Positron emission tomography for nine cancers (bladder, brain, cervical, kidney, ovarian, pancreatic, prostate, small cell lung, testicular)

Ospina MB, Horton J, Seida J, Vandermeer B, Liang G

CRD summary
This generally well-conducted review concluded there is some evidence that $^{18}$fluorodeoxyglucose positron emission tomography (with or without CT) may be useful for diagnosing, staging or detecting recurrences for some types of cancer. The cautious conclusions and recommendations for research seem appropriate given the paucity of evidence available.

Authors' objectives
To evaluate the use of $^{18}$FDG-PET for diagnosis, staging, restaging and monitoring response to treatment of nine types of cancer.

Searching
MEDLINE, EMBASE, SCOPUS, Cochrane Central Register of Controlled Trials (CENTRAL) and NHS EED were searched without language restrictions (although only English-language studies were included) from 2003 to March 2008; the search strategy was available from downloadable appendices (See URL For Additional Data).

Study selection
Prospective or retrospective studies that compared $^{18}$fluorodeoxyglucose positron emission tomography ($^{18}$FDG-PET) with or without computed tomography (CT) to a reference standard (magnetic resonance imaging; CT; biopsy/histology; X-ray; ultrasound; PET with other radioisotope tracer; clinical follow-up) for diagnosis of cancer of the bladder, brain, cervix, kidney, ovary, pancreas, prostate, lung and testicles in at least 12 adults (>16 years), were eligible for inclusion. The primary outcomes were of diagnostic performance (this abstract focuses on these outcomes). Secondary outcomes included: need for additional diagnostic work-up; treatment decisions and management strategy; changes in therapy; clinical outcomes (such as survival, quality of life); and economic outcomes.

Diagnostic accuracy studies evaluated tests for initial diagnosis and/or recurrence and included patients with all stages of cancer. Participant age ranged from 17 to 93 years. The most common reference standard used was biopsy/histology (with or without clinical follow-up, conventional imaging or CT/bone scintigraphy). Doses of $^{18}$FDG varied across studies; some used fixed doses and others calculated by weight.

One of four reviewers evaluated the title and abstracts. Two independent reviewers evaluated studies retrieved as full papers. Disagreements were resolved by consensus.

Assessment of study quality
Study quality was assessed (dependant on the study design) using either: Scottish Intercollegiate Guidelines Network (SIGN) assessment tool for diagnostic studies, which is based on QUADAS; criteria related to study design, randomisation, allocation concealment, baseline comparability, characteristics of tests used, use of cointerventions, blinding and withdrawals/dropouts; or Consensus on Health Economic Criteria.

Two independent reviewers performed the quality assessment; disagreements were resolved by consensus.

Data extraction
Data were extracted to construct a 2x2 table of test performance from which sensitivity, specificity, positive and negative predictive values (PPV and NPV), and positive and negative likelihood ratios (LR+ and LR-) were calculated with 95% confidence intervals (CI).

Data were extracted by one reviewer and checked by a second; disagreements were resolved by consensus.
Methods of synthesis
Studies were combined in a narrative synthesis. Differences between studies were discussed in the text and results displayed visually using forest plots without pooled results. Studies were grouped depending on the diagnostic purpose of the test; primary diagnosis, staging, restaging or recurrence. Pooled estimates of the LR+ and LR-, with 95% CI, were calculated using the DerSimonian and Laird random-effects method. Heterogeneity was assessed using the $I^2$ statistic. Summary receiver operating characteristic (sROC) curves were produced; the model used was not reported.

Results of the review
One hundred and eight studies met the inclusion criteria. Of these, 104 reported on some aspect of diagnostic accuracy (n=5,935, range 15 to 517): 68 prospective studies and 36 retrospective. Unless specified, outcomes are for $^{18}$FDG PET. Of the 104 diagnostic accuracy studies, 43 recruited a representative patient spectrum, 100 avoided verification bias, 97 incorporation bias and 33 differential verification bias, and 86 reported some level of blinding (most commonly of the interpretation of the index test). Results for the impact on decision making and management were also reported.

Bladder cancer (two prospective and one retrospective studies):

For staging (two prospective, n=88), sensitivity was 53% (95% CI 27% to 79%) and 77% (95% CI 46% to 95%) and specificity 72% (95% CI 51% to 88%) and 94% (95% CI 81% to 99%). The pooled LR+ was 4.68 (95% CI 0.65 to 33.90) and LR- was 0.43 (95% CI 0.15 to 1.19); significant heterogeneity was observed.

Brain cancer (three prospective and two retrospective studies):

For staging (two prospective, n=42), sensitivity was 63% (95% CI 35% to 85%) and 63% (95% CI 38% to 84%). Specificity was reported in one study as 100% (95% CI 59% to 100%). Pooled likelihood ratios could not be calculated.

Cervical cancer (21 prospective and 12 retrospective studies):

For staging (five prospective, three retrospective, n=561), sensitivity ranged from 10% (95% CI 0% to 45%) to 100% (95% CI 77% to 100%) and specificity 76% (95% CI 53% to 92%) to 100% (95% CI 59% to 100%). From the prospective studies, pooled LR+ was 8.22 (95% CI 2.59 to 26.08) and LR- 0.38 (95% CI 0.12 to 1.20); significant heterogeneity was observed.

For detecting recurrence (three prospective and three retrospective, n=627), sensitivity ranged from 50% (95% CI 1% to 99%) to 100% (95% CI 16% to 100%) and specificity 76% (95% CI 70% to 82%) to 100% (95% CI 79% to 100%) depending on site of recurrence. From the prospective studies, pooled LR+ ranged from 15.24 (95% CI 5.63 to 41.27) to 45.89 (95% CI 14.09 to 149.49) and LR- from 0.09 (95% CI 0.02 to 0.40) to 0.37 (95% CI 0.22 to 0.60) depending on site of recurrence.

Kidney cancer (three prospective and five retrospective studies):

For staging (three retrospective, n=42), sensitivity ranged from 60% (95% CI 32% to 84%) to 90% (95% CI 55% to 100%) and specificity from 80% (95% CI 28% to 99%) to 100% (95% CI 3% to 100%). Pooled LR+ was 3.95 (95% CI 1.14 to 13.73) and LR- was 0.30 (95% CI 0.12 to 0.79); no significant heterogeneity was observed.

For primary diagnosis and staging (two retrospective, n=27), sensitivity was 60% (95% CI 32% to 84%) and 89% (95% CI 52% to 100%) and specificity was 100% (95% CI 3% to 100%). The pooled LR+ was 3.48 (95% CI 0.60 to 20.15) and LR- was 0.42 (95% CI 0.21 to 0.84); no significant heterogeneity was observed.

Ovarian cancer (14 prospective and six retrospective studies):

For detecting recurrences (six prospective and five retrospective), sensitivity for $^{18}$FDG PET/CT ranged from 73% (95% CI 39% to 94%) to 100% (95% CI 72% to 100%) and specificity 40% (95% CI 5% to 85%) to 100% (95% CI 29% to 100%). From four prospective studies using any reference standard, pooled LR+ was 6.97 (95% CI 1.94 to 25.0) and LR- 0.12 (95% CI 0.06 to 0.26).
Pancreatic cancer (14 prospective and three retrospective studies):

For primary diagnosis and staging (seven prospective, n=479), sensitivity ranged from 73% (95% CI 62% to 83%) to 97% (95% CI 91% to 99%) and specificity 41% (95% CI 18% to 67%) to 97% (95% CI 86% to 100%). Pooled LR+ was 2.77 (95% CI 1.62 to 4.73) and LR- 0.19 (95% CI 0.10 to 0.43).

Small cell lung cancer (six prospective and four retrospective studies):

For staging (three prospective and two retrospective, n=276), sensitivity was 100% (95% CI 48% to 100%). Specificity was reported in one study as 98% (95% CI 91% to 100%).

Testicular (three prospective and one retrospective studies) and prostate cancer (two prospective and two retrospective studies): No results were presented due to insufficient data. Results for other outcomes were reported, as were sROC curves.

Cost information
No economic evaluations were identified.

Authors' conclusions
For some types of cancer, there is some evidence of the utility of \(^{18}\)FDG PET (with or without CT) for diagnosing, staging or detecting recurrences.

CRD commentary
The authors addressed a clear, although very wide, research question, supported by appropriate inclusion criteria. Several relevant sources were searched, however, only English language studies were included, and there was no specific search for unpublished studies. Despite the large number of studies included in the review, the potential language and publication bias may still impact on the results as few studies reported the same outcomes for the same cancer site. Apart from the screening of titles and abstracts, the review process was conducted in duplicate, reducing the potential for error and bias where this method was employed. Study quality was assessed using appropriate criteria. Not all of the diagnostic accuracy studies seem to be included in the synthesis, although results were provide in an appendix. There seemed to be some discrepancies between the text and the on-line appendices. This was a generally well-conducted review, and the cautious conclusions and recommendations for research seem appropriate given the paucity of evidence available.

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

Research: The authors made a number of recommendations around the need for confirm the clinical and cost-effectiveness of \(^{18}\)FDG PET (with and without CT), for each potential use of its application, and for specific cancers where data is particularly lacking.

Funding
Agency for Healthcare Research and Quality.

Bibliographic details

Original Paper URL
http://www.cms.hhs.gov/determinationprocess/downloads/id54TA.pdf

Additional Data URL

**Indexing Status**
Subject indexing assigned by CRD

**MeSH**
Brain Neoplasms; Humans; Kidney Neoplasms; Lung Neoplasms; Ovarian Neoplasms; Pancreatic Neoplasms; Positron-Emission Tomography; Prostatic Neoplasms; Testicular Neoplasms; Urinary Bladder Neoplasms; Uterine Cervical Neoplasms

**AccessionNumber**
12009108218

**Date bibliographic record published**
11/11/2009

**Date abstract record published**
24/03/2010

**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.