Does this patient have hearing impairment?
Bagai A, Thavendiranathan P, Detsky A S

CRD summary
This review evaluated the accuracy of bedside clinical manoeuvres for diagnosing hearing impairment. The authors concluded that elderly individuals who acknowledge a hearing impairment require audiometry, while others require the whispered-voice test. Those failing the whispered-voice test require audiometry, while others need no further testing. Weber and Rinne tests should not be used. The conclusions are largely based on problematic evidence and caution is required.

Authors' objectives
To evaluate the accuracy and precision of bedside clinical manoeuvres for diagnosing hearing impairment. Only the diagnostic accuracy data will be abstracted for this abstract.

Searching
MEDLINE and EMBASE were searched from inception to April 2005. In addition, the references of identified articles, pertinent articles on the topic and a clinical skills textbook were screened, and experts in the field were contacted. Some search terms were given. The review was limited to articles reported in the English language.

Study selection
Study designs of evaluations included in the review
Original accuracy or precision articles were eligible for inclusion. All but one of the included diagnostic studies were prospective evaluations.

Specific interventions included in the review
Studies examining the accuracy and/or precision of bedside screening questions or physical examinations for hearing impairment were eligible for inclusion. The included studies investigated the diagnostic value of a self-reported screening question (e.g. 'Do you have trouble hearing?'), the Hearing Handicap Inventory for the Elderly - Screening Version (HHIE-S), audioscopes, whispered-voice test, Weber tuning fork test, Rinne tuning fork test, or a combination of these.

Reference standard test against which the new test was compared
Studies that used another screening test evaluated in the review as the reference standard were excluded. The included studies used an audiogram as reference standard or, in one case, an audiologic examination and audiometry; the definition of hearing impairment varied.

Participants included in the review
Studies of participants aged 16 years or older were eligible for inclusion; the patients could be symptomatic or asymptomatic. The patients in the included studies came from community populations, nursing homes or specific care settings. Where reported, the mean age in the samples ranged from 32 to 80.8 years.

Outcomes assessed in the review
To be eligible for inclusion, studies had to enable the calculation of likelihood ratios (LRs). The review presented positive and negative LRs as measures of diagnostic accuracy.

How were decisions on the relevance of primary studies made?
One reviewer screened the identified studies; a second reviewer checked the included studies independently.

Assessment of study quality
The studies were classified using a 5-level grading scheme that took sample size, use of consecutive patients and other features into account. Studies investigating the self-reported screening question and the HHIE-S had to be of level 1 or 2 (blind independent comparisons of a test with a valid reference standard in a sample consecutive patients) to be eligible for inclusion in the review. For the other tests, the best available evidence studies were included (several level 3 or 4 studies: blind independent comparisons of the test with a reference standard, but patients enrolled in non-consecutive fashion - using a subset or smaller group who may have had the condition or a 'grab sample' of patients believed to have the condition).

The authors did not state how the validity assessment was performed.

**Data extraction**

Two reviewers independently extracted the data, with any differences resolved by consensus. Positive and negative LRs were calculated from the raw data.

**Methods of synthesis**

How were the studies combined?

Studies with similar definitions of hearing impairment and using similar screening tests were combined using a random-effects model to compute summary accuracy data and confidence intervals (CIs).

How were differences between studies investigated?

The reviewers stated that statistical heterogeneity was computed.

**Results of the review**

Twenty-four studies (n=12,645) were included in the review.

The CI around the point estimate was considered sufficiently narrow to allow appropriate inferences from the pooled result despite the presence of heterogeneity.

**Self-reported screening question.**

With a reference standard of more than 25 dB (4 identified studies), the positive LR was 2.2 (95% CI: 1.8, 2.8) and the negative LR 0.45 (95% CI: 0.36, 0.56). A reference standard of over 30 dB (2 studies) resulted in a positive LR of 2.4 (95% CI: 1.6, 3.8) and a negative LR of 0.49 (95% CI: 0.41, 0.59). A reference standard of over 40 dB (2 studies) showed a positive LR of 2.5 (95% CI: 1.7, 3.6) and a negative LR of 0.13 (95% CI: 0.09, 0.19).

**HHIE-S.**

Defining hearing impairment as more than 8 points on the HHIE-S and a pure tone average threshold of over 25 dB resulted in a positive LR of 4.5 (95% CI: 3.1, 6.6; 6 studies) and a negative LR of 0.55 (95% CI: 0.45, 0.67; 5 studies). Defining hearing impairment as more than 8 points on the HHIE-S and a pure tone average threshold of over 40 dB resulted in a positive LR of 3.8 (95% CI: 3.0, 4.8; 5 studies) and a negative LR of 0.38 (95% CI: 0.29, 0.51; 4 studies). Defining impairment as more than 24 points on the HHIE-S and the reference standard as more than 25 dB resulted in a positive LR of 4.3 (95% CI: 1.7, 10.4; 1 study) and a negative LR of 0.76 (no CI reported; 1 study). Defining hearing impairment as more than 24 points on the HHIE-S and a pure tone average threshold of over 40 dB resulted in a positive LR of 4.0 (95% CI: 2.6, 6.2; 2 studies) and a negative LR of 0.64 (95% CI: 0.50, 0.80; 1 study).

**Weber tuning fork test.**

One study met the inclusion criteria. An abnormal Weber test result has an LR of 1.6 (95% CI: 1.0, 2.3) to 1.7 (95% CI: 1.0, 2.9), depending on the tuning fork used (256 or 512 Hz). A normal result lowered the probability, with an LR of 0.70 (95% CI: 0.48, 1.0) to 0.76 (95% CI 0.57, 1.0).

**Rinne tuning fork test.**
Five lower quality studies reporting on multiple definitions presented positive LRs ranging from 2.7 (95% CI: 2.0, 3.5) to 62 (95% CI: 3.9, 970) and negative LRs ranging from 0.01 (95% CI: 0, 0.15) to 0.85 (95% CI: 0.76, 0.95).

Whispered-voice test.

When applying various test definitions and a reference standard of over 30 dB, the positive LR was 6.1 (95% CI: 4.5, 8.4; 3 pooled studies) and the negative LR 0.03 (95% CI: 0, 0.24; 3 pooled studies).

Audioscope.

When applying different test and reference standard definitions, the positive LR was 2.4 (95% CI: 1.4, 4.1; 4 pooled studies) and the negative LR 0.07 (95% CI: 0.03, 0.17; 3 pooled studies).

Authors’ conclusions
Elderly individuals who indicate they have hearing impairment require audiometry, while those who do not should be screened with the whispered-voice test. Those who perceive the whispered voice require no further testing, while those who do not require audiometry. The Weber and Rinne tests should not be used for general screening.

CRD commentary
The research question and inclusion criteria for the diagnostic accuracy studies were clear. The search was limited to two databases and English language publications only, thus it is possible that publication and language bias were present at the study selection stage and that relevant studies were missed. The reviewers took some measures to avoid bias and errors in the selection of studies and data extraction, but this was not reported for the assignment of the level of evidence; this is unfortunate as studies were also selected on the basis of these quality features.

The results of the included studies were presented in detail, providing pooled and individual study results, which was appropriate as the study results differed widely and the pooled results alone may not be clinically meaningful. Statistical heterogeneity was present but no details were reported. The evidence basis for the conclusion about the Weber tuning fork test was one poorly described study. Given the limitations in the search, scope (English language only) and review methodology, the conclusions drawn from the review should be regarded with caution.

Implications of the review for practice and research
Practice: The authors presented a screening algorithm for determining the need for formal audiometric testing and stated explicitly that the Rinne and Weber tuning fork tests should no longer be part of the medical curriculum.

Research: The authors stated that research is required to improve the standardisation of the whispered-voice test.

Bibliographic details
Bagai A, Thavendiranathan P, Detsky A S. Does this patient have hearing impairment? JAMA 2006; 295(4): 416-428

PubMedID
16434632

DOI
10.1001/jama.295.4.416

Original Paper URL
http://jama.ama-assn.org/

Other publications of related interest
These additional published commentaries may also be of interest. Glasziou P. Review: self report of hearing loss and the whispered voice test are useful for screening for hearing impairment. Evid Based Med 2006;11:116. Martin N.
Review: self report of hearing loss and the whispered voice test are useful for screening for hearing impairment. Evid Based Nurs 2006;9:120.

Indexing Status
Subject indexing assigned by NLM

MeSH
Aged; Hearing Loss /diagnosis /physiopathology; Hearing Tests /methods; Humans

AccessionNumber
12006000706

Date bibliographic record published
30/04/2007

Date abstract record published
30/04/2007

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.