Systematic review of the diagnostic accuracy of 99mTc-HMPAO-SPECT in dementia
Dougall N J, Bruggink S, Ebmeier K P

CRD summary
The authors' conclusion, that brain single-photon enhanced computed tomography (SPECT) is less sensitive but more specific than clinical criteria in differentiating Alzheimer disease from other dementias, relies upon data from a post-mortem study published elsewhere. The recommendation for clinical follow-up studies to provide more reliable accuracy data and to establish the predictive role of SPECT with respect to disease progression and treatment response appears reasonable.

Authors' objectives
To assess the diagnostic performance of ^{99m}technetium-hexamethyl-propylenamine oxime (Tc-HMPAO) single-photon enhanced computed tomography (SPECT) imaging in distinguishing Alzheimer disease (AD) from other dementias. The secondary objectives were to compare discriminatory power between AD and other dementias with the diagnostic accuracy obtained in distinguishing AD from normal, and to determine which study characteristics contribute to the variation in observed diagnostic accuracy.

Searching
MEDLINE and EMBASE were searched for studies published between January 1985 and December 2002; no search terms were reported. Additional studies were identified from personal collections and from the bibliographies of narrative reviews.

Study selection
Peer-reviewed studies of more than 10 participants, with more than 5 participants in each subject group, were eligible for inclusion; conference proceedings and review articles were excluded. The included studies were of diagnostic cohort and diagnostic case-control designs. All included studies were required to evaluate the ^{99m}Tc-HMPAO-SPECT imaging examination in a population that included participants with AD. A variety of criteria were used to define ^{99m}Tc-HMPAO-SPECT diagnosis in the included studies, and these are described in the publication. The reference standard method used to determine the clinical diagnosis was not specified. The included studies were required to report sufficient data to populate 2x2 contingency tables (i.e. numbers of true positives, false negatives, false positives and true negatives).

Two reviewers independently assessed all identified studies against the specified inclusion criteria; any differences were resolved by consensus between the three authors.

Assessment of study quality
Two reviewers independently applied validity assessment criteria; any differences were resolved by consensus between the three authors. The criteria assessed: validity of the reference standard (clearly defined clinical diagnosis); blinding (assessment of SPECT blind to clinical diagnosis and vice versa); independence of patient selection from SPECT result; and treatment paradox (effect of treatment, after clinical diagnosis, on SPECT result). These internal validity criteria were used to assign an overall quality score on an ascending 3-point scale (-, +, ++). A similar scoring method was applied to external validity (extent to which the results of a study are applicable to the general population).

Data extraction
The authors did not state how the data were extracted for the review, or how many of the reviewers performed the data extraction. 2x2 Contingency data were extracted for each included study and comparison (some studies reported multiple data sets). Calculated values for sensitivity and specificity were also reported.

Methods of synthesis
Diagnostic odds ratios from each data set, along with 95% confidence intervals (CIs) were displayed on forest plots. The effect of diagnostic threshold on performance was investigated by assessing the correlation of sensitivity and

specificity. Where no evidence for threshold effect was observed and where there was no evidence for between-study heterogeneity (χ^2 test), a pooled estimate of the diagnostic odds ratio (DOR) was generated using the Mantel-Haenszel fixed-effect model. Where there was evidence of threshold effect, a summary receiver operating characteristic curve and multiple regression analysis were used to explore the effect of study level variables on DOR. Funnel plots were used to assess the extent of publication bias.

Results of the review
A total of 48 studies (41 diagnostic cohorts and 7 diagnostic case-controls) were included in the review.

Accuracy of SPECT in distinguishing AD from vascular dementia (13 studies, 768 participants).

The mean age of the study participants, where reported, ranged from 62.8 to 80.2 years. All studies scored + or ++ for both internal and external validity. There was no evidence of a threshold effect. The pooled estimates of sensitivity and specificity were 71.3% (95% CI: 67.5, 75.2) and 75.9% (95% CI: 70.8, 81.1), respectively. Regression analysis found no explanatory variables for the observed between-study heterogeneity in DOR. The funnel plot indicated an absence of publication bias.

Accuracy of SPECT in distinguishing AD from fronto-temporal dementia (7 studies, 424 participants).

The mean age of the study participants, where reported, ranged from 55.6 to 72.8 years. Six studies were graded ++ and one – for internal validity, whereas four were graded ++ and three + for external validity. There was no evidence of a threshold effect or significant between-study heterogeneity in DOR. The pooled estimates of sensitivity and specificity were 71.5% (95% CI: 66.3, 76.7) and 78.2% (95% CI: 71.2, 85.2), respectively. Regression analysis was not attempted. With only 7 studies, the funnel plot was inconclusive.

Accuracy of SPECT in distinguishing AD from non-dementia patients (13 studies, 1,082 participants).

The mean age of the study participants, where reported, ranged from 56.0 to 77.4 years. Control groups varied and included normal controls and/or patients with headaches and dizziness, memory problems, mild dementia, or psychiatric disorders. All studies scored + or ++ for internal validity and all but one for external validity. There was no evidence of a threshold effect. The pooled estimates of sensitivity and specificity were 65.7% (95% CI: 62.2, 69.3) and 79.1% (95% CI: 75.1, 83.1), respectively. Regression analysis found no explanatory variables for the observed between-study heterogeneity in DOR. The funnel plot indicated an absence of publication bias.

Authors' conclusions
Pathological verification studies (elsewhere) indicate that the application of clinical criteria may be a more sensitive method of detecting AD than brain SPECT; however, SPECT has a higher specificity. SPECT may therefore be helpful in the differential diagnosis of AD. Clinical follow-up studies are needed to establish the predictive role of SPECT with respect to disease progression and treatment response in AD.

CRD commentary
This was a generally well-conducted review that addressed a series of relevant and clearly stated research questions. The inclusion criteria were clearly stated and relevant details of individual included studies were well reported. The search strategy was reasonable, though the large number of additional studies identified from sources other than the electronic searches may indicate a poor rate of retrieval. The review process was described clearly and appropriate measures were employed to minimise error and bias in the study selection and validity assessment processes. Criteria relevant to diagnostic accuracy studies were used to assess the methodological quality of the included studies. However, the results of the quality assessment were only reported as summary scores, which limits their usefulness to the reader.

The methods used to assess between-study heterogeneity and to generate pooled estimates of measures of diagnostic performance were broadly appropriate. The generation of pooled estimates of sensitivity and specificity in the presence of significant, unexplained between-study heterogeneity is of questionable value. The use of standard funnel plots is not recommended as a method of assessing publication bias in systematic reviews of diagnostic accuracy studies. The authors' conclusions, that clinical examination is more sensitive but less specific than brain SPECT, rely upon data from a post-mortem study published elsewhere and are not derived from the data presented in the review; their reliability is

therefore uncertain.

Implications of the review for practice and research

Practice: The authors made no specific recommendations for practice.

Research: The authors stated that clinical follow-up studies are needed to validate initial clinical diagnoses and thus provide more reliable accuracy data, and to establish the predictive role of SPECT with respect to disease progression and treatment response in AD.

Funding

Bibliographic details

PubMedID

DOI

Indexing Status

Subject indexing assigned by NLM

MeSH

Aged; Aged, 80 and over; Alzheimer Disease /radionuclide imaging; Brain /radionuclide imaging; Dementia /radionuclide imaging; Dementia, Vascular /radionuclide imaging; Diagnosis, Differential; Female; Humans; Male; Middle Aged; Predictive Value of Tests; Technetium Tc 99m Exametazime; Tomography, Emission-Computed, Single-Photon

AccessionNumber

12004002046

Date bibliographic record published

03/05/2007

Date abstract record published

09/08/2008

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.