Promises and challenges of positron emission tomography for assessment of sarcoma in daily clinical practice

A.C.M. van de Luijtgaarden, J.W.J. de Rooy, L.F. de Geus-Oei, W.T.A. van der Graaf and W.J.G. Oyen

Abstract

A correct histological diagnosis, careful staging and detection of tumour response to treatment are all crucial in the management of sarcomas. Imaging is important in all of these stages. Sarcomas have distinct biological and treatment-related features posing challenges for imaging. For example, size measurements may not adequately reflect response rates. Techniques which can measure tissue function rather than generate merely anatomical data such as positron emission tomography (PET) are rapidly gaining interest. We discuss the importance of imaging in different stages of patient management, emphasising the unique characteristics of sarcoma. Furthermore, we discuss the potential of PET for the various indications, focussing on therapy evaluation.

Keywords: Diagnostic imaging; radionuclide imaging; positron emission tomography; sarcoma; gastrointestinal stromal tumors; therapeutics.

Introduction

Sarcomas are a rare and heterogeneous group of mesenchymal-derived tumours with distinct molecular features. They are subclassified into bone and soft-tissue sarcoma (STS). The latter group consists of over 50 subtypes including gastro-intestinal stromal tumour (GIST)\(^1\). In STS, treatment consists of surgery and in selected subtypes or stages of disease also of radiotherapy and/or chemotherapy. In bone sarcoma, such as Ewing’s sarcoma and osteosarcoma, treatment always includes (neo)-adjuvant chemotherapy. Targeting underlying molecular events may provide spectacular benefits, as demonstrated in GIST and dermatofibrosarcoma protuberans (DFSP)\(^2-5\). Prognosis depends on the extent of the disease, requiring optimal staging, and the possibility for radical resection of the primary tumour. Prognosis drops dramatically once the sarcoma is metastatised or worsens, in case of Ewing’s sarcoma and osteosarcoma, if the histological response to neoadjuvant chemotherapy is limited. Early, adequate therapy evaluation prevents prolonged exposure to toxic yet ultimately unsuccessful treatment, which in some cases may be substituted by an alternative, more effective one. An ideal evaluation method should provide information in an objective and reproducible fashion. At present, especially since the introduction of targeted therapies, the call for functional rather than mere anatomical imaging is increasing.

Diagnosis, grading and staging

Histological classification is a crucial first step in suspected sarcomas since tumour type and grade have
Response evaluation and restaging

Available imaging modalities for therapy evaluation are essentially the same as in staging and grading and the choice is guided by tumour or metastasis localisation. Bone scintigraphy, however, is not sufficiently specific for assessment of response. Much work has been done to standardise the interpretation of radiological evaluation methods on the basis of size. This resulted in the World Health Organization (WHO) criteria, later replaced by the simplified European Organisation for Research and Treatment of Cancer Response Evaluation Criteria in Solid Tumours (EORTC/RECIST). These criteria have imperfections, both in general and specifically for sarcoma. First, they were originally based on the change in tumour size which could reliably be detected by palpation. Therefore marked size reduction is required before a tumour is considered responsive to therapy. Although much smaller changes can now be detected, these RECIST criteria are still adhered to because of a supposed relationship between tumour size reduction and clinical benefit. The inter- and intratumoural heterogeneity and the rarity of sarcomas have prevented a reliable scientific foundation of the relationship between size and effects, although there are several reasons to question whether such a relationship exists.

First, sarcomas differ from other tumours as they contain large volumes of non-malignant cells and other stromal materials, which maintain a certain size even if all malignant components disappear. Also, with the disappearance of tumour cells, rather than shrinkage, replacement with fibrous materials or calcification can be seen. In bone sarcomas, surrounding bone has limited capacity to return to its normal size (Fig. 2). In GIST, progression under treatment may present as a nodule within a cystic mass, instead of mass enlargement. Finally, the development of targeted tyrosine kinase inhibitors and antibodies needs consideration. These treatments are cytostatic rather than cytoreductive and hence cause consolidation rather than reduction of tumour size.

In summary, the large changes in size required by WHO and RECIST criteria may be too stringent for the detection of progression or response in sarcoma. Furthermore, size reduction takes time. Thus, identification of tumour response or lack thereof may require several weeks or months. This time delay causes unnecessary costs and toxicity and may prevent a timely switch to alternative treatments. Moreover, the absence of progression of disease is increasingly considered as a relevant endpoint for clinical trials, replacing response rate.

To overcome the possibility that size does not represent tumour viability, histological evaluation can be performed. Histological evaluation for therapy-related changes, including the percentage of necrosis, provides adequate insight into the response to previous therapy and correlates well with prognosis. Limitations are that standardised approaches are only available for osteosarcoma and Ewing’s sarcoma. Heterogeneity of the tumour might again hamper a representative and reliable histological assessment based on small biopsies.

New developments

Improvements in anatomical imaging and interpretation

The above-mentioned restrictions of anatomical imaging during therapy are not only challenging for existing criteria, they are also demanding for the development of new ones. The appearance of a nodule within a mass has been proposed as a sign of recurrent GIST. Choi and colleagues have shown that a 10% tumour size decrease or 15% tumour density decrease as determined by measuring the attenuation coefficient are sensitive and specific methods for assessing targeted therapy response in patients with GIST (the so-called ‘Choi criteria’) (Fig. 1). Dynamic contrast-enhanced MRI (DCE-MRI) is a technique sensitive to alterations in
vascular permeability and blood flow. It has been reported as a sensitive imaging method for the evaluation of response to chemotherapy and it might help in differentiating viable tumour from vascularised granulation tissue\(^{13,33}\). It is, however, not widely used in clinical practice, as it is labour intensive and technically challenging\(^{34}\). These approaches incorporate an increasingly popular concept of imaging; i.e. it is not (only) size that matters, but more so the underlying tissue function, cell biology, physiology and biochemistry.

**Positron emission tomography**

Positron emission tomography (PET) is a technique with large potential because of its ability to image biological characteristics based on the differential utilisation of various substrates by cancer cells and normal tissue. Numerous PET radiopharmaceuticals are available, but the most widely used agent is \(^{18}\)F-fluorodeoxyglucose (FDG). This lead position is in part due to its approval as a tracer by the US Food and Drug Administration (FDA) for routine clinical use\(^{35}\), its early development, and wide availability\(^{36}\). There is a clear rationale for its use. Mammalian cells depend on glucose as a major source of energy and of carbons. Glucose is transported into the cell via facilitative transporters (GLUT) present in all cell types. Many GLUT isoforms exist with tissue-specific expression, subject to environmental control (e.g. hypoxia)\(^{37}\). After membrane transport, glucose is phosphorylated by hexokinases to glucose 6-phosphate and is further metabolised in the glycolysis pathway. Increased glucose utilisation of malignant cells has been recognised for decades. Like glucose, FDG is transported into the cell cytoplasm where it is phosphorylated and becomes trapped inside the cell as dephosphorylation hardly occurs.

Another well-characterised radiopharmaceutical is \(^{18}\)F-fluorodeoxythymidine (FLT)\(^{38-40}\). FLT is a pyrimidine analog that utilises the salvage pathway of DNA synthesis. Much like FDG, it is taken up through

**Figure 1** Patient with liver metastasis from gastric GIST. Top, PET; middle, contrast-enhanced diagnostic CT; bottom, fused images. After 18 days of imatinib there was a complete metabolic response on PET, while CT showed stable disease according to RECIST (−15%) and partial response according to Choi et al.\(^{32}\).
facilitated transport and diffusion and phosphorylated by a cell cycle-regulated enzyme, thymidine kinase 1 (TK1) and then becomes trapped in the cell. TK1 activity is higher in malignant cells than in normal cells; therefore, the uptake of FLT is a reflection of proliferative activity. Recent data indicate that the sensitivity to detect most tumour types is lower than that of FDG PET. However, specificity could be higher since FLT does not accumulate in inflammatory cells, although the latter is not undisputed.

PET scans can be evaluated both qualitatively (visual assessment) or (semi)quantitatively. Qualitative assessment is more practical for clinical use, but obviously less objective. Semiquantitative measurements of maximum standardised uptake value (SUV), average SUV and tumour-to-background ratio (TBR) have all been used. In recent years a gradual replacement of PET scans with hybrid PET-CT scanners has occurred.

PET in the various stages of imaging in sarcoma

Grading and staging

In theory, increased FDG uptake represents a metabolically active site. Thus, PET scanning could aid biopsy guidance and should discriminate between sarcoma and benign conditions. However, non-malignant processes like inflammation and areas with variable physiologic FDG turnover such as brown fat and muscle may interfere with image analysis. The inherently limited spatial resolution of PET compared to anatomical imaging has been largely solved by using hybrid PET-CT technology. Still, sensitivity for pulmonary and intrahepatic lesions may be relatively limited.

To date, only one meta-analysis has been performed, indicating that FDG-PET can indeed discriminate between sarcomas and benign tumours and low and high grade sarcomas, although the methodological quality of the studies included was generally poor. Thus, there is an urgent need for further evidence to support the routine clinical use of FDG-PET for diagnosing and staging sarcomas.

Treatment evaluation

The same principles that make FDG PET an interesting new option in the diagnostic phase apply to evaluation during treatment. Particularly, PET allows quantification of tumour viability or proliferation. When listing and evaluating studies that have investigated the value of PET or PET/CT for therapy evaluation in sarcoma, it becomes clear that there is no generally accepted definition for a metabolic response in sarcoma on FDG-PET. Preliminary criteria have been published by the EORTC/RECIST group and the National Cancer Institute, but the imaging protocols, measures of activity and definitions of response have varied and study size has been modest.

Bone sarcomas

Schulte et al., Franzius et al. and Nair et al. used TBR to determine metabolic response to neoadjuvant chemotherapy in a total of 64 patients. In the first...
two studies, the decrease of FDG uptake after chemotherapy correlated well with histological response. Nair et al. found that tumour necrosis was accurately predicted on PET scan in 15/16 patients by visual assessment, 14/15 patients by TBR value on presurgery scans, and 7/15 patients using percent change of TBR on serial scans. Hawkins et al. measured changes in the tumour SUV$_{\text{max}}$ in 69 patients and found a correlation with both post-chemotherapy SUV$_{\text{max}}$ and post- to pre-treatment SUV$_{\text{max}}$ ratio with histological response$^{[58,59]}$. Iagaru implemented the RECIST criteria for PET analysis in a heterogeneous group of bone and soft tissue sarcoma patients in a protocol including both PET and PET/CT imaging. The pathological degree of necrosis after chemotherapy was concordant with PET in 57.1% of cases; the observed discrepancies were attributed to chemotherapy-induced inflammation. Inflammation can be induced by radiotherapy and certain cytotoxic agents$^{[60-62]}$ sometimes presenting as the so-called fibrous pseudocapsule with inflammatory tissue that can form around the tumour$^{[63]}$. Uptake in the latter has been attributed to passive accumulation of FDG as a result of altered cell membrane permeability in the initial phase of irreversible cell death$^{[64]}$.

For the detection of recurrent disease, Arush et al.$^{[65]}$ and Gerth et al.$^{[66]}$ recently studied a total of 72 patients including many Ewing’s sarcoma cases. The first study confirms the high accuracy of FDG-PET/CT in the diagnosis of local relapse of sarcoma, while it failed in the detection of metastases in three patients. Gerth et al. compared PET and PET/CT and found that the sensitivity, specificity, and accuracy of single-modality PET were 71%, 95%, and 88%, respectively; the corresponding values for the hybrid PET/CT technique were 87%, 97%, and 94% ($P<0.0001$). PET/CT thus was significantly more accurate than PET alone for the detection and localisation of lesions.

**Soft-tissue sarcomas**

In this group of tumours, PET has been used to evaluate not only chemotherapy effects, but also the effect of isolated limb perfusion (ILP), a technique with the possibility of local high-dosed therapy to facilitate limb-sparing surgery. For ILP, Nieweg and colleagues first reported a case of liposarcoma in which PET suggested complete response which was later confirmed by histological examination$^{[67]}$. In larger groups of STS patients, it was shown by Van Ginkel et al. that based on the pre-treatment glucose consumption in soft-tissue sarcomas, one could predict the probability of a patient achieving complete response after ILP, although uptake in inflammatory tissue hampered the evaluation$^{[68]}$. To overcome this, PET with $L$-[11C]tyrosine was used by the same group$^{[69]}$. They were able to predict histological response by post-treatment uptake rates and inflammatory tissue did not interfere with viable tumour. More recently, a study from this group investigated the possibilities of FLT by PET/CT. Interestingly, uptake was correlated with the mitotic index of the tumours ($r=0.82$ and $P=0.004$ for SUV$_{\text{max}}$; $r=0.87$ and $P=0.001$ for SUV$_{\text{mean}}$). After HILP, the uptake of FLT decreased significantly ($P=0.008$ for SUV$_{\text{max}}$ and $P=0.002$ for SUV$_{\text{mean}}$). Tumours with initially high FLT uptake showed a better response to HILP ($r=0.64$, $P<0.05$)$^{[41]}$.

In the evaluation of chemotherapy, Jones et al. described a homogeneously decreased FDG uptake throughout the tumour in responsive cases. Again, despite complete necrosis, persistent tumour FDG uptake was observed in fibrous pseudocapsules$^{[63]}$. Shields describes two patients who underwent chemotherapy; in the responding patient a decrease of FDG uptake of 40% was seen while in the non-responding case uptake increased by 69%. Change in $[^{11}C]$thymidine incorporation was more marked in the responder but remained stable in the non-responding case$^{[70]}$. Changes in tumour SUV$_{\text{max}}$ predicted outcomes in 46 patients with localised extremity STS by PET scanning in a study by Schuetze et al.$^{[36]}$. Not only was a change in SUV$_{\text{max}}$ >40% correlated with the amount of residual viable tumour, also multivariate analysis found a correlation between lack of response and increased risk of disease recurrence, metastasis and death. Peng et al. confirmed the association between permanent uptake and rapid relapse versus decreased uptake and favourable response in rhabdomyosarcoma$^{[71]}$, as did Kasper et al., by demonstrating a significant difference in the progression-free survival for patients with a decrease in the standardised uptake value in comparison with patients with an increased or stable SUV$^{[72]}$.

Hybrid PET/CT scanning has been performed in the reports by Evilevitch et al. and Park et al.$^{[35,73]}$. In the Evilevitch study, FDG-PET was significantly more accurate than size-based criteria (RECIST) at assessing response to neoadjuvant therapy, correctly identifying all of the responders and 71% of the non-responders while only 25% of responding tumours were identified by size-based criteria. Moreover, threshold values ranging from 50% to 70% of baseline FDG uptake allowed assessment of response, thereby limiting the effect of remaining uptake in inflammatory lesions. Park et al. report that PET or PET/CT was highly effective in discriminating true recurrence in patients with suspected recurrence and was highly sensitive in detecting recurrence in asymptomatic patients.

**Gastro-intestinal stromal tumours**

The separate consideration of GIST from STS appears rather artificial, but there are some distinct features to FDG-PET in GIST. Comparability is better because the same tumour under the same treatment regimen is studied and adherence to standardised response criteria$^{[53]}$ has been rather strict. Also, directly from the first availability of imatinib for GIST treatment, it was shown that
FDG-PET seems to predict response very early ([Fig. 1](#)) and therefore it has been incorporated in a relatively large amount of study protocols, the larger of which we have listed in Table 1. Overall results of these studies have been that the main limitations are the occasional lack of pre-treatment FDG avidity ([31,75]) and the lower sensitivity for pulmonary and hepatic lesions ([76,77]). Still, all agree that PET scanning is a sensitive and rapid indicator of response preceding size-based response by weeks or months ([23,74,76]). Furthermore, response on PET scans is closely related to clinical symptom relief ([74]) and predicts clinical outcome ([32,75,79]).

FDG-PET in GIST exceeds the role of a staging and restaging modality. Different mutations with different therapeutic impact can exist synchronously in a patient ([81]). Furthermore, a change in microenvironment can result in replacement of tumour cells with drug-resistant variant cells ([27,82,83]). In this context, FDG-PET evaluation can be used to select only progressive lesions for resection in patients with ongoing response of the remainder of the metastases ([84]), an approach that is still controversial versus the generally accepted surgical approaches in oncology.

### Conclusion

Sarcomas have unique properties which not only increase demands on imaging but also pose specific problems. In the phase of diagnosis and staging, anatomical imaging techniques such as MRI for local tumour characterisation and CT for detection of pulmonary metastasis remain indispensable and reliable techniques. During treatment, the limitations of these techniques and their size-based evaluation in sarcoma become clearer. Although studies on the value of PET are of limited size and quality, PET is a promising modality especially for treatment evaluation in sarcoma, providing a rapid and reliable indication of response. The possibility to non-invasively detect tumour progression has already influenced clinical practice in GIST in a revolutionary way, with clear impact on patient management. PET scanning has inherent limitations which fortunately do not entirely overlap with those of anatomical imaging. Therefore, these techniques should be regarded as complementary. The studies comparing PET versus PET/CT underscore this statement. For future studies, the availability of objective criteria for response evaluation with PET would be highly instrumental to implement PET in a cost effective way for patient tailored sarcoma treatment.

### References


