Performance of imaging modalities in diagnosis of liver metastases from colorectal cancer: a systematic review and meta-analysis

Florian I, Torri V, Rulli E, Garavaglia D, Compagnoni A, Salvolini L, Giovagnoni A

CRD summary
This review assessed and compared performance of imaging modalities (magnetic resonance imaging, MRI; computed tomography, CT; 18F-fluorodeoxyglucose positron emission tomography, FDG-PET; ultrasound) for the diagnosis of colorectal cancer liver metastases. The authors concluded that the evidence supported MRI. This conclusion should be viewed cautiously as comparative data, particularly for specificity and for MRI compared with FDG-PET, were limited.

Authors' objectives
To assess the performance of different imaging modalities in the diagnosis of colorectal cancer liver metastases.

Searching
MEDLINE and EMBASE were searched, without language restrictions, from January 2000 to August 2008. The full search strategy was reported. Bibliographies of trials, narrative reviews, and systematic reviews, identified by electronic searching, were screened to identify additional studies.

Study selection
Studies that compared at least two of specified imaging modalities (magnetic resonance imaging, MRI; computed tomography, CT; 18F-fluorodeoxyglucose positron emission tomography, FDG-PET; or ultrasound) in patients with a diagnosis or suspicion of liver metastases from colorectal cancer were eligible for inclusion. Included studies were required to use an adequate reference standard, defined as histopathological finding (surgical specimen, core biopsy, or positive cytological finding), intraoperative ultrasonography, or clinical and imaging follow-up. Studies of patients with solid tumours other than colorectal cancer were included if it was possible to extrapolate results obtained on colorectal cancer liver metastases.

The majority of included studies were of colorectal cancer patients only. The mean age of participants, where reported ranged from 56 to 67.9 years. Where details of the method were reported, studies of CT used multi-detector CT or helical CT. All MRI studies used contrast media (liver-specific or extra-cellular). Four of the six ultrasound studies used no contrast agent.

Studies were assessed for inclusion by two independent reviewers, a statistician and a radiologist. Disagreements were resolved by consensus.

Assessment of study quality
The methodological quality of included studies was assessed using the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) tool.

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

Data extraction
Data were extracted on the sensitivity and specificity, with 95% confidence intervals (CIs), of each imaging modality in each included study. Data were extracted per-lesion, and/or per-patient, as reported.

Data were were independently extracted by two radiologists and two statisticians, using a piloted data extraction form. Differences were resolved by consensus with a fifth reviewer, referring back to the original article.
Methods of synthesis
Pooled estimates of sensitivity, specificity, and positive and negative likelihood ratios were estimated using a fixed-effect or random-effects model, depending on the presence of statistical heterogeneity. Statistical heterogeneity was defined as an I² greater than 50%.

Odds ratios (ORs) were used to compare performance between tests, where the odds for sensitivity was defined as true positive rate/false negative rate and the odds for specificity was defined as true negative rate/false positive rate.

Heterogeneity was also assessed using the Cochrane Q statistic. Where statistical heterogeneity was identified, possible sources were explored using sub-group analyses, e.g. type of CT (CT versus CT arterial portography) and type of contrast media used for MRI (liver-specific contrast media versus extra-cellular contrast media).

Publication bias was assessed using a funnel plot.

Results of the review
Twenty-five studies, with a total of greater than 1,774 participants (number not reported for two studies), were included in the review; only six studies had a sample size of 100 or more participants. Ultrasound was assessed in six studies. CT was assessed in 24 studies and 26 data sets (12 helical, eight multi-detector CT, two CT arterial portography, two not unique, and two unknown). MRI was assessed in 11 studies (three gadolinium-chelate-enhanced, seven liver-specific contrast medium, one with both and one unknown). 18F-fluorodeoxyglucose positron emission tomography (FDG-PET) was assessed in 14 studies. Methodological quality assessment indicated verification bias could not be excluded in eight of 25 studies and that interpretation of results was not blinded in 15 of 25 studies.

Performance of ultrasound: The pooled estimate of sensitivity per-patient was 63.0% (95% CI 56.0 to 70.0; five studies) and specificity per-patient was 97.6% (95% CI 95.6 to 99.5; four studies). The overall positive likelihood ratio was 16.88 (95% CI 9.85 to 28.92) and the overall negative likelihood ratio was 0.34 (95% CI 0.20 to 0.58). There was no evidence of statistically significant heterogeneity. Only one study reported a per-lesion sensitivity for ultrasound of 86.3% (95% CI 76.3 to 93.2); no specificity data were reported.

Performance of CT: The pooled estimate of sensitivity per-patient was 74.8% (95% CI 71.2 to 78.3; 12 studies) and specificity per-patient was 95.6% (95% CI 93.4 to 97.8; seven studies). The overall positive likelihood ratio was 11.66 (95% CI 7.74 to 17.55) and the overall negative likelihood ratio was 0.38 (95% CI 0.25 to 0.58). There was no evidence of statistically significant heterogeneity. The pooled per-lesion sensitivity was 82.6% (95% CI 80.9 to 84.4; 19 studies) and the per-lesion specificity was 58.6% (95% CI 51.3 to 65.9; three studies); there was evidence of statistically significant heterogeneity in specificity values.

Performance of MRI: The pooled estimate of sensitivity per-patient was 81.1% (95% CI 76.0 to 86.1; five studies) and specificity per-patient was 97.2% (95% CI 94.5 to 99.9; two studies). The overall positive likelihood ratio was 29.16 (95% CI 15.04 to 56.56) and the overall negative likelihood ratio was 0.35 (95% CI 0.18 to 0.69). There was no evidence of statistically significant heterogeneity. The per-lesion estimate of sensitivity was 86.3% (95% CI 84.5 to 88.2; 11 studies) and the per-lesion estimate of specificity was 87.2% (95% CI 81.9 to 92.6; two studies).

Performance of FDG-PET: The pooled estimates of sensitivity per-patient was 93.8% (95% CI 90.0 to 97.7; six studies) and specificity per-patient was 98.7% (95% CI 97.2 to 100; six studies). The overall positive likelihood ratio was 51.53 (95% CI 31.99 to 82.99) and the overall negative likelihood ratio was 0.008 (95% CI 0.005 to 0.013). There was some evidence of statistically significant between study heterogeneity (p=0.0361). The per-lesion estimate of sensitivity was 86.0% (95% CI 83.2 to 88.8; nine studies) and the per-lesion estimate of specificity was 97.2% (95% CI 94.0 to 100; two studies).

Comparative performance: MRI showed a better sensitivity than CT in both per-patient analysis (OR 0.69, 95% CI 0.47 to 0.99; data sets from five studies) and in per-lesion analysis (OR 0.66, 95% CI 0.55 to 0.80; 11 studies). In per-lesion analysis, the difference was higher when liver-specific contrast agents were used for MRI. No other comparison showed a statistically significant difference in performance for any measure.

No association was found between sample size and diagnostic performance, and the authors concluded that there was
no evidence of publication bias.

**Authors' conclusions**
Available evidence supported MRI use for the detection of colorectal cancer liver metastases.

**CRD commentary**
The review question was defined by appropriate inclusion criteria and relevant studies were sought from a number of sources without restriction by language. Measures were taken to minimise error and/or bias in study selection and data extraction, but it was unclear whether similar measures were applied to quality assessment.

The methodological quality of included studies was assessed and summarised in the text, but results of quality assessment were not reported in full. Details of included studies and results were reported fully. The methods used for meta-analyses were reasonable, given the limited size and quality of the data set, and comparisons of test performance were derived from direct comparison data only.

The authors' conclusion (in favour of the use of MRI) should be viewed cautiously and requires some qualification, in that comparative data were limited; in particular, there was very little data that compared the performance of MRI with FDG-PET and data on specificity were limited.

**Implications of the review for practice and research**
The authors made no specific recommendations for practice or future research.

**Funding**
Not stated.

**Bibliographic details**

**PubMedID**
20027569

**DOI**
10.1002/jmri.22010

**Original Paper URL**

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Colorectal Neoplasms /diagnosis /epidemiology; Diagnostic Imaging /utilization; Humans; Liver Neoplasms /diagnosis /epidemiology /secondary; Reproducibility of Results; Sensitivity and Specificity

**AccessionNumber**
12010001281

**Date bibliographic record published**
23/06/2010

**Date abstract record published**
06/10/2010

**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.