the T_1-W acquisition; TR 7771 ms, TE 105 ms, slice thickness 5 mm, spacing 5.5 mm, matrix 512 × 512 for the T_2-W series.

Quantitative comparison of texture features was also performed, and will be described in the next paragraph.

Table 3
Relaxation times* of human tissues measured *in vivo* (upper) and of background and inserts measured in PETER PHAN V1 and V2 (lower).

		T_1 [ms]*	T_2 [ms]*
In vivo T_1 and T_2 mapping	Muscle	1128 (806–1378)	51 (40–65)
	Tumour	1637 (1396–2121)	94 (79–101)
Phantom validation	$MnCl_2$	1277 (768–2120)	57 (55–59)
	INS1	1366 (6–1999)	137 (106–183)
	INS2	1345 (22–2000)	123(101–192)
	INS3	1525 (16–1996)	117 (101–150)
	INS4	1447 (19–2000)	120 (89–167)

*mean (range).

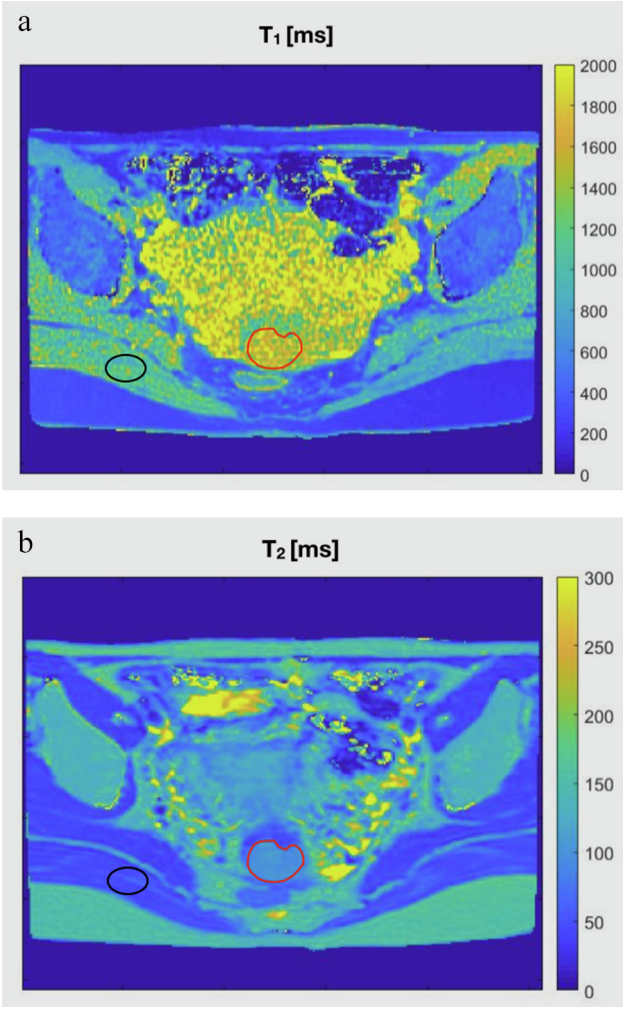

Fig. 2. (a) T_1 and (b) T_2 maps of a patient with cervical cancer. The tumour site region of interest (ROI) is highlighted with a red contour and the muscle ROI contour in black. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

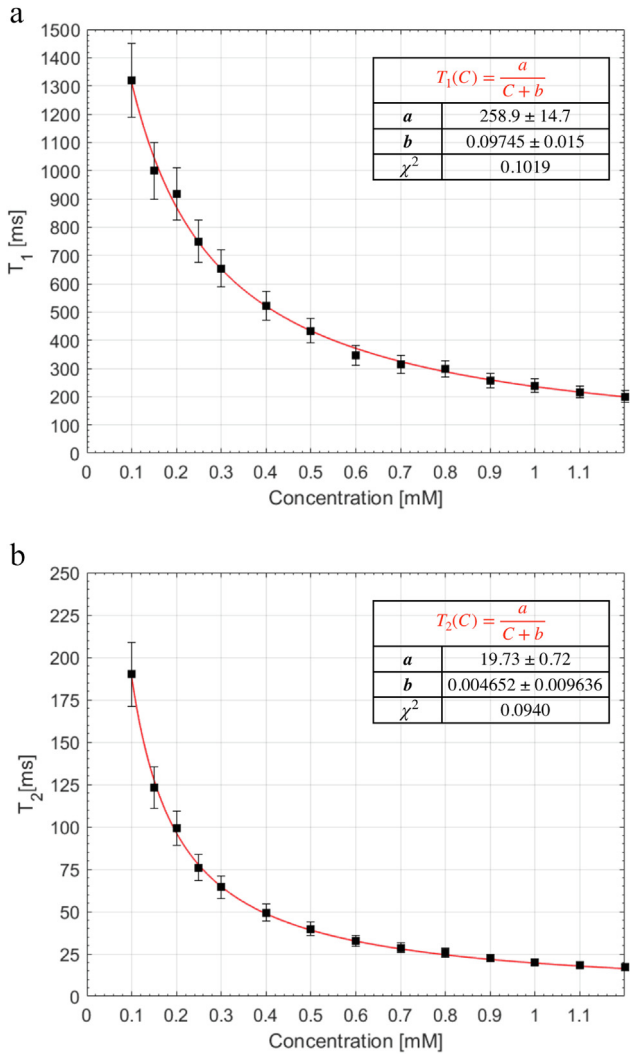

Fig. 3. (a) T_1 and (b) T_2 relaxation times for solutions with different $MnCl_2$ concentrations. Black dots are the experimental data. The red line represents the fitted model. The χ^2 values quantify the goodness of the fit. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.5. PETER PHAN for radiomic studies optimisation: Proof of concept

PETER PHAN images acquired with the T_1-W and T_2-W sequences of the clinical protocol described in the previous paragraph were exploited to carry out a first assessment of the radiomic features repeatability. For each sequence, the phantom acquisition was repeated ten times (acq. 1–10) without changing the experimental setup nor the scanning parameters. Subsequently, the phantom was repositioned and the ten acquisitions were repeated identically (acq. 11–20). Sixteen cylindrical ROIs of four sizes were drawn on the four phantom inserts for the first acquisition images (Fig. 6a) using 3D Slicer [40] (version 4.10.1, Brigham and Women's Hospital, Harvard University, Boston, USA) and considering three consecutive slices. The images of the other acquisitions (with phantom repositioning) were co-registered with the images of the first acquisition and the same ROIs used. PyRadiomics software [41] (version 2.2.0, Brigham and Women's Hospital, Harvard Medical School, Boston, USA) was exploited for images normalisation and to extract radiomic features.

The normalisation function was

$$f(x) = \frac{s(x - \mu_x)}{\sigma_x}$$

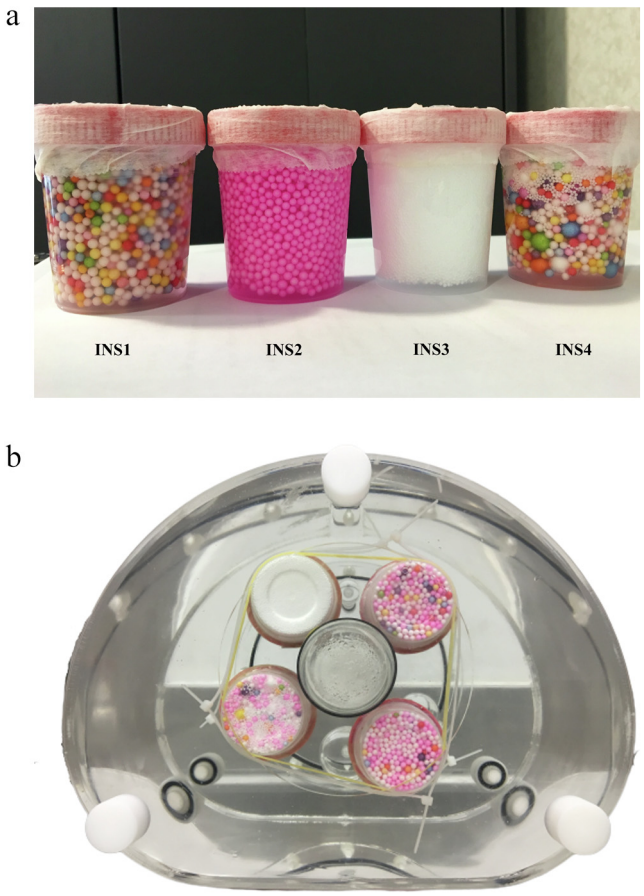

Fig. 4. (a) Inserts prepared for PETER PHAN V2, representing tumours with various textures. (b) Axial view of PETER PHAN assembled.

where x is the original intensity, $f(x)$ the normalised intensity, μ_x and σ_x the mean and the standard deviation of the image intensity values, s is a scaling factor which was set to 100 in this study. The radiomic features extracted from each ROI belonged to the categories: Shape, First Order, GLCM, GLRLM, GLSZM, NGTDM, GLDM[2]. The extraction was performed on both original images and images filtered with the following kernels: Laplacian of Gaussian, Wavelets, Exponential, Logarithm, Square, Square Root and 2D Local Binary Pattern. The filter width σ used for the Gaussian kernel was set to 6.0 mm and one level was used for the wavelet decompositions. The spatial resolution of the Local Binary Pattern operator was set to one. The fixed bin width was chosen as grey level discretisation technique, as suggested in PyRadiomics documentation [42]. The bin width was optimised for each extraction to obtain a number of bins in the range (30–130). Its value was set to 2 and 8 for the T_1-W and T_2-W images of the phantom respectively.

The interclass correlation coefficient (ICC) for absolute agreement [43] was calculated pairwise for each radiomic feature of the first set of acquisitions (1–2, 3–4, …, 9–10) and the second one (11–12, 13–14, …, 19–20) to test repeatability of the radiomic features with fixed phantom position. The ICC was then calculated matching the measurements from the two sets of acquisitions (1–11, 2–12, …, 10–20), to test the impact of repositioning on repeatability. According to the ICC thresholds provided by Koo and Li [43], the radiomic features extracted from the

[2] GLCM = Grey Level Co-occurence Matrix; GLRLM = Grey Level Run Length Matrix; GLSZM = Grey Level Size Zone Matrix; NGTDM = Neighbouring Grey Tone Difference Matrix; GLDM = Grey Level Dependence Matrix. For each matrix, at fixed angle, the distance between the central voxel and the neighbour, calculated with the infinity norm, was set to 1.

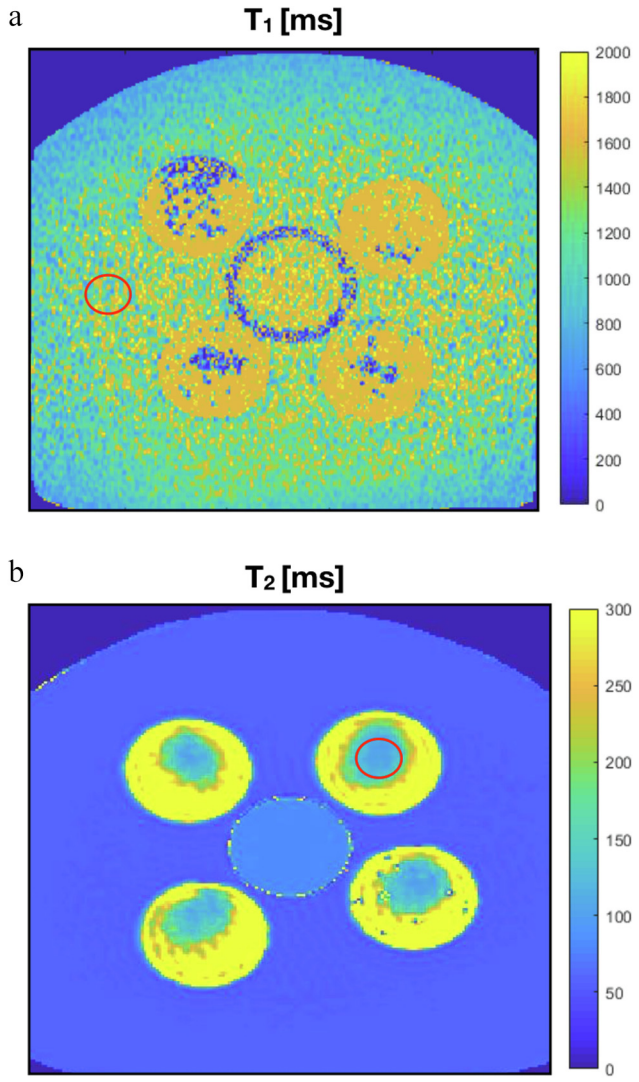

Fig. 5. (a) T_1 map of PETER PHAN V1. An example of region of interest (ROI) in the background compartment is highlighted with a red contour. (b) T_2 map of PETER PHAN V2. An example of ROI drawn inside one of the inserts is highlighted with a red contour. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

phantom images were classified into four categories: features showing excellent (ICC > 0.9), good (0.75 < ICC ≤ 0.9), moderate (0.5 < ICC ≤ 0.75) and poor (ICC ≤ 0.5) repeatability.

The features showing the best performance in terms of repeatability were selected and, following an approach similar to the one proposed by Samei et al. [44], the range of the features values in the phantom and in a set of patients was compared. The procedure described for extracting the features from the phantom images was used to extract the radiomic features from the tumour volume on the patients' images (both T_1-W and T_2-W acquisitions, as in Fig. 6. Patient listed in Table 1), with 3 as bin width. In addition, a statistical test of compatibility between the two samples (the values extracted from the phantom images and those obtained from the patients' images), was performed to test the hypothesis that the two populations have equal means. Since each radiomic feature is given by the sum of independent variables of finite variance, the distributions from which the data derive can be supposed to be normal, in agreement with the Central Limit Theorem. Since the variances of the two normal distributions are unknown, we performed the general two-sample *t*-test, known as Welch's *t*-test, which applies to two samples, deriving from normal distributions, with unequal sample sizes (as here) and/or unequal variances [45].

The comparison in terms of texture properties between the phantom inserts and the patients' lesions allowed a further validation in addition to the work described in the previous paragraph.

3. Results

3.1. *In vivo T_1 and T_2 mapping*

The range of T_1 and T_2 values obtained for the tissues of interest are summarized in the upper part of Table 3 and examples of the *in vivo* T_1 and T_2 maps are given in Fig. 2. The values reported are the mean over all the patients and the minimum and maximum values.

3.2. *Relaxometry of $MnCl_2$ solutions*

The calibration curves, linking the concentration of the $MnCl_2$ solution with its relaxation times, are reported in Fig. 3. The chi-square values confirm that the data are well described by the hypothesized model.

3.3. *Phantom design and assembly*

Pictures in Fig. 4a show the inserts used to simulate lesions with different texture (V2) and Fig. 4b the assembled PETER PHAN. Based on the fitted model of T_1 relaxation as a function of $MnCl_2$ concentration, a 0.15 mM $MnCl_2$ solution was chosen to fill the background of PETER PHAN V1 to reproduce the muscle T_1. Similarly, the phantom V2 was prepared with a 0.4 mM solution of $MnCl_2$ to reproduce the muscle T_2.

3.4. *Validation*

T_1 and T_2 maps of PETER PHAN V1 and V2 can be seen in Fig. 5a and 5b respectively. The mean relaxation time values inside the four inserts and in the surrounding background are summarised in the lower part of Table 3. For the background solution representing the muscle, the values are the average of the results obtained on four identical ROIs identified in different positions on the fluid-filled region to assess the homogeneity of the relaxation time maps.

Comparing the phantom with patients' results (Table 3), the agreement between muscle and $MnCl_2$ is within 13% for T_1 mean values and 12% for T_2 mean values. The measurements performed on the inserts showed the following absolute percentage variations (T_1, T_2) with respect to the tumour mean value: INS1 (17%, 46%), INS2 (18%, 31%), INS3 (7%, 24%) and INS4 (12%, 28%). Fig. 6 offers a comparison of images of the phantom and a patient obtained with the clinical T_1-W and T_2-W sequences. These results showed that the inserts reproduced a typical lesion in terms of its relaxation times. The results presented in the following paragraph confirmed that the inserts are representative of a typical lesion also in terms of its texture properties.

3.5. *PETER PHAN for radiomic studies optimisation: Proof of concept*

The features classification in terms of repeatability for the T_1-W and the T_2-W sequences, with and without phantom repositioning, is shown in Table 4. In total, 1034 features were extracted for each ROI. The majority of the features showed excellent repeatability in the experiment with fixed setup: 773 (74.8%) in the T_1-W acquisition and 985 (95.3%) in the T_2-W acquisition. A consistent loss of features repeatability is evident after phantom repositioning, both for the T_1-W and T_2-W acquisitions. In this case the features with excellent repeatability decreases to 337 (32.6%) for the T_1-W acquisition and 254 (24.6%) for the T_2-W acquisition. 118 (11.4%) radiomic features show excellent repeatability in both experiments (with and without phantom repositioning) and both sequences (T_1-W and T_2-W acquisitions). Of these, the majority were extracted from original (non-filtered) images.

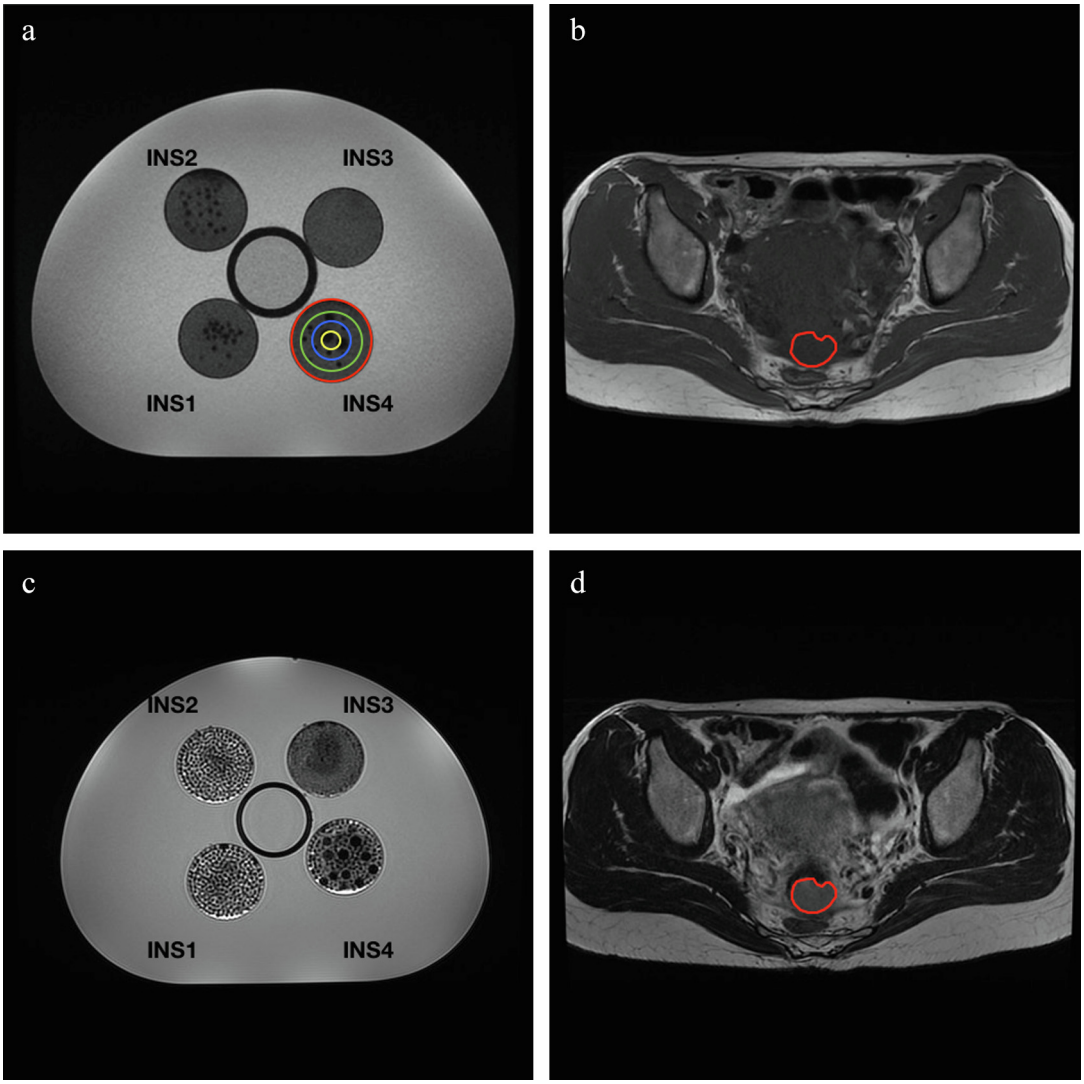

Fig. 6. (a-b) T_1-W image of PETER PHAN V1, representing tumours with various textures, and of a patient with cervical cancer. In (a) selected ROIs of four sizes identified on the phantom inserts for radiomic analysis are shown (yellow d = 12 mm; blue d = 24 mm; green d = 36 mm, red d = 48 mm). In (b) the tumour site in highlighted in red. (c-d) T_2-W image of PETER PHAN V2 and of the same patient. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Number (percentage) of radiomic features showing excellent, good, moderate and poor repeatability when extracted from the 16 ROIs segmented on the phantom inserts on T_1-W images of PETER PHAN V1 and T_2-W images of PETER PHAN V2 without and with phantom repositioning.

	PETER PHAN V1	PETER PHAN V1 after repositioning	PETER PHAN V2	PETER PHAN V2 after repositioning
Excellent	773 (74.8%)	337 (32.6%)	985 (95.3%)	254 (24.6%)
Good	162 (15.7%)	244 (23.6%)	36 (3.5%)	218 (21.0%)
Moderate	74 (7.1%)	258 (24.9%)	8 (0.7%)	278 (26.9%)
Poor	25 (2.4%)	195 (18.9%)	5 (0.5%)	284 (27.5%)

Except for shape features, intrinsically stable due to the fact that the same ROIs are used on all the images, there is no evidence of a class providing more repeatable features than others. From this set of 118 features, we selected the five showing the highest number of recurrences as features with excellent repeatability across the various image type (original and filtered images). These features are '*first-order_Energy*', '*glrlm_RunLengthNonUniformity*', '*glszm_GrayLevelNonUniformity*', '*glszm_SizeZoneNonUniformity*' and '*gldm_DependenceNonUniformity*'. The values of these features extracted from the 16 ROIs of the phantom and from the patients' lesions identified on the original images are represented as a box plot in Fig. 7 (PETER PHAN V1) and Fig. 8 (PETER PHAN V2).

The phantom features show overlap in magnitude with the patients' features in all cases. All the features extracted from the inserts of PETER PHAN V2 (Fig. 8) and considered for this analysis cover a range exceeding the patients' distribution.

For all the ten couples of samples (five for PETER PHAN V1 and five for PETER PHAN V2) the Welch's *t*-test confirmed the null hypothesis of equal mean, with a significance test level of 5%. Therefore, the radiomic features in the phantom inserts are in agreement with the features extracted from the patients' lesions.

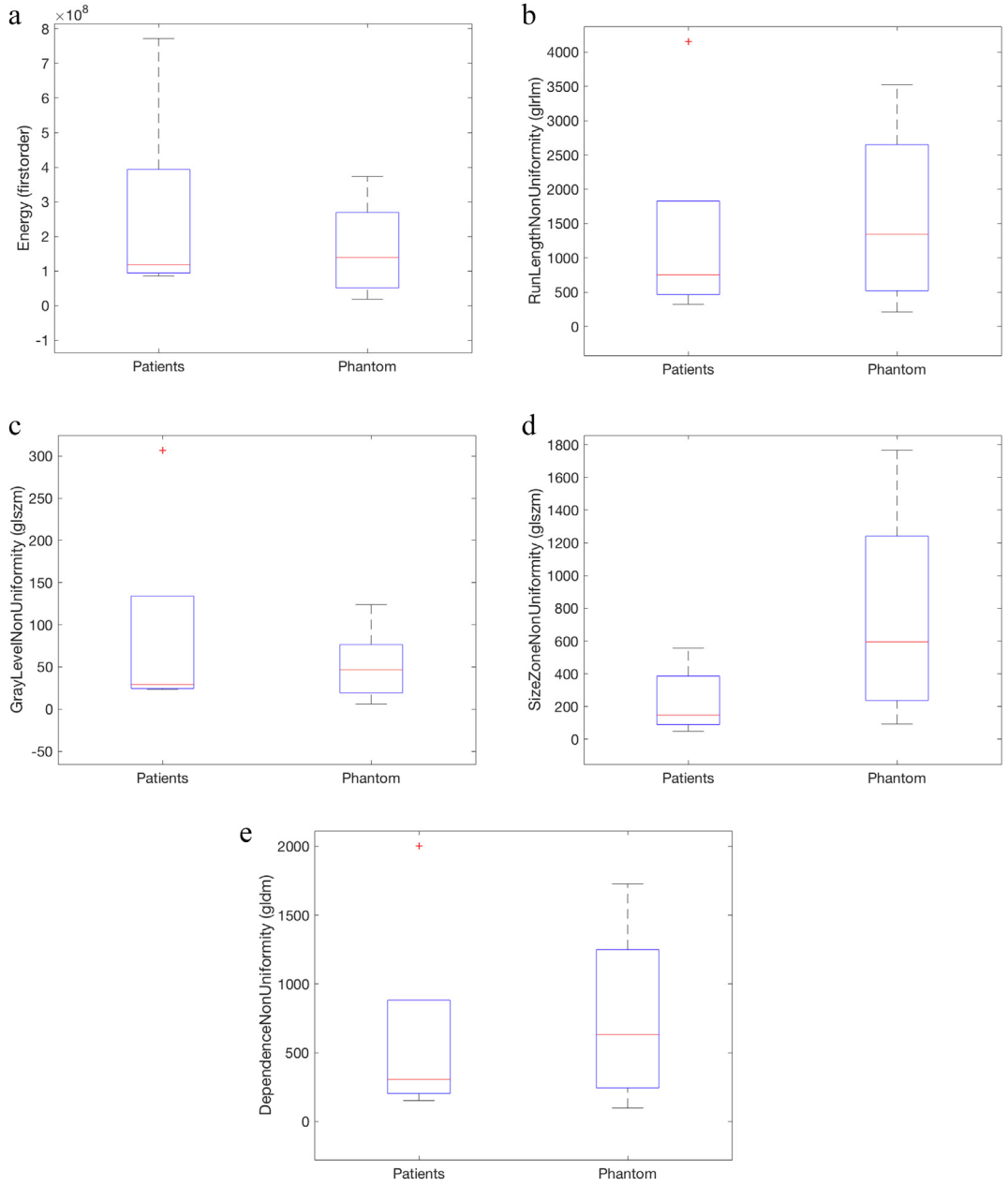

Fig. 7. Each plot (a-e) shows the comparison of the selected radiomic features values extracted (left) from the tumour site of the patients and (right) from the 16 ROIs identified on the phantom inserts. All the images were acquired with a T1-W sequence typically included in the clinical protocol for pelvis diagnostic imaging.

4. Discussion

In developing PETER PHAN, the first phase of the work was the measurement of the relaxation times of both muscle (normal tissue) and tumoural tissue in the case of female patients affected by a pelvic tumour. The maps performed on patients and healthy volunteers indicate a muscle (normal tissue) T_1 to be in the range 806–1378 ms, and T_2 in the range 40–65 ms. As mapping of T_1 and T_2 is not commonly included in the clinical protocols, T_1 and T_2 reference values are limited and not easily available. However, a comparison can be made with the values published by Kato *et al.* [46], that for muscle are in the range 797–1206 ms for T_1 and 31–47 ms for T_2. For the tumour the results obtained in this work are in the range of 1396–2121 ms for T_1 and 79–101 ms for T_2. The values reported in the same reference for miscellaneous tumors and fibrosarcoma are in the range 1011–1083 ms for T_1 and 65–87 ms for T_2. The range indicated do not overlap with our results and the reason can reside in the different nature of the tumours considered for the analysis. Precisely because this work was specifically focused on pelvis tumours, it was decided to measure the relaxation times on patients to obtain the reference values for the phantom development.

The $MnCl_2$ solution concentrations chosen to simulate the muscle tissue resulted in T_1 and T_2 values within 13% and 12% of the corresponding average *in vivo* values. In addition, we incorporated inserts with different textures with which we sought to mimic both the mean signal and the texture observed for lesion during *in vivo* clinical imaging. INS3, consisting of 1 mm diameter polystyrene beads and agar, was the closest in reproducing both these properties, as can be also seen visually in both T_1-W and T_2-W images shown in Fig. 6. Although INS1, INS2 and INS4 were characterized by textures less similar to the clinical images considered in this study, they were introduced with the aim to test the radiomic features ability to distinguish different textures. Moreover, they could be useful for reproducing more heterogeneous lesions observed in populations different from the one considered for

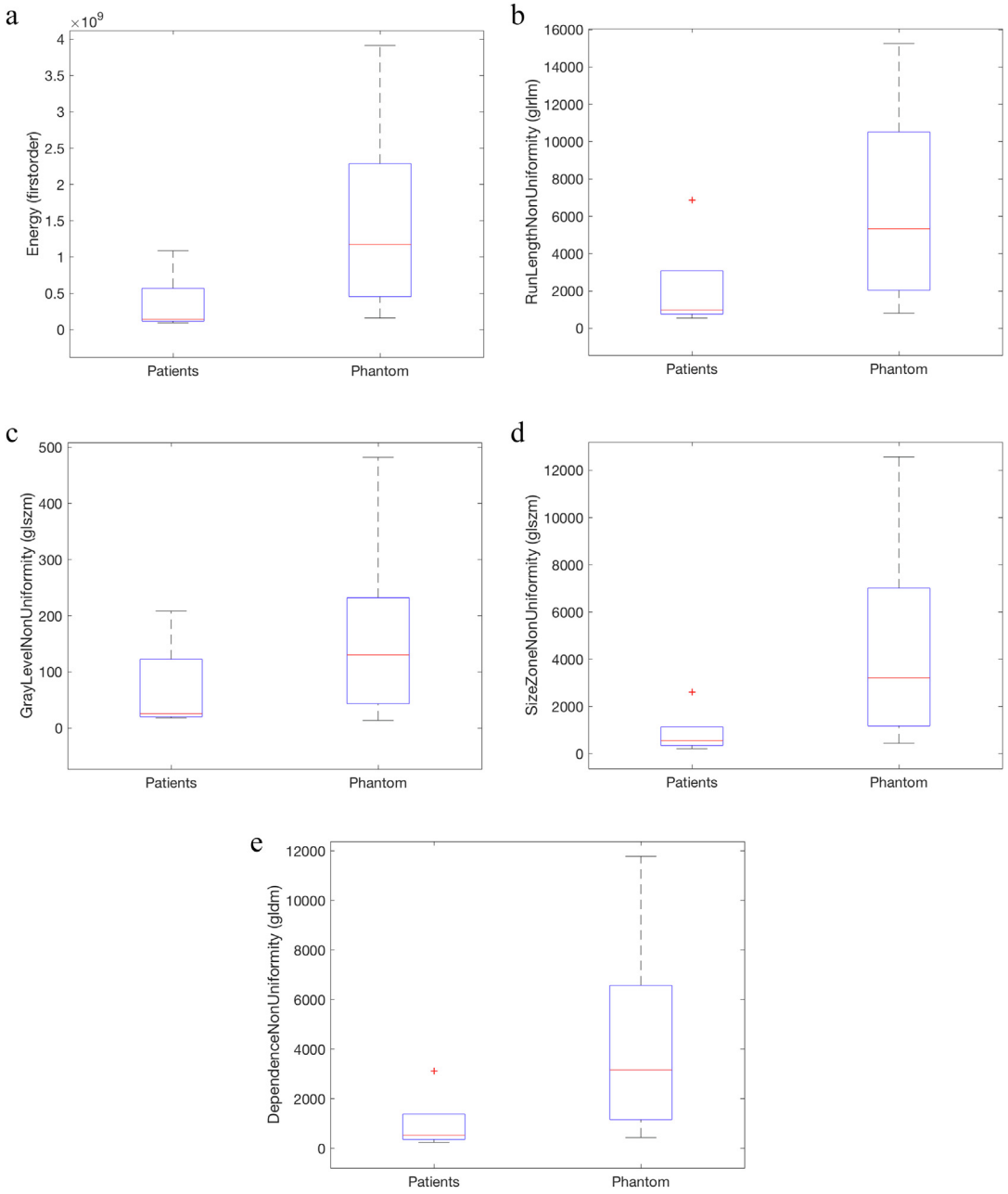

Fig. 8. Each plot (a-e) shows the comparison of the selected radiomic features values extracted (left) from the tumour site of the patients and (right) from the 16 ROIs identified on the phantom inserts. All the images were acquired with a T2-W sequence typically included in the clinical protocol for pelvis diagnostic imaging.

this study.

To provide an example of the possible phantom usage for optimisation of radiomic studies, we showed the results of the investigation of radiomic features on phantom images. We found that lots of features do not show high repeatability when the phantom is repositioned, both in the case of T_1-W and T_2-W images. The reason for this could be differences, despite small, in the phantom orientation in the magnetic field inside the scanner or signal variations before and after repositioning due to a variety of factors (e.g. changes in room temperature and/or in signal-to-noise ratio). These results highlight the importance of conducting a methodological study to support each radiomic analysis on a patients' database, since the features we could identify as predictive of a certain clinical endpoint might not be reliable due, for example, to their insufficient repeatability.

Similarly, the inclusion of unreliable features into a model may introduce confounding factors, preventing the possibility to identify significant associations with clinical outcomes, thus it is fundamental to identify and exclude *a priori* non-robust features.

Selecting five radiomic features among the most repeatable, we proved that they show the same order of magnitude when extracted from the phantom inserts and the tumour volume of the patients considered. Despite the number of patients used for this assessment is limited, this first quantitative evaluation confirms that both phantom versions (V1 and V2) adequately reproduce, for the purpose of a radiomic study, not only the average signal of healthy tissue and tumour, but also the texture seen in a female patient with a pelvic tumour.

Summing up, we successfully created a phantom that mimics the female pelvis and lesions to inspect radiomic properties, and offered a first demonstration of its applicability for the investigation of radiomic features robustness.

The radiomic investigation performed on the present version of

PETER PHAN, as well as its future developments, may allow the exploration of the complex binomial radiomics-MRI in pelvis imaging, contributing to the achievement of a robust methodology to be included in the clinical research. PETER PHAN could be easily assembled and acquired on different scanners and/or at different centres, in order to compare the radiomic response in relation to the scanner vendor, scanner type and magnetic field strength. The radiomic features extracted from the phantom's images will be studied not only for their repeatability but also reproducibility, potential dependence on the MR sequence parameters and on the process of image acquisition in its entirety. Different approaches in image processing will be compared (algorithms for image normalisation or field correction) and different options for the choice of parameters in the features extraction phase will be tested. The possibility to probe the relationship between the ROI volume and the features stability is offered along with the chance to assess the features ability to discriminate between different textures. PETER PHAN represents also a test object to be used for a comparison in the performance of platforms available for the extraction of the radiomic features.

Despite it was originally thought for reproducing the female pelvis, the phantom can easily be extended to the application in male pelvis studies, upon dedicated T_1 and T_2 measurement of male pelvis tumours, or more generally to other cancers occurring in the pelvis district.

A limitation of the phantom is that it was separately optimized for T_1 and T_2. The combined use of two paramagnetic agents may allow the simultaneous matching of T_1 and T_2, which would simplify the use of the phantom. Modifying the relaxation times of the agar gels with similar relaxation agents may also improve the modest agreement with *in vivo* tumour values, seen for T_2 values of interest. Instead, we opted to match image contrast by adjusting the concentration of the polystyrene spheres.

Future versions of the phantom could include inserts with combinations of other sizes of spheres, in order to widen the range of textures reproduced and cover the variability seen in patients. In operational terms, starting from the partial overlap shown in Fig. 7, a new edition of PETER PHAN V1 will aim at reaching a complete overlap of the texture features range offered by the phantom inserts with the range seen in the patients' lesions. Furthermore, the number of patients considered for this evaluation will be increased to make the phantom more representative of the patients' population. From this perspective, also PETER PHAN V2, already showing an optimum overlap with the patients considered, could be optimised.

Additionally, in a recent paper, R. Rai et al. [47] presented a work about an MR texture phantom produced with 3D printing technology, which could be considered for further development and refinement on PETER PHAN inserts, guaranteeing the possibility to realise a variety of tumor-like shapes and textures, in a more reproducible way. The employment of this strategy requires additional investigation on the relaxation times of the materials used as printer ink, in order to develop a potential material (not yet available) that properly reproduces the relaxation of the tumoural tissue of our interest as achieved, despite in a hand-crafted fashion, in PETER PHAN.

5. Conclusions

In this work we have studied the design and assembly of an MRI innovative phantom suitable for radiomic studies of the pelvis. As opposed to standard quality assurance MRI phantoms, PETER PHAN has the double requirement of (i) providing image intensities similar to those of human pelvis tissues on routine clinical scans and (ii) including structures that show heterogeneous signal in the MR image, mimicking the texture properties of a typical pelvic tumour. Given that other MR phantoms dedicated to radiomic analysis do not exist presently, PETER PHAN represents a first inexpensive and accessible tool, that in the future could be improved and refined. Providing the necessary data to build the phantom, this work aims to promote the diffusion of the methodological studies needed in this field.

The phantom represents a valuable tool to support clinical radiomic studies involving pelvic tumours, allowing the assessment of the radiomic features analytical validity. Specifically, the phantom can be used to evaluate the radiomic features repeatability, as shown in this work, and reproducibility. Coupling the specific clinical applications with the methodological analysis by the use of such phantom could guarantee the development of robust and reliable predictive models, taking a step forward in the establishment of the radiomic discipline applied to MR. Further developments will consider the use of more advanced techniques, including 3D printing, and the design of MRI phantoms specific for different anatomical sites for further radiomic approaches in cancer research.

Declaration of Competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Acknowledgements

The authors gratefully acknowledge GE Healthcare for providing the T_2 map sequence run under research mode on the MR scanner, Alfio Cascio for technical support in the phantom assembly and Dr. Mario Pietro Carante for his help with the statistical analysis.

This work was partially supported by the Italian Ministry of Health with Ricerca Corrente and 5x1000 funds.

References

[1] WHO. World Health Statistics Monitoring Health for SDGs. 2018. doi:Geneva: World Health Organization; 2018. Licence: CC BY-NC-SA 3.0 IGO.
[2] Gillies RJ, Anderson AR, Gatenby RA, Morse DL. The biology underlying molecular imaging in oncology: from genome to anatome and back again. Clin Radiol 2010;65:517–21. https://doi.org/10.1016/j.crad.2010.04.005.
[3] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RGPM, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6. https://doi.org/10.1016/j.ejca.2011.11.036.
[4] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5. https://doi.org/10.1038/ncomms5006.
[5] Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. Radiology 2016;278:563–77. https://doi.org/10.1148/radiol.2015151169.
[6] Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: The process and the challenges. Magn Reson Imaging 2012;30:1234–48. https://doi.org/10.1016/j.mri.2012.06.010.
[7] Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. Eur Radiol Exp 2018;2:36. https://doi.org/10.1186/s41747-018-0068-z.
[8] Wibmer A, Hricak H, Gondo T, Matsumoto K, Veeraraghavan H, Fehr D, et al. Haralick Texture Analysis of prostate MRI: Utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason Scores. Eur Radiol 2015;25:2840–50. https://doi.org/10.1007/s00330-015-3701-8.Haralick.
[9] Parekh VS, Jacobs MA. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. npj Breast Cancer 2017;3:1–8. https://doi.org/10.1038/s41523-017-0045-3.
[10] Gevaert O, Mitchell LA, Achrol AS, Xu J, Echegaray S, Steinberg GK, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. 313-313 Radiology 2015;276. https://doi.org/10.1148/radiol.2015154019.
[11] Scalco E, Marzi S, Sanguineti G, Vidiri A, Rizzo G. Characterization of cervical lymph-nodes using a multi-parametric and multi-modal approach for an early prediction of tumor response to chemo-radiotherapy. Phys Medica 2016;32:1672–80. https://doi.org/10.1016/j.ejmp.2016.09.003.
[12] Li Z, Li H, Wang S, Dong D, Yin F, Chen A, et al. MR-Based radiomics nomogram of cervical cancer in prediction of the lymph-vascular space invasion preoperatively. J Magn Reson Imaging 2018:1–7. https://doi.org/10.1002/jmri.26531.
[13] Kan Y, Dong D, Zhang Y, Jiang W, Zhao N, Han L, et al. Radiomic signature as a predictive factor for lymph node metastasis in early-stage cervical cancer. J Magn Reson Imaging 2018:1–7. https://doi.org/10.1002/jmri.26209.
[14] Meng J, Liu S, Zhu L, Zhu L, Wang H, Xie L, et al. Texture Analysis as Imaging Biomarker for recurrence in advanced cervical cancer treated with CCRT. Sci Rep 2018;8:1–9. https://doi.org/10.1038/s41598-018-29838-0.

[15] Zhou X, Yi Y, Liu Z, Cao W, Lai B, Sun K, et al. Radiomics-based pretherapeutic prediction of non-response to neoadjuvant therapy in locally advanced rectal cancer. Ann Surg Oncol 2019. https://doi.org/10.1245/s10434-019-07300-3.
[16] Jeon SH, Song C, Chie EK, Kim B, Kim YH, Chang W, et al. Delta-radiomics signature predicts treatment outcomes after preoperative chemoradiotherapy and surgery in rectal cancer. Radiat Oncol 2019;14:43. https://doi.org/10.1186/s13014-019-1246-8.
[17] Meng X, Xia W, Xie P, Zhang R, Li W, Wang M, et al. Preoperative radiomic signature based on multiparametric magnetic resonance imaging for noninvasive evaluation of biological characteristics in rectal cancer. Eur Radiol 2018. https://doi.org/10.1007/s00330-018-5763-x.
[18] Rizzo S, Femia M, Buscarino V, Franchi D, Garbi A, Zanagnolo V, et al. Endometrial cancer: an overview of novelties in treatment and related imaging keypoints for local staging. Cancer Imaging 2018;18:1–12. https://doi.org/10.1186/s40644-018-0180-6.
[19] Rizzo S, Buscarino V, Origgi D, Summers P, Raimondi S, Lazzari R, et al. Evaluation of diffusion-weighted imaging (DWI) and MR spectroscopy (MRS) as early response biomarkers in cervical cancer patients. Radiol Medica 2016;121:838–46. https://doi.org/10.1007/s11547-016-0665-y.
[20] Rizzo S, Botta F, Raimondi S, Origgi D, Buscarino V, Colarieti A, et al. Radiomics of high-grade serous ovarian cancer: association between quantitative CT features, residual tumour and disease progression within 12 months. Eur Radiol 2018;28:4849–59. https://doi.org/10.1007/s00330-018-5389-z.
[21] Mayerhoefer ME, Szomolanyi P, Jirak D, Materka A, Trattnig S. Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: an application-oriented study. Med Phys 2009;36:1236–43. https://doi.org/10.1118/1.3081408.
[22] Larue RTHM, Defraene G, De Ruysscher D, Lambin P, Van Elmpt W. Quantitative radiomics studies for tissue characterization: A review of technology and methodological procedures. Br J Radiol 2017; 90. Doi: 10.1259/bjr.20160665.
[23] Bane O, Hectors SJ, Wagner M, Arlinghaus LL, Aryal MP, Cao Y, et al. Accuracy, repeatability, and interplatform reproducibility of T 1 quantification methods used for DCE-MRI: results from a multicenter phantom study. Magn Reson Med 2018;79:2564–75. https://doi.org/10.1002/mrm.26903.
[24] Hellerbach A, Schuster V, Jansen A, Sommer JMRI. Phantoms – are there alternatives to agar? PLoS ONE 2013;8. https://doi.org/10.1371/journal.pone.0070343.
[25] Baeßler B, Weiss K, Dos Santos DP. Robustness and reproducibility of radiomics in magnetic resonance imaging: a phantom study. Invest Radiol 2019;54:221–8. https://doi.org/10.1097/RLI.0000000000000530.
[26] Niebuhr NI, Johnen W, Echner G, Runz A, Bach M, Stoll M, et al. The ADAM-pelvis phantom - an anthropomorphic, deformable and multimodal phantom for MRgRT. Phys Med Biol 2019;64.. https://doi.org/10.1088/1361-6560/aafd5f.
[27] Yang F, Dogan N, Stoyanova R, Ford JC. Evaluation of radiomic texture feature error due to MRI acquisition and reconstruction: a simulation study utilizing ground truth. Phys Medica 2018;50:26–36. https://doi.org/10.1016/j.ejmp.2018.05.017.
[28] Ford J, Dogan N, Young L, Yang F. Quantitative radiomics: impact of pulse sequence parameter selection on mri-based textural features of the brain. Contrast Media Mol Imaging 2018;2018. https://doi.org/10.1155/2018/1729071.
[29] Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor Brian, et al. Measuring computed tomography scanner variability of radiomics features. Invest Radiol 2015;50:757–65. https://doi.org/10.1097/RLI.0000000000000180.
[30] Forgacs A, Pall Jonsson H, Dahlbom M, Daver F, Difranco MD, Opposits G, et al. A study on the basic criteria for selecting heterogeneity parameters of F18-FDG PET images. PLoS ONE 2016;11:1–14. https://doi.org/10.1371/journal.pone.0164113.
[31] Messroghli DR, Radjenovic A, Kozerke S, Higgins DM, Sivananthan MU, Ridgway JP. Modified look-locker inversion recovery (MOLLI) for high-resolution T 1 mapping of the heart. Magn Reson Med 2004;52:141–6. https://doi.org/10.1002/mrm.20110.
[32] Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, et al. Magnetic resonance fingerprinting. Nature 2013;495:187. https://doi.org/10.1002/mrm.27558.
[33] Deoni SCL, Rutt BK, Peters TM. Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state. Magn Reson Med 2003;49:515–26. https://doi.org/10.1002/mrm.10407.
[34] Ramos-Llorden G, Vegas-Sanchez-Ferrero G, Bjork M, Vanhevel F, Parizel PM, San Jose Estepar R, et al. NOVIFAST: A Fast Algorithm for Accurate and Precise VFA MRI T 1 Mapping. IEEE Trans Med Imaging 2018; 37: 2414–27. doi: 10.1109/TMI.2018.2833288.
[35] Thangavel K, Saritaş EÜ. Aqueous paramagnetic solutions for MRI phantoms at 3 T: A detailed study on relaxivities. Turkish J Electr Eng Comput Sci 2017;25:2108–21. https://doi.org/10.3906/elk-1602-123.
[36] Hahn EL. Spin echoes. Phys Rev 1950;80:580.
[37] Carr Herman Y, Purcell EM. Effects of diffusion on free precession in nuclear magnetic resonance experiments. Phys Rev 1954;94:630.
[38] Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. Rev Sci Instrum 1958;29:688–91. https://doi.org/10.1063/1.1716296.
[39] Jirák D, Dezortová M, Hájek M. Phantoms for texture analysis of MR images. Long-term and multi-center study. Med Phys 2004;31:616–22. https://doi.org/10.1118/1.1646231.
[40] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 2012;30:1323–41. https://doi.org/10.1016/j.mri.2012.05.001.
[41] Van Griethuysen JJ, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017;77:e104–7. https://doi.org/10.1016/j.physbeh.2017.03.040.
[42] PyRadiomics documentation, https://pyradiomics.readthedocs.io; 2016 [accessed 9 December 2019].
[43] Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
[44] Samei E, Hoye J, Zheng Y, Solomon JB, Marin D. Design and fabrication of heterogeneous lung nodule phantoms for assessing the accuracy and variability of measured texture radiomics features in CT. 021606 J Med Imaging 2019;2. https://doi.org/10.1117/1.
[45] Welch BL. The generalization ofstudent's' problem when several different population variances are involved. Biometrika 1947;34:28–35. https://doi.org/10.1093/biomet/34.1-2.28.
[46] Kato H, Kuroda M, Yoshimura K, Yoshida A, Hanamoto K, Kawasaki S, et al. Composition of MRI phantom equivalent to human tissues. Med Phys 2005;32:3199–208. https://doi.org/10.1118/1.2047807.
[47] Rai R, Wang YF, Manton D, Dong B, Deshpande S, Liney GP. Development of multipurpose 3D printed phantoms for MRI. 075010 Phys Med Biol 2019;64. https://doi.org/10.1088/1361-6560/ab0b49.