The impact of fluor-18-deoxyglucose-positron emission tomography in the management of colorectal liver metastases: a systematic review and metaanalysis

Wiering B, Krabbe P F, Jager G J, Oyen W J, Ruers T J

CRD summary
This review assessed the impact of fluorine-18-deoxyglucose-positron emission tomography (FDG-PET) in the management of colorectal liver metastases. The authors concluded that FDG-PET is useful in the diagnostic workup of these patients and is superior to computed tomography. Poorly described inclusion criteria and review methodology, and limited detail of the included studies, make it difficult to assess the reliability of the review's findings.

Authors' objectives
To assess the role of fluorine-18-deoxyglucose-positron emission tomography (FDG-PET) in the selection of patients for resection of colorectal liver metastases.

Searching
MEDLINE and EMBASE were searched up to January 2004. The search strategies were reported in full.

Study selection
Study designs of evaluations included in the review
Systematic reviews were excluded.

Specific interventions included in the review
Studies of FDG-PET were eligible for inclusion. The review also included studies of computed tomography (CT).

Reference standard test against which the new test was compared
No inclusion criteria relating to the reference standard of diagnosis were specified for the diagnostic accuracy studies. The methods used to determine the true disease status of the participants in each study were not reported.

Participants included in the review
Studies of patients with recurrent colorectal cancer were eligible for inclusion. The review appeared to focus on patients with potentially resectable colorectal liver metastases. The primary studies included patients with intrahepatic and extrahepatic metastases.

Outcomes assessed in the review
Studies were included in the review if they described the impact of FDG-PET on management, or reported FDG-PET results. The review assessed lesion-based sensitivity and specificity and the percentage of patients whose management was changed following FDG-PET.

How were decisions on the relevance of primary studies made?
The authors did not state how the papers were selected for the review, or how many reviewers performed the selection.

Assessment of study quality
The review used a methodological quality scoring tool, developed by a panel of experts (a hepatic surgeon, a nuclear medicine physician, a methodologist and a radiologist) specifically for use in this review. The scoring system assessed five 'domains':

patient population (patient and lesion characteristics, co-morbidity, sample size, management decision after normal workup and excluded patients);
study design (research question, inclusion criteria, study design, randomisation, study period and blinding);

statistics (descriptive statistics, sensitivity and specificity of PET and CT or magnetic resonance imaging for hepatic and extrahepatic metastases, and confidence interval, CI);

technology and imaging (pre-operative follow-up, surgical work-up, interval CT and PET, interval diagnostics and surgery, technical specifications and patient preparation for PET and CT, definition of positive and negative PET findings, number of reviewers, distribution of disease and lesions studied); and

confirmation (pre-operative ultrasound, histopathological confirmation or clinical follow-up, post-operative follow-up, losses to follow-up and change in management).

Items in the quality assessment tool were scored 0 when not satisfied, 0.5 when partially satisfied, and 1 when satisfied. Individual items were weighted and a final score awarded to each study. The items were adapted from the Standards for the Reporting of Diagnostic Accuracy Studies (STARD) statement. Studies scoring more than 10 points above the mean score were considered the best quality studies.

The authors did not state how quality was assessed, or how many reviewers performed the quality assessment.

**Data extraction**
The authors did not state how the data were extracted for the review, or how many reviewers performed the data extraction.

For each study, lesion-based sensitivity and specificity values were calculated directly from the reported number of true-positive, false-positive, true-negative and false-negative results. Inconclusive results were considered false positive.

**Methods of synthesis**
How were the studies combined?
Pooled estimates of sensitivity and specificity were calculated for FDG-PET and CT after correcting for the variable number of patients in each study.

A pooled estimate of the percentage of patients whose clinical management changed as a result of FDG-PET was also calculated, after correcting for the number of patients with liver metastases in the individual studies.

How were differences between studies investigated?
Pooled sensitivity and specificity estimates were calculated separately for hepatic and extrahepatic lesions and for the 6 studies with the highest quality scores. Lesion-based sensitivity and specificity values were illustrated graphically, by imaging modality, for each included study.

Pooled estimates of percentage change in management were stratified by quality score (separate analyses of studies with a quality score above the mean, studies with a score below the mean, and the 5 studies with the highest quality scores).

**Results of the review**
Thirty-two retrospective and prospective studies were included in the review. One study did not report the number of participants; the remaining 34 studies involved a total of 1,843 participants, of whom 987 were patients with liver metastases. The review did not include any randomised controlled trials. Twenty-nine studies reported data on diagnostic performance (sensitivity and specificity), while 17 studies reported data on changes in the clinical management of liver metastases.

The studies had several methodological limitations. Such limitations included poor or inadequate description of the methods used to deal with excluded patients, descriptive statistics, CIs, workup of patients for resection, technical aspects, preparation of patients for imaging, means and intervals of post-operative follow-up, impact of PET on clinical management, co-morbidity and the characteristics of primary tumours.
Diagnostic performance.

FDG-PET for hepatic lesions (n29 studies): the pooled estimate of lesion-based sensitivity was 88.0% (95% CI: 88, 98) and the corresponding estimate of specificity was 96.1% (95% CI: 70.4, 104.3).

FDG-PET for extra-hepatic lesions (22 studies): the pooled estimate of lesion-based sensitivity was 91.5% (95% CI: 84.3, 96.2) and the corresponding estimate of specificity was 95.4% (95% CI: 71.4, 98.4).

CT for hepatic lesions (24 studies): the pooled estimate of lesion-based sensitivity was 82.7% (95% CI: 64.2, 88.6) and the corresponding estimate of specificity was 84.1% (95% CI: 68.2, 97.0).

CT for extra-hepatic lesions (18 studies): the pooled estimate of lesion-based sensitivity was 60.9% (95% CI: 44.4, 68.9) and the corresponding estimate of specificity was 91.1% (95% CI: 66.0, 92.8).

The results were similar when only the 6 studies with the highest quality scores were pooled: FDG-PET had a sensitivity and specificity of 79.9% and 92.3%, respectively, for hepatic disease and 91.2% and 98.4% for extrahepatic disease; CT had a sensitivity and specificity of 85.8% and 88.3%, respectively, for hepatic disease and 55.3% and 95.6% for extrahepatic disease.

Change in clinical management.

After correcting for the number of patients with liver metastases in each study, the change in clinical management was 32.9% for studies with a quality score below the mean and 30.8% for those with a score above the mean. The pooled change of management for the 5 highest scoring studies was 25.4%.

Authors’ conclusions

FDG-PET is useful in the diagnostic workup of patients with potentially resectable colorectal liver metastases. It has superior sensitivity and specificity to CT for the detection of extrahepatic disease and a clear influence on clinical management in the majority of the included studies.

CRD commentary

The review addressed a broad clinical question. Few inclusion criteria were specified, though the included studies appeared appropriate for the question under consideration. Only two databases were searched and this might have resulted in the omission of other relevant studies. The search strategy was clearly reported, but did not include any attempt to identify unpublished studies. The review methodology was poorly reported, making it difficult to assess the potential impact of bias or error introduced during the review process. The development and use of a review-specific quality assessment tool was described, and the results of the assessment were presented graphically.

There was little detail on the individual included studies, in particular, a lack of data on the methods used to establish diagnosis (reference standard); this contributed to the difficulty in assessing the appropriateness of pooling and also impeded an assessment of the generalisability of the review’s findings. The results of the individual studies were depicted graphically; no clear numerical values were reported. Pooled estimates of sensitivity and specificity were calculated, but it was unclear whether between-study heterogeneity was assessed, or whether the included studies shared a common diagnostic threshold; simple pooling of sensitivity and specificity is generally considered inappropriate where there is variation in the diagnostic threshold. Pooled estimates of sensitivity and specificity calculated from the total numbers of true-positive, false-positive, true-negative and false-negative results in the individual studies assume that these data are derived from a common sample population, an assumption which cannot be tested without assessment of between-study heterogeneity. The impact of the methodological quality of the included studies on the outcome measures was evaluated using quality scores, which are often considered to be unreliable in comparison with individual quality items.

The authors’ conclusions appeared somewhat firm on the basis of the data presented and should be treated with caution given the limitations described. In particular, the authors’ conclusion that FDG-PET demonstrated superior sensitivity and specificity to CT may not be reliable since the CI for specificity estimates overlap.
Implications of the review for practice and research
Practice: The authors stated that multidisciplinary oncological meetings should be held to review clinical and imaging information on individual patients.

Research: The authors stated that randomised controlled trials are required to examine the place of FDG-PET scanning in patients with colorectal liver metastases and to assess outcomes such as survival and cost-effectiveness. They also stated that future research should address the methodological limitations identified in this review (see Results).

Bibliographic details

PubMedID
16315241

DOI
10.1002/cncr.21569

Indexing Status
Subject indexing assigned by NLM

MeSH
Colorectal Neoplasms /mortality /pathology /therapy; Female; Fluorodeoxyglucose F18; Hepatectomy /methods; Humans; Liver Neoplasms /mortality /radionuclide imaging /secondary /surgery; Male; Neoplasm Recurrence, Local /mortality /radionuclide imaging /surgery; Neoplasm Staging; Positron-Emission Tomography; Prognosis; Risk Assessment; Sensitivity and Specificity; Survival Analysis; Treatment Outcome

AccessionNumber
12006000130

Date bibliographic record published
30/11/2006

Date abstract record published
30/11/2006

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.