18F-fluorodeoxyglucose positron emission tomography to evaluate recurrent gastric cancer: a systematic review and meta-analysis


CRD summary
This review concluded that 18F-fluorodeoxyglucose positron emission tomography (PET) had good performance for detection of recurrent gastric cancer but performed less well than contrast enhanced computed tomography (CT). Combined PET/CT may improve performance. The conclusions do not appear to be supported by the data and should be viewed with caution.

Authors' objectives
To assess the diagnostic accuracy of 18F fluorodeoxyglucose positron emission tomography (18F-FDG PET) for detection of recurrent gastric cancer.

Searching
MEDLINE and EMBASE were searched from January 2001 to July 2011 and CancerLit and The Cochrane Library were searched. Search terms were reported. Bibliographies of included studies and review articles were screened for additional studies. Only studies published in English were included.

Study selection
Studies that assessed 18F-FDG PET for detection of recurrent gastric cancer and included 10 or more participants were eligible for inclusion. Included studies were required to report patient level data on the numbers of true positive, false negative, false positive and true negative test results. Only studies that met more than nine out of the 14 quality criteria of the QUADAS tool were eligible for inclusion.

Most studies were retrospective. Study participants were aged between 27 and 87 years and around 65% were male. Most of the included studies used a combination of histology and clinical follow-up (not defined) as the reference standard to confirm diagnosis; two studies used clinical follow-up alone.

Two reviewers assessed studies for inclusion.

Assessment of study quality
The methodological quality of included studies was assessed using the QUADAS tool.

Two reviewers independently assessed study quality. Any disagreements were resolved by consultation with a third reviewer.

Data extraction
Data were extracted on per patient numbers of true and false positive and true and false negative test results. These data were used to calculate sensitivity and specificity estimates, with 95% confidence intervals (CIs), for each study. Where reported, data were also extracted for contrast-enhanced computed tomography (CT) and PET/CT.

A value of 0.5 was added to all cells of studies that contained a count of zero to enable calculation of sensitivity and specificity.

Authors of abstracts and studies that did not report sufficient data were contacted to request additional information.

Two reviewers independently extracted data. Any disagreements were resolved by consultation with a third reviewer.

Methods of synthesis
Pooled estimates of sensitivity and specificity, with 95% CIs, were calculated for 18F-FDG PET, CT and PET/CT. A random-effects model was used where significant heterogeneity was observed. Summary positive and negative likelihood ratios were derived from these estimates.
A summary receiver operating characteristic (SROC) curve was generated for $^{18}$F-FDG PET, which appeared to have used a hierarchical model.

Between-study heterogeneity was assessed using the $X^2$ test and the $I^2$ statistic. Potential sources of heterogeneity were investigated using meta-regression analysis. Subgroup analyses were performed for patient enrolment type (consecutive versus non-consecutive or not reported), blinding (blinded interpretation of test results versus no blinding or not reported) and reference standard (histological and clinical follow-up versus clinical follow-up only). Publication bias was assessed using Deeks’ funnel plots.

**Results of the review**

Nine studies (526 participants, range 18 to 139) were included in the review. Eight studies were retrospective. Only three studies reported enrolling patients consecutively. Three studies reported interpretation of $^{18}$F-FDG PET results blind to other test results and clinical data. Seven studies did not use the same reference standard regardless of index test result.

$^{18}$F-FDG PET: Pooled estimates of sensitivity and specificity were 76% (95% CI 70% to 81%; $I^2$=58%) and 82% (95% CI 77% to 87%; $I^2$=9%). Pooled positive and negative likelihood ratios were 3.52 (95% CI 2.68 to 4.63) and 0.32 (95% CI 0.22 to 0.46). Results of subgroup analyses were reported.

Contrast-enhanced CT: Pooled estimates of sensitivity and specificity were 72% (95% CI 62% to 80%) and 84% (95% CI 77% to 90%) based on five studies; heterogeneity assessment not reported. Pooled positive and negative likelihood ratios were 3.32 (95% CI 2.16 to 5.12) and 0.35 (95% CI 0.19 to 0.67).

CT and PET combined: Pooled estimates of sensitivity and specificity were 75% (95% CI 67% to 82%) and 85% (95% CI 79% to 90%). Pooled positive and negative likelihood ratios were 4.04 (95% CI 2.63 to 6.19) and 0.29 (95% CI 0.16 to 0.54).

SROC analysis found no evidence of threshold effect. The authors reported that differential verification was the most important variable in the meta-regression analysis; the direction of effect and results for other variables assessed were not reported.

There was no evidence of publication bias.

**Authors' conclusions**

$^{18}$F-FDG PET had good performance for detection of recurrent gastric cancer but still had some limited performance compared with contrast CT. The combination of $^{18}$F-FDG PET and CT might improve performance.

**CRD commentary**

This study defined a clear objective, specified appropriate inclusion criteria and searched a range of sources for relevant studies. The restriction to studies in English raised the possibility of language bias and meant that relevant studies may have been missed. The review process incorporated measures to minimise error and bias throughout. Methodological quality of included studies was assessed and reported. The reporting of the meta-analyses could have been clearer; the method used to generate the SROC curve was not clearly stated and it was unclear whether summary estimates of sensitivity were derived from this analysis or from a simple random-effects model (results differed between tables and text). Use of a hierarchical model throughout may have provided a more effective method of comparing performance between tests. The authors drew conclusions on the relative performance of $^{18}$F-FDG PET, CT and PET/CT but did not report any statistical comparison to support these conclusions; the test performance characteristics reported appeared similar across the three tests.

The authors' conclusions do not appear to be supported by the data and should be viewed with caution.

**Implications of the review for practice and research**

**Practice:** The authors did not specify any recommendations for clinical practice.

**Research:** The authors stated that larger studies might better clarify whether there was indeed an incremental diagnostic improvement with $^{18}$F-FDG PET combined with CT over contrast CT alone in recurrent gastric cancer.
Funding
Shanghai Leading Academic Discipline Project; Shanghai Jiao Tong University School of Medicine Leading Academic Discipline Project.

Bibliographic details

PubMedID
21916986

DOI
10.1111/j.1440-1746.2011.06919.x

Original Paper URL

Indexing Status
Subject indexing assigned by NLM

MeSH
Contrast Media; Fluorodeoxyglucose F18; Humans; Likelihood Functions; Neoplasm Recurrence, Local /radiography /radionuclide imaging; Positron-Emission Tomography; Radiopharmaceuticals; Sensitivity and Specificity; Stomach Neoplasms /radiography /radionuclide imaging; Tomography, X-Ray Computed

AccessionNumber
12012013247

Date bibliographic record published
10/07/2012

Date abstract record published
01/11/2012

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.