The diagnostic accuracy of CT and MRI in the staging of pelvic lymph nodes in patients with prostate cancer: a meta-analysis

Hovels AM, Heesakkers RA, Adang EM, Jager GJ, Strum S, Hoogeveen YL, Severens JL, Barentsz JO

CRD summary
This review concluded that computed tomography (CT) and magnetic resonance imaging (MRI) demonstrated equally poor performance in detection of lymph node metastases from prostate cancer. The review suffered from a number of methodological weaknesses and results are likely to be out of date (the most recent included study was 2003). The results of this review should be interpreted with caution.

Authors' objectives
To compare the diagnostic accuracy of computed tomography (CT) and magnetic resonance imaging (MRI) in the diagnosis of lymph node metastases in prostate cancer.

Searching
MEDLINE and The Cochrane library were searched for articles published between 1980 and 2003. Search terms were reported. Bibliographies of included studies were checked manually to identify additional articles. Only English-language articles were included.

Study selection
Studies to determine the accuracy of CT and/or MRI for nodal staging in patients with prostate cancer were eligible for inclusion. Included studies were required to use histopathological evaluation of the lymph nodes as the reference standard to establish diagnosis. Included studies had to report (or provide sufficient data to calculate) numbers of true positive, true negative, false negative and false positive test results. Population characteristics were sparsely reported by the included studies. The threshold for a positive lymph node varied between 0.5cm and 1.5cm. All studies used pelvic lymph node dissection (PLND) as the reference standard; 10 out of 24 studies also used fine-needle aspiration biopsy (FNAB).

Studies were assessed for inclusion by two reviewers. Disagreements were resolved by consensus.

Assessment of study quality
Methodological quality of included studies was assessed based on the criteria: sample size; publication year; consecutively enrolled patients; prospective study design; reference tests; blind interpretation of test results; and a clear description of the test.

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

Data extraction
Data were extracted on numbers of true positive, true negative, false negative and false positive test results for each included study. These were used to calculate sensitivity, specificity and diagnostic odds ratio (DOR).

The authors did not state how many reviewers were involved in data extraction.

Methods of synthesis
Pooled estimates of sensitivity and specificity, with 95% confidence intervals (CIs), were calculated, for CT and MRI using a random-effects model. In cases of zero values, a correction factor of 0.5 was added to all cells in the 2x2 table for that study.

Summary receiver operating characteristic (SROC) curves were estimated (using the Moses and Littenberg method) and...
Q* (an overall measure of accuracy indicating maximum joint sensitivity and specificity) values were estimated for MRI and CT. Q* values and their standard errors were used to assess differences in performance between CT and MRI. In case of a symmetrical SROC curve, the Mantel-Haenszel method was used to generate pooled diagnostic odds ratios.

Pooled estimates of positive and negative likelihood ratios were calculated (method not reported) and used to calculate post-test probabilities.

Subgroup analyses were used to determine whether certain methodological or clinical characteristics affected diagnostic accuracy; differences between subgroups were assessed by comparing log diagnostic odds ratios using unpaired t-tests or Mann-Whitney U test, as appropriate.

Results of the review
Twenty four studies were included in the review. Four studies directly compared CT and MRI. Ten studies that used MRI (n=628 participants) and 18 studies that used CT (n=1,024 participants) were included in the meta-analyses. Most of the included studies (16 out of 24) were prospective, but only eight of these recruited consecutive series of patients. Only four studies reported blinded interpretation of test results. All except two studies provided a clear description of the test.

For CT, pooled sensitivity was 0.42 (95% CI 0.26 to 0.56) and pooled specificity was 0.82 (95% CI 0.80 to 0.83). For MRI, pooled sensitivity was 0.39 (95% CI 0.22 to 0.56) and the pooled specificity was 0.82 (95% CI 0.79 to 0.83).

For CT, estimated Q* value of was 0.77 (95% CI 0.69 to 0.83) and for MRI estimated Q* value was 0.77 (95% CI 0.73 to 0.80); there was no significant difference in performance between CT and MRI.

Average prevalence of lymph node metastases in studies included in this analysis was 0.17 for CT and 0.3 for MRI. When these numbers were used as pre-test probabilities and combined with pooled estimates of positive and negative likelihood ratios, post-test probabilities of a positive test were 0.31 (95% CI 0.23 to 0.40) for CT and 0.47 (95% CI 0.30 to 0.58) for MRI. Post-test probabilities for a negative test were 0.12 (95% CI 0.10 to 0.16) for CT and 0.23 (95% CI 0.18 to 0.29) for MRI.

Diagnostic accuracy of CT (as indicated by DOR) was higher in the five studies with greater than 50 participants. No other subgroups showed significant differences.

Authors' conclusions
CT and MRI demonstrated an equally poor performance in detection of lymph node metastases from prostate cancer. Reliance on either CT or MRI would misrepresent the patient's true status regarding nodal metastases and thus misdirect the therapeutic strategies offered to the patient.

CRD commentary
The review addressed a clearly stated research question defined by appropriate inclusion criteria. The literature search was restricted to two bibliographic databases and English-language studies, which left open the possibility of language bias and the likelihood that relevant studies were omitted. Measures were taken during study selection to minimise potential for error and/or bias; it was unclear whether similar measures were applied throughout the review process. Methodological quality of included studies was assessed and incorporated into interpretation of results. No formal heterogeneity assessment was presented, but values of sensitivity and specificity reported for individual studies varied widely; the simple pooling of these values to generate overall estimates was of limited utility. The summary ROC curve represented a more appropriate approach. Hierarchical/bivariate models (now recommended for generation of summary ROC curves) allow simultaneous generation of overall sensitivity and specificity estimates. The authors' conclusions broadly reflected the data presented, but it should be noted that there was considerable between-study heterogeneity and insufficient technical details were presented for the tests used in individual studies to allow adequate interpretation. In addition, the most recent studies included in this review were published in 2003; in a rapidly evolving field such as diagnostic imaging, the results of this review are likely to already be out of date.

Implications of the review for practice and research
Practice: Because of the low post-test probabilities of a positive test, assessment of lymph node involvement should not be done using either CT or MRI.

Research: The authors made no recommendations for future research.

Funding
Not stated.

Bibliographic details

PubMedID
18325358

DOI
10.1016/j.crad.2007.05.022

Original Paper URL
http://www.clinicalradiologyonline.net/article/S0009-9260(07)00334-0/abstract

Indexing Status
Subject indexing assigned by NLM

MeSH
Humans; Lymph Nodes; Lymphatic Metastasis; Magnetic Resonance Imaging /standards; Male; Neoplasm Staging; Pelvic Neoplasms /diagnosis /pathology /secondary; Prostatic Neoplasms /pathology; ROC Curve; Sensitivity and Specificity; Tomography, X-Ray Computed /standards

AccessionNumber
12008103970

Date bibliographic record published
23/12/2008

Date abstract record published
09/06/2010

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.