The value of FDG-PET in the detection, grading and response to therapy of soft tissue and bone sarcomas: a systematic review and meta-analysis

CRD summary
This review found that fluorodeoxyglucose positron emission tomography can grade tumours and that its accuracy for the detection of sarcomas is good. However, there were insufficient data to assess response to therapy. The review suffered from a number of limitations, including failure to describe the methods of analysis or to investigate variability between the studies; the results should therefore be interpreted with caution.

Authors' objectives
To evaluate the diagnostic value of fluorodeoxyglucose (FDG) positron emission tomography (PET) in the detection, grading and therapeutic response of soft tissue and bone sarcomas.

Searching
PubMed, MEDLINE, EMBASE and the Cochrane Library were searched from 1990 to June 2002. The search terms were reported in the paper and included diagnostic filters. The reference lists of identified studies and relevant reviews were screened for additional studies.

Study selection
Study designs of evaluations included in the review
Clinical studies were eligible for inclusion. Case reports, reviews, editorials and studies of animals were excluded.

Specific interventions included in the review
Studies of FDG PET were eligible for inclusion. Studies that used other radiopharmaceuticals were excluded. The included studies assessed detection, grading and response to therapy. Some studies also compared PET with other tests such as magnetic resonance imaging, computed tomography, scintigraphy, ultrasound and single-photon emission tomography.

Reference standard test against which the new test was compared
No inclusion criteria relating to the reference standard were specified. Most of the studies used histopathological examination alone as the reference standard. In some studies this was combined with clinical follow-up.

Participants included in the review
Studies of patients assessed for soft tissue or bone sarcomas were eligible for inclusion. The age of the patients ranged from 1 to 89 years; overall, the studies included 57% male patients. The location of soft tissue sarcomas was mainly in the extremities; bone and mixed sarcomas were also located in the trunk.

Outcomes assessed in the review
No inclusion criteria relating to the outcomes were specified; only studies that reported sufficient data to calculate the sensitivity and specificity and standard uptake value (SUV) or doses uptake ratio (DUR) were included in the meta-analysis.

How were decisions on the relevance of primary studies made?
Four reviewers independently screened titles and abstracts for relevance. Final decisions on inclusion were based on the full article.

Assessment of study quality
Three reviewers independently assessed study quality. The criteria of the Cochrane Methods Working Group on
Systematic Reviews of Screening and Diagnostic Tests, as modified by Mijnhout and colleagues (see Other Publications of Related Interest), were used. The items assessed included: independent, blind comparison with the reference test; measurement of the reference test in all patients; application of the reference test in a standardised manner; demographic characteristics; disease stages; and clear inclusion and exclusion criteria. Any disagreements were resolved through discussion. Each item that a study fulfilled was scored as 1, giving a maximum possible score of 25. Each item was given a weighting based on the weights ascribed by four independent reviewers.

Data extraction
One reviewer extracted the data using a standardised form. Data were extracted on the number of patients, location of sarcoma, age, sensitivity, specificity, accuracy, SUV or DUR. Standard deviations were extracted or derived if not reported. The method of quantitative interpretation of PET images was divided into six categories.

Methods of synthesis
How were the studies combined?
The studies were classified as being of soft tissue sarcoma, bone sarcoma, or mixed sarcoma (studies that reported combined data on soft tissue and bone sarcomas that could not be separated). The authors stated that a meta-analysis was conducted using the program Comprehensive Meta-analysis, but provided no further details of the methods used.

How were differences between studies investigated?
The authors did not report a method for investigating differences between the studies.

Results of the review
Twenty-nine studies (n=1,278) were included: 11 on soft tissue sarcoma, 6 on bone sarcomas and 12 on mixed sarcomas. The sample sizes ranged from 5 to 202.

The methodological quality of most of the included studies was poor; the scores ranged from 8 to 20 out of a possible 25. A very detailed overview of the quality of the included studies was provided in the article.

Only 17 of the 29 studies reported sufficient data to calculate sensitivity and specificity. These studies assessed the detection and grading of sarcomas.

Detection of sarcomas.
All studies (17 studies): the sensitivity ranged from 50 to 100% and the specificity from 43 to 100%. The pooled sensitivity was 91% (95% confidence interval, CI: 89, 93), the pooled specificity was 85% (95% CI: 82, 88) and the pooled accuracy was 0.88 (95% CI: 0.86, 0.90).

Soft tissue sarcomas (9 studies): the sensitivity ranged from 74 to 100% and the specificity from 43 to 100%. The pooled sensitivity was 88% (95% CI: 83, 93) and the pooled specificity was 86% (95% CI: 81, 91).

Mixed sarcomas (8 studies): the sensitivity ranged from 86 to 100% and the specificity from 67 to 92%. The pooled sensitivity was 94% (95% CI: 91, 96) and the pooled specificity was 74% (95% CI: 68, 80).

Bone sarcomas (3 studies): the sensitivity ranged from 50 to 100% and the specificity from 96 to 98%. Pooled values were not reported.

Grading of sarcomas.
All studies (10 studies): there was a statistically significant difference in mean SUV between malignant and benign tumours (p<0.0005, p=0.006) and low- and high-grade tumours (p<0.0005).

Soft tissue sarcomas (6 studies): there was a statistically significant difference in mean SUV between malignant and benign tumours (p=0.006), but not between low- and high-grade tumours (p=0.104).
Bone sarcomas (1 study): there were differences in mean SUV between malignant and benign tumours, but the statistical significance of this difference was not reported.

**Authors' conclusions**
The review was limited by the poor quality of the included studies and the fact that only a few studies reported comparable outcomes. The results suggested that PET can discriminate between sarcomas and benign tumours and low- and high-grade tumours based on mean SUV, but there were insufficient data to assess response to therapy.

**CRD commentary**
This review assessed a broad objective and stated some inclusion criteria that could have been defined more clearly. The literature search included diagnostic filters, so relevant studies might have been missed. In addition, attempts to identify unpublished studies were limited, thus the review may be subject to publication bias. Details of the review process were reported and these did include steps to minimise bias, although the data extraction was only carried out by one reviewer and may therefore be subject to bias. A very detailed quality assessment was carried out and the results of this formed the basis for the 'Results' section; it was also discussed at length in the 'Discussion'. However, despite this in-depth analysis, study quality was not incorporated into the analysis.

The analysis conducted was limited and details of the methods used were not reported. It is therefore not possible to determine whether this was carried out appropriately. In addition, there was no formal assessment or investigation of heterogeneity. Given the very wide range in estimates of sensitivity and specificity, this makes it very difficult to interpret the results. The authors' conclusions should therefore be interpreted with caution.

**Implications of the review for practice and research**
Practice: The authors did not state any implications for practice.

Research: The authors stated that further clinical studies of good methodological quality are needed to confirm the review results; prospective outcome-oriented study designs are needed to develop guidelines for the cost-effective use of PET; and the value of FDG PET in the evaluation of therapy needs further exploration.

**Bibliographic details**

**PubMedID**
14766127

**DOI**
10.1016/j.ctrv.2003.07.004

**Other publications of related interest**

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Adolescent; Adult; Aged; Biopsy, Needle; Bone Neoplasms /mortality /pathology /radionuclide imaging /therapy; Combined Modality Therapy; Disease Progression; Female; Fluorodeoxyglucose F18; Humans; Immunohistochemistry;
Male; Middle Aged; Neoplasm Staging; Randomized Controlled Trials as Topic; Risk Assessment; Sarcoma /mortality /pathology /radionuclide imaging /therapy; Sensitivity and Specificity; Soft Tissue Neoplasms /mortality /pathology /radionuclide imaging /therapy; Survival Analysis; Tomography, Emission-Computed

**AccessionNumber**
12004000501

**Date bibliographic record published**
31/12/2007

**Date abstract record published**
31/12/2007

**Record Status**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.