Accuracy of magnetic resonance imaging for the diagnosis of multiple sclerosis: systematic review

CRD summary
This well-conducted review assessed the accuracy of magnetic resonance imaging (MRI) criteria for the early diagnosis of multiple sclerosis (MS) in patients with suspected disease. The authors concluded that MRI appears to be a relatively poor test for both ruling in and ruling out MS. This conclusion is likely to be reliable.

Authors' objectives
To determine the accuracy of magnetic resonance imaging (MRI) criteria for the early diagnosis of multiple sclerosis (MS) in patients with suspected disease.

Searching
Twelve unspecified databases were searched from inception to September or November 2004 for published and unpublished studies. The search terms were reported and no language restrictions were applied. Further studies were sought through a citation search on an article reporting the McDonald 2001 criteria, and by scanning the reference lists of included studies and studies in the National Institute of Clinical Excellence guidelines for MS.

Study selection
Diagnostic accuracy studies of any design were eligible for inclusion.

Specific interventions included in the review
Studies of MRI, or diagnostic criteria incorporating such imaging, were eligible for inclusion.

Reference standard test against which the new test was compared
The studies were required to use a reference standard for the diagnosis of MS; no further criteria were specified. Most of the included studies used clinical follow-up as the reference standard, using either the Poser or McDonald 1977 criteria.

Participants included in the review
No inclusion criteria for the participants were specified. The included studies varied in relation to their participant inclusion criteria: some included only patients presenting with a particular clinically isolated syndrome, while others included all patients with suspected MS.

Outcomes assessed in the review
The studies were required to report sufficient data to enable the construction of a 2x2 table of test performance.

How were decisions on the relevance of primary studies made?
Two reviewers independently screened titles and abstracts for relevance. Screening for inclusion was carried out by one reviewer and checked by a second.

Assessment of study quality
Studies were assessed using 13 items from the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) criteria (see Other Publications of Related Interest). One reviewer carried out the quality assessment and a second reviewer checked it.
Data extraction
One reviewer carried out the data extraction and a second reviewer checked it. Data for 2x2 tables were extracted, and sensitivity, specificity and likelihood ratios were calculated for each table. For each study, the median diagnostic odds ratio (DOR) was identified.

Methods of synthesis
How were the studies combined?
The results from all studies were plotted on a receiver operating characteristic (ROC) plot of sensitivity against specificity. The studies were split into two groups: prospective cohort studies that enrolled patients with suspected MS and studies of other designs. Summary DORs with 95% confidence intervals (CIs) were calculated for each group (from the median DORs for each study) using a random-effects meta-analysis.

How were differences between studies investigated?
A hierarchical summary ROC method was used to assess the effect of duration of follow-up on accuracy and threshold among the cohort studies, using a reference standard diagnosis of clinically definite MS arrived at by clinical information alone. ROC plots were generated for groups of these studies evaluating commonly reported MRI criteria (Barkhof, Fazekas, Paty) or for combined clinical and MRI criteria (McDonald 2001). Individual study ROC plots were produced for studies with at least 10 years' clinical follow-up and the area under the curves was compared.

Results of the review
The review included 29 studies: 18 cohort studies and 11 studies of other designs. The total number of participants involved was unclear.

Study quality was generally poor; three items on which studies scored badly were blinding, the use of an appropriate reference standard, and the availability of clinical information.

Cohort studies produced lower estimated sensitivity and specificity than studies of other designs. The pooled DOR was 9 (95% CI: 5, 16) for the cohort studies and 213 (95% CI: 85, 535) for studies of other designs. These were significantly different (P<0.001, permutation test).

Among the 15 cohort studies that used a reference standard of clinically definite MS diagnosed using clinical data alone, it was found that studies with longer follow-up produced higher sensitivity and lower specificity.

The negative likelihood ratios for studies using the Barkhof, Fazekas and Paty criteria ranged from 0.2 to 0.5 and the positive likelihood ratios were all less than 5. This suggested that these criteria are of limited utility for either predicting or ruling out the development of MS (within the 3- to 6-year follow-up duration of these studies). Studies using the McDonald 2001 criteria gave similar negative likelihood ratios (range: 0.1 to 0.5). The positive likelihood ratios ranged from 2.7 to 8.7, suggesting that the combined MRI and clinical criteria have more potential for predicting the development of MS than the MRI criteria alone.

An analysis of the 2 studies with at least 10 years' follow-up suggested that MRI is of limited utility for predicting or ruling out a diagnosis of MS.

Authors' conclusions
MRI appears to be a relatively poor test for both ruling in and ruling out MS.

CRD commentary
The review question and inclusion criteria were clearly stated. Several relevant sources were searched for primary studies and there were no language or publication status restrictions, making it unlikely that relevant studies were missed. Some attempt was made to reduce the introduction of reviewer bias and error during the study selection and data extraction processes by having a second reviewer check the work. The quality of the included studies was formally assessed.
assessed, and the results were clearly reported and discussed. The studies were grouped and combined appropriately, taking account of the differences between the studies in terms of design and clinical factors. The conclusions follow from the results presented and they are likely to be reliable.

Implications of the review for practice and research
Practice: The authors stated that neurologists should discuss with their patients the effects of potential errors of false-positive and false-negative MRI results.

Research: The authors stated that high-quality research using improved MRI techniques with a complete description of participants and long term clinical follow-up is warranted for further quantitative assessment of MRI in the diagnosis of MS.

Funding
Medical Research Council Health Services Research Collaboration.

Bibliographic details

PubMedID
16565096

DOI
10.1136/bmj.38771.583796.7C

Original Paper URL
http://bmj.bmjjournals.com/cgi/content/full/332/7546/875

Other publications of related interest

These additional published commentaries may also be of interest. Role of MRI in diagnosing multiple sclerosis [letters]. BMJ 2006;332:1034-5. Weinschenker BG. Review: Magnetic resonance imaging alone is of limited usefulness in diagnosing multiple sclerosis. ACP J Club 2006;145:51.

Indexing Status
Subject indexing assigned by NLM

MeSH
Case-Control Studies; Cohort Studies; Early Diagnosis; Humans; Magnetic Resonance Imaging /standards; Multiple Sclerosis /diagnosis; Sensitivity and Specificity

AccessionNumber
12006008162

Date bibliographic record published
31/07/2006

Date abstract record published
31/07/2006
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.