Published text:

Position Paper

Evaluation of response after pre-operative radiotherapy in soft tissue sarcomas; the European Organisation for Research and Treatment of Cancer — Soft Tissue and Bone Sarcoma Group (EORTC — STBSG) and Imaging Group recommendations for radiological examination and reporting with an emphasis on magnetic resonance imaging

C. Messiou a,*,1,2, S. Bonvalot b,1, A. Gronchi c, D. Vanel d, M. Meyer e,2, P. Robinson f, C. Morosi g, J.L. Bloem h,2, P.H. Terrier i, A. Lazar j, C. Le Péchoux k, E. Wardelman l, J.M. Winfield m, B. Boulet i, J. Bovée n, R.L. Haas o

a Department of Radiology, The Royal Marsden Hospital, London, UK
b Department of Surgery Institut Curie, Paris, France
c Department of Surgery, Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy
d Department of Radiology, The Rizzoli Institute, Bologna, Italy
e Institute of Clinical Radiology and Nuclear Medicine, University Medical Centre Mannheim, Medical Faculty Mannheim — Heidelberg University, Mannheim, Germany
f Department of Radiology, Leeds Teaching Hospitals NHS trust, Leeds, UK
g Department of Radiology, Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy
h Department of Radiology, The Leiden University Medical Center, Leiden, The Netherlands
i Institut Gustave Roussy, Paris, France
j University of Texas MD Anderson Cancer Center, Houston, TX, USA
k Department of Radiotherapy Institut Gustave Roussy, Paris, France
l Gerhard-Domagk-Institute of Pathology of the University Hospital, Münster, Germany
m Department of Physics, The Royal Marsden Hospital, London, UK
n Department of Pathology, Leiden University Medical Centre, Leiden, The Netherlands
o Department of Radiotherapy, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Received 13 October 2015; accepted 8 December 2015
Available online 20 January 2016
KEYWORDS
Sarcoma;
MRI;
Radiotherapy

Abstract
At present, there is no standardised approach for the radiological evaluation of soft tissue sarcomas following radiotherapy (RT). This manuscript, produced by a European Organisation for Research and Treatment of Cancer—Soft Tissue and Bone Sarcoma Group (EORTC—STBSG) and Imaging Group endorsed task force, aims to propose standardisation of magnetic resonance imaging techniques and interpretation after neoadjuvant RT for routine use and within clinical trials.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction
For patients with soft tissue sarcomas (STS) an en-bloc resection with negative margins will usually provide long-term local control in several patients even without preoperative or postoperative radiotherapy (RT) [1]. However for the management of the more aggressive STS subtypes (especially of intermediate to high grade histology, larger than 5 cm and/or deep seated), surgery alone results in inferior local control rates as compared to surgery in combination with RT [2]. Historically but also based upon the Canadian SR-2 trial [3], for preoperative RT, a conventionally fractionated RT regimen of once-daily 1.8–2 Gy up to a total dose of 50–50.4 Gy over 5 weeks is chosen. The absence of a standardised approach to response assessment in this setting impacts not only clinical practice but also clinical trials investigating novel RT delivery. Regimens applying both hyperfractionation (smaller fraction doses usually prescribed twice daily) and hypofractionation (greater than 2 Gy per fraction [4–6]) are under investigation. Also, in an attempt to sensitise sarcoma cells and/or sarcoma stroma to radiation, chemotherapy and targeted agents are being applied concurrently with RT, such as antracycline/ifosfamide [7,8], bevacizumab [9], sorafenib [10,11], sunitinib [12], pazopanib [13] and more recently nanoparticles [14]. In order to appreciate the true value of these new regimens a standardised and reproducible radiological evaluation is critical.

This manuscript which focuses on the use of magnetic resonance imaging (MRI) for assessing response in STS following RT is targeted at radiologists and clinical oncologists but will serve as a guide to all health professionals in the STS community.

2. Imaging objectives
In an era of evolving targeted RT techniques such as intensity modulated RT and image guided RT, imaging is crucial for achieving accurate treatment volumes with sparing of adjacent normal tissues at risk of toxicity. STS are notorious for their marked inter- and intra-tumoural heterogeneity and some investigators are already starting to interrogate the clinical value of “dose painting” regimens where non-uniform radiation is distributed based on functional or molecular imaging [15]. RT planning is not within the remit of this article, however the MRI protocol before and after RT should be identical and performed on the same MRI scanner. Following RT, imaging can be used not only to assess response but also to revisit surgical staging as increases in tumour size secondary to oedema, haemorrhage and necrosis are not uncommon [16] although this does not influence local control rates [17].

2.1. Timing of post RT MRI studies
Currently imaging during and immediately post RT should be avoided as the complex imaging features in this setting can be misleading (Fig. 1). Imaging should be performed as close to the surgical date as possible and most clinical protocols assume that an interval for surgery of 4–6 weeks following RT is optimal.

2.2. Routine MRI protocol
For limb sarcomas a combination of T2-weighted (W), T1-W, T1-W with fat saturation, T1-W fat-saturated (FS) contrast-enhanced and short-tau inversion recovery (STIR) should be used (Fig. 2). Some institutions may choose to use T2 with fat saturation or Dixon water T2-W in preference to STIR. These sequences can be supplemented by diffusion weighted MRI (DWI) sequences (b 50, 600 and 900 s/mm²) and apparent diffusion coefficient (ADC) maps depending on local expertise. Coverage should include the joint above and below the sarcoma. Choice of imaging for truncal STS is site dependent but a similar MRI protocol can be used for masses in the anterior abdominal wall for example. Imaging may be performed either at 1.5T or 3T. At other sites such as the mediastinum, contrast enhanced computed tomography (CT) can be more suitable since CT is less prone to motion artefacts. For retroperitoneal sarcomas most centres use CT but the MRI protocols used for limb sarcomas can easily be translated to the retroperitoneum. This can be particularly helpful to address targeted questions regarding local invasion to structures such as psoas, neural foramina and bone and will become increasingly
relevant if the practice of neoadjuvant RT for retroperitoneal sarcomas increases. Careful attention should be made to ensure reproducible patient positioning and to avoid deforming the contours of the tumour i.e. patients with buttock tumours should be imaged prone.

2.3. Assessing response

Histopathological changes including necrosis, cystic change, haemorrhage, hyalinisation and fibrosis which occur following RT may influence dimension based assessments of response significantly. In fact with the exception of myxoid liposarcomas [18] (Fig. 3) significant dimensional radiologic responses after preoperative RT are rare events and have been reported as low as 0% [19]. Miki et al [20] showed that 31% of tumours increased in size by more than 10% but this was not associated with a deterioration in local recurrence free survival, event free survival or overall survival. Look et al [16] failed to show any correlation with Response

Fig. 1. Pseudoprogression post radiotherapy. Contrast enhanced T1-W MRI with fat saturation of an anterior compartment pleomorphic sarcoma before (a), 6 weeks following radiotherapy (b) and 10 weeks following radiotherapy (c). At 6 weeks post radiotherapy (b) the increase in size and enhancement gives a false positive for tumour progression. Surgery was delayed for medical reasons and MRI at 10 weeks confirms tumour regression. Five to 10% viable tumour was present in the resection specimen.

Fig. 2. Recommended clinical MRI protocol. Axial T1-W (a), T2-W (b), T1 fat saturated (FS)(c), T1 FS+ contrast (d); coronal STIR (e); DWI b50(f), b900 (g), ADC map (h). Note that for morphological imaging both limbs are included but DWI suffers extreme artefacts if both limbs are included.
Evaluation Criteria in Solid Tumours (RECIST) and outcome measures and demonstrated that tumours could show significant reductions in size despite demonstrating predominantly viable tumour whereas stable or growing tumours could show dramatic histopathological response. In fact growing tumours have been shown to have a higher incidence of mosaic high and low T2 signal changes on MRI compared to tumours with stable or reducing size which correlated with cystic change and haemorrhage, respectively on histopathology [20]. Le Grange [17] reported that RECIST was also a poor reflection in overall changes in tumour volume classifying 89% of tumours as stable disease despite 80% of tumours reducing in volume.

The most reliable measure of tumour response to pre-operative RT in STS has not yet been determined. Efforts should be directed towards distinguishing non-viable elements from visible tumour and new surrogate end-points for pathological response should be defined. Such an approach has already been taken for gastrointestinal stromal tumours where Choi criteria incorporate tumour attenuation changes as well as size [21]. The pilot study by Stacchiotti et al [22] has provided preliminary evidence that this may also be useful in a post chemoradiotherapy setting for STS. However in synovial sarcoma they were not able to differentiate the cystic component of tumour from treatment effect. In such cases functional imaging techniques may provide additional information. A small study of fluorodeoxyglucose (FDG) positron-emission tomography/CT in high-grade sarcomas confirmed the inability of CT volumetric measurements to identify histopathological responders but did show some promise for FDG uptake measurements of standardised uptake value (SUV) max and SUV mean [23].

It is important to recognise that in current clinical practice in the absence of metastatic disease, appearances on post RT imaging are unlikely to alter the decision to operate. Furthermore, neoadjuvant RT is very unlikely to render an inoperable STS resectable [24]. This imaging primarily serves to guide operative management particularly in cases where the STS has increased in size. However recommended guidelines for response assessment of STS to RT in routine clinical practice include:

1. Post RT imaging should not be performed earlier than 4 weeks post RT (later if possible).
2. Images acquired in the same plane should be performed with identical planning and slice thickness to allow correlation between sequences.
3. With the exception of myxoid liposarcomas, size and volume measurements should not be used to reflect histopathological response.
4. Internal signal/density characteristics should be used in combination to assess response. For example diminished enhancement and reduction in size of restricted components/rising ADC on DWI may be interpreted as response (Fig. 4).
5. Areas of new enhancement should be interpreted with caution as this can arise secondary to vascular disruption following RT and does not necessarily reflect progression.
6. Not all areas of diminished enhancement following RT represent necrosis and therefore attention to terminology is suggested. The term ‘treatment effects’ may be more appropriate encompassing several processes including necrosis, cystic change (liquefaction), hyalinisation etc.

For clinical trials where novel regimens are under interrogation quantitative measures may be desirable and this will be discussed in the following section.

2.4. Functional imaging and response assessment in clinical trials

As discussed previously dimensional assessments of response are not appropriate in this setting. In other
words RECIST should not be used to evaluate response to neoadjuvant management of STS. STS can be usefully interrogated by multiparametric MRI but both intertumoural and intratumoural heterogeneity can mask treatment effects if data are analysed for cohort effects. It seems reasonable to expect that different sarcoma subtypes may behave differently and therefore subgroup analysis is strongly recommended. Repeatability and reproducibility of different functional MRI parameters is also profoundly influential and therefore for quantitative studies the ideal and most robust approach is for each patient to undergo two baseline studies thereby defining their own repeatability [25–29]. This is also an approach which can be used in multicenter trials as repeatability is a critical factor influencing the decision to pool data although this does of course have resource and time implications. Repeatability studies are also important for future proofing data for post hoc analysis, which is particularly important for rare tumour types where data is scarce.

Although the proposed link between ADC derived from DW MRI and cellularity makes this arguably the most promising parameter it is important to remember that it is likely to be influenced by many factors including water motion in different compartments, microarchitecture and heterogeneity, membrane integrity, presence of macromolecules leading to water binding, the cellular volume fraction, active transport as well as the MRI acquisition parameters and differences in MRI scanner hardware. Although there is paucity of sarcoma specific data it remains important to notice that the inverse correlation between cell density in soft tissues and ADC has not been so impressive in all cell types [30,31]. It is likely therefore that ADC is a complex function of tissue microarchitecture which is influenced by several components. A significant advantage of ADC is speed of acquisition and good reproducibility with a reported coefficient of variation of 4.8% in one patient study [32] and slightly greater in volunteer studies of different organs with CVs ranging from 7 to 16.9% outside of the liver [33]. Choice of b values has been shown to influence ADC in several tissues [34,35]. Consensus recommendations for DW-MRI as a cancer biomarker recommend that protocols should be optimised to maximise signal-to-noise ratio, minimise artefacts from ghosting and distortion, optimise fat suppression, ensure ADC values can be measured accurately and reproducibly and, ideally, use parameters which can be replicated on other platforms [27]. However development of protocols for multicenter studies necessitates additional trade-offs between optimisation and standardisation of protocols on different platforms. Some parameters are difficult or impossible to standardise such as diffusion gradient scheme, diffusion gradient strengths and timings, and methods of fat suppression and parallel imaging. It is not possible to standardise the method of fat suppression across scanners from different manufacturers although spectral methods may be employed on all scanners. Spectral methods may be preferable over inversion recovery as the latter reduces the overall signal and introduces T1 weighting. However STIR may be advantageous where B0 homogeneity is poor.

It is likely that for any multicenter imaging study, a lead site will be nominated. It is the responsibility of this site to disseminate a final protocol, taking into account the capabilities of scanner types and field strengths at participating sites. Sites should agree on aspects of the protocol that should be fixed e.g. slice thickness, orientations, b-values for DWI, TEs for R2*. Some aspects
may have to be different between scanners so that images are optimised from all scanners e.g. diffusion encoding gradient schemes (monopolar versus double spin-echo). The process of producing the final protocol for a study will require development work and the challenges of this may depend on the scanners involved and the anatomy being imaged. This also includes defining what common elements are possible. It is important that this process should be built into the schedule and budget of a study as this is too often overlooked.

Donati et al [33] have addressed the questions surrounding the use of different field strengths and vendor platforms in DWI multicenter studies. Critically they found no significant difference between ADCs at different field strengths but they did find that vendor variability had greater influence at 3T. Additionally the findings of the Innovative Medicines Initiative QuIC-ConCePT (Quantitative Imaging in cancer: Connecting cellular Processes with Therapy) project are eagerly awaited. This aims to qualify imaging biomarkers including DW MRI and a standardisation procedure of diffusion MRI is being established in collaboration between QuIC-ConCePT in Europe and QIBA (quantitative imaging biomarker association) in the United States.

Ice-water phantoms containing water or sucrose solutions have previously been used to compare ADC estimates between scanners [36,37]. Ice-water provides a simple and inexpensive method to control temperature. Sucrose restricts diffusion of water molecules and can be used to reduce the ADC of water to values comparable to ADCs observed in vivo. Phantoms containing sucrose solutions have also been used to assess long-term repeatability of ADC estimates [38]. The use of a large Field of View (FOV) imposes an additional requirement for good homogeneity of ADC estimates across the FOV in addition to accurate ADC estimates near the isocentre. Although phantoms are likely to have an important role in multicenter studies they are unlikely to replace the usefulness of repeatability studies in subjects.

It has become apparent that ADC is influenced also by perfusion effects at low b values, which can be interrogated by more sophisticated non-monoeponential models such as intravoxel incoherent motion (IVIM) and stretched exponential [39]. However our understanding of the physiology which underpins these measurements is limited.

Although experience of ADC metrics in other tumours is now quite advanced, the presence of lipid and myxoid elements in many STS will mean that the most sensitive metrics will have to be redifined. The role of diffusion weighted imaging in assessing response to RT in STS is yet to be fully explored. Prospective studies with histopathological correlation and outcome data are much needed.

Other options for exploring tumour vascularity include enhancing fraction or more formal measures of perfusion from dynamic contrast enhanced MRI (DCE-MRI) such as K\text{trans} which is thought to reflect microvascular permeability. These have intrapatient coefficients of variation of 8.6% and 13.9%, respectively [26]. A major advantage of enhancing fraction over DCE MRI is whole tumour coverage in the case of large tumours. R2* measures of hypoxia from blood oxygen level dependent (BOLD) MRI are becoming increasingly relevant as hypoxia is a key mechanism leading to radioresistance. Hypoxia mapping techniques therefore have great potential for improving dose delivery [40]. In preclinical fibrosarcoma models R2* studies provided a completely non-invasive prognostic indicator of radiotherapeutic response but reproducibility in the clinical studies is not well defined [41]. Recommended parameters for reporting in clinical trials are listed as follows:

2.5. Recommended parameters for reporting multiparametric MRI in clinical trials (optional parameters in italics)

- Maximum axial dimension (mm)
- Tumour volume (mm$^3$)
- ADC: mean, median, (10th, 25th, 75th, 90th percentiles, skew and kurtosis.)
- Contrast enhanced MRI: enhancing fraction
- $\pm K\text{trans}$, iAUC$_{60}$, $K_{ep}$, $V_e$, $V_p$.
- Non-monoeponential models for example D, f and D* from IVIM.
- \textit{BOLD: R2*}

For other soft tissue tumours either high b value DWI images or contrast-enhanced images are most often selected for drawing regions of interest (ROIs). However for STS elements such as fat and myxoid components may not be identifiable on such sequences and therefore T2-W sequences may be more appropriate for drawing ROIs with use of other sequences for correlation.

Although multiparametric MRI is extremely promising currently data analysis times are prohibitive. Drawing ROIs on image slices through the whole tumour volume at multiple time points, in addition to post processing and data analysis is time consuming. Advances in data informatics and workflow design are much needed to enable translation of these technologies from small to larger scale use.

Until further validation studies are available, use of multiparametric MRI in clinical trials should be regarded as exploratory. The majority of planned and current neoadjuvant studies have the benefit of resection specimens which can be used for histopathological assessment of response as primary or secondary outcome measures along with progression free survival or overall survival.
2.6. Recommendations for multiparametric MRI in clinical trials

1.5T is preferable for planning multicentre studies. 3T is possible but greater vendor variability should be considered.

Subgroup analysis of broad histopathological subtypes is recommended where possible.

The MRI protocol should include clinical sequences supplemented with DWI and quantitative contrast enhanced MRI (enhancing fraction +/- DCE MRI with high temporal resolution of at least 3 s). Non-monoexponential DWI models and BOLD MRI are interesting but optional.

Coverage should include the entire tumour. For large tumours this is not possible for all sequences such as DCE MRI and BOLD but in such circumstances at least three slices of data should be acquired and analysed.

Ideally repeatability should be established using two baseline studies for each patient performed within 7 days of each other. This is feasible for non contrast enhanced components such as DW MRI and BOLD however the invasive nature of DCE MRI will usually preclude this.

ROIs should be drawn throughout the whole tumour volume on T2-W images with use of other sequences for visual correlation.

Recommended b values for DWI are 50, 600, 900 s/mm² with spectral fat saturation techniques.

Recommended b values for non-monoexponential models of DWI are 0, 25, 50, 80, 150, 300, 500, and 800 s/mm².

Recommended TEs for BOLD: 5, 10, 30, 45, and 55 ms

3. Discussion

Here, to the best of our knowledge for the first time in the literature, a systematic process for MRI of STS following neoadjuvant RT is described. To date, radiological series post RT have used heterogeneous methodology, which makes comparisons extremely challenging.

The next research strategy will be to correlate the imaging characteristics after completion of neoadjuvant management of STS with pathological phenomena in the definitive resection specimen and also with local control rates and overall survival. Until such a time, functional imaging in clinical trials of STS should be regarded as exploratory. This should take place within prospective studies, using double baseline and optimised MR protocols with broad stratification of sarcoma subtypes. Power calculations to determine the number of patients should make use of repeatability data from dual baseline and effect size measured in pilot studies. Whole volume coverage may be difficult to achieve in large tumours and it may be necessary to adopt protocols that include whole volume coverage for some measurements e.g. ADC and local measurements for some others, e.g. R2*. Once the pathological data can be reliably predicted by preoperative imaging, the sarcoma scientific community can start to implement these read-outs in individual patient care. Hopefully in the future, all these investigations and techniques, can guide us either to stop an ineffective neoadjuvant therapy if so designated, or to de-intensify a regimen in case of shown efficacy on-treatment, in an attempt to decrease both postoperative morbidity and long term sequellae.

For this purpose, this multidisciplinary task force, came to consensus on MRI evaluation of response after neoadjuvant RT in STS.

Role of a funding source

Christina Messiou is supported by NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden Hospital.

Conflict of interest statement

RH: GSK/Novartis and Nanobiotix Company research grants to the institution;
SS: Pharmamar research grants to the institution. None of the other authors have declared conflicts of interest.

Acknowledgements

The authors kindly acknowledge the Nanobiotix Company, Paris, France, for organising a multidisciplinary workshop, as a result of which this consensus paper was produced. However, Nanobiotix had no part whatsoever in writing this manuscript.

References


