18F-FDG PET, combined FDG-PET/CT and MRI for evaluation of bone marrow infiltration in staging of lymphoma: a systematic review and meta-analysis

Wu LM, Chen FY, Jiang XX, Gu HY, Yin Y, Xu JR

CRD summary
This review concluded that 18F-fluorodeoxyglucose-positron emission tomography/computed tomography (FDG-PET/CT) had better diagnostic performance than FDG-PET or magnetic resonance imaging, and was highly sensitive and specific for staging bone marrow involvement in patients with lymphoma. The review had a number of methodological weaknesses and used indirect comparisons of tests, so the authors' conclusions should be interpreted cautiously.

Authors' objectives
To assess the diagnostic performance of 18F-fluorodeoxyglucose-positron emission tomography (FDG-PET), FDG-PET/computed tomography (CT), and magnetic resonance imaging (MRI) in evaluating bone marrow infiltration in the staging of patients with primary lymphoma or recurrent lymphoma after complete remission.

Searching
MEDLINE and EMBASE were searched from January 1995 to July 2010; The Cochrane Library and China biomedicine databases (for Chinese articles) were also searched. Search terms were reported. Searches included methodological terms for test accuracy studies. ScienceDirect, SpringerLink, Scopus, and China Science and Technological Journal database were scanned for further articles. Bibliographies of retrieved articles were screened for additional studies. Only studies in English or Chinese were included.

Study selection
Studies that assessed the use of FDG-PET, FDG-PET/CT, or MRI (alone or in combination, but not in sequence) to identify and characterise bone marrow infiltration in the staging of patients with primary lymphoma or recurrent lymphoma after complete remission were eligible for inclusion. Studies were required to use histopathology and/or clinical and imaging follow-up for a minimum of six months as the reference standard to confirm diagnosis. Studies had to include at least ten participants and report sufficient data to derive numbers of true positive, false negative, false positive, and true negative imaging test results (per patient or per lesion). Studies that assessed different imaging modalities in combination (where the performance of an individual modality could not be differentiated), and studies on the diagnosis of bone marrow involvement in lymphoma with other co-existing diseases (where separate data on bone marrow involvement could not be extracted) were excluded.

In included studies, the mean age of participants ranged from 15.5 to 62.4 years; participants were predominately men (where reported). Most studies included both patients with Hodgkin's lymphoma and patients with non-Hodgkin's lymphoma.

Two reviewers independently assessed studies for inclusion; any disagreements were resolved by consensus.

Assessment of study quality
The methodological quality of included studies was assessed using the 14-item QUADAS (Quality Assessment of Diagnostic Accuracy Studies) tool. Only studies which met at least 9 of the 14 QUADAS criteria were included.

The authors did not state how many reviewers performed the quality assessment.

Data extraction
For each study and imaging technique, data were extracted to calculate sensitivity and specificity, with 95% confidence intervals (CIs) for the detection of bone marrow infiltration.

The authors did not state how many reviewers performed the data extraction.
Methods of synthesis

Pooled estimates of sensitivity, specificity, and diagnostic odds ratios with 95% confidence intervals were calculated for each imaging modality (model not specified). Summary receiver operating characteristic curves were estimated using the Moses and Littenberg model. A z test was used to assess whether measures of test performance differed between imaging modalities.

Between study heterogeneity was assessed using $\chi^2$ and $I^2$.

Publication bias was assessed using funnel plots.

Results of the review

Thirty-two studies (1,826 patients) were included in the review. Twenty studies evaluated FDG-PET, five studies evaluated FDG-PET/CT, and eight studies evaluated MRI (one study evaluated both FDG-PET/CT and MRI). Half of the included studies did not recruit consecutive participants. Most studies were not blinded or did not report blinding status.

FDG-PET ($^{18}$F-fluorodeoxyglucose-positron emission tomography; 19 studies): The pooled sensitivity estimate was 81.5% (95% CI 77.3 to 85.3) and specificity was 87.3% (95% CI 84.9 to 89.5).

FDG-PET/CT (five studies): The pooled sensitivity estimate was 91.6% (95% CI 85.1 to 95.9) for FDG-PET/CT and specificity was 90.3% (95% CI 85.9 to 93.7).

MRI (eight studies): The pooled sensitivity estimate was 90.3% (95% CI 82.4 to 95.5) and specificity estimates was 75.9% (95% CI 69.8 to 81.2.

The sensitivity and specificity estimates for FDG-PET/CT were significantly higher than those for FDG-PET and for MRI. There was statistically significant between study heterogeneity in all measures, but no evidence of a threshold effect.

Pooled estimates of diagnostic odds ratios and summary receiver operating characteristic plots were also presented.

There was no evidence of publication bias.

Authors’ conclusions

FDG-PET/CT had high sensitivity and specificity for diagnosing bone marrow infiltration in patients with lymphoma. Compared with MRI and FDG-PET alone, FDG-PET/CT could play an important role in the staging of lymphoma.

CRD commentary

The review reported a clear research objective. Inclusion criteria were specified for study participants, index test(s), and reference standard. A number of sources were searched for relevant studies, but (as noted by the authors) the restriction to English and Chinese language raised the possibility of language bias and omission of relevant studies. The study selection process included measures to minimise error and bias, but it was unclear whether similar measures were applied to data extraction and quality assessment. The methodological quality of included studies was assessed; results were used to select studies for inclusion and were reported for included studies.

Studies were included if they reported per patient or per lesion data, but it was unclear from the information provided whether reported estimates of test performance were per patient or per lesion. Statistically significant heterogeneity was present for all measures of test performance in all imaging modalities, so the pooled estimates reported were of questionable validity. Comparisons of diagnostic performance between imaging modalities were based on indirect comparisons; only one of the included studies evaluated more than one imaging modality.

Given a number of methodological weaknesses and the use of indirect comparisons of tests, the authors’ conclusions should be interpreted cautiously.

Implications of the review for practice and research

Practice: The authors did not specify any recommendations for clinical practice.
The authors stated that analysis of cost-effectiveness should be conducted for the diagnosis of bone marrow involvement in lymphoma.

**Funding**
Shanghai Leading Academic Discipline Project, China; Shanghai Jiaotong University School of Medicine Leading Academic Discipline Project, China.

**Bibliographic details**

**PubMedID**
21145680

**DOI**
10.1016/j.ejrad.2010.11.020

**Original Paper URL**
http://www.ejradiology.com/article/S0720-048X(10)00574-7/abstract

** indexing assigned by NLM**
Bone Marrow Neoplasms /epidemiology /pathology; Comorbidity; Female; Fluorodeoxyglucose F18; Humans; Lymphoma /epidemiology /pathology; Magnetic Resonance Imaging /statistics & numerical data; Male; Neoplasm Staging; Positron-Emission Tomography /statistics & numerical data; Prevalence; Radiopharmaceuticals; Reproducibility of Results; Sensitivity and Specificity; Subtraction Technique /statistics & numerical data; Tomography, X-Ray Computed /statistics & numerical data

**AccessionNumber**
12012004751

**Date bibliographic record published**
24/03/2012

**Date abstract record published**
18/10/2012

**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.