

# Primary care

## Whispered voice test for screening for hearing impairment in adults and children: systematic review

Sandi Pirozzo, Tracey Papinczak, Paul Glasziou

### Abstract

**Objective** To determine the accuracy of the whispered voice test in detecting hearing impairment in adults and children.

**Design** Systematic review of studies of test accuracy.

**Data sources** Medline, Embase, Science Citation Index, unpublished theses, manual searching of bibliographies of known primary and review articles, and contact with authors.

**Study selection** Two reviewers independently selected and extracted data on study characteristics, quality, and accuracy of studies. Studies were included if they had cross sectional designs, at least one of the index tests was the whispered voice test, and the reference test (audiometry) was performed on at least 80% of the participants.

**Data extraction** Data were used to form 2×2 contingency tables with hearing impairment by audiometry as the reference standard.

**Data synthesis** The eight studies that were found used six different techniques. The sensitivity in the four adult studies was 90% or 100% and the specificity was 70% to 87%. The sensitivity in the four childhood studies ranged from 80% to 96% and specificity ranged from 90% to 98%.

**Conclusion** The whispered voice test is a simple and accurate test for detecting hearing impairment. There is some concern regarding the lower sensitivity in children and the overall reproducibility of the test, particularly in primary care settings. Further studies should be conducted in primary care settings to explore the influence of components of the testing procedure to optimise test sensitivity and to promote standardisation of the testing procedure.

### Introduction

Hearing impairment is a common problem in elderly people, affecting almost 40% of people over the age of 60 and 90% over the age of 80.<sup>1-3</sup> If not detected and treated, it can have an appreciable impact on the social and emotional functioning of the individual.<sup>4-5</sup> These negative effects can, however, be reversed once hearing impairment is detected and treated.<sup>6</sup> The prevalence of permanent hearing impairment in children is relatively low—from 1% in 3 year olds to 2% in children aged 9-16 years.<sup>7</sup> In addition, at any given time, 5-7% of young children have a temporary 25 dB hearing loss

associated with otitis media with effusion.<sup>8</sup> Hearing impairment can severely affect young lives by retarding language acquisition and cognitive development.<sup>9</sup> Screening tests provide a quick and cost effective way to separate people into two groups: those who pass the screening test and are presumed to have no hearing loss and those who fail the screening test and are in need of an in-depth evaluation by an audiologist. Thus, screening for hearing impairment in elderly people and children is an integral part of overall health assessment and can be accomplished with a variety of simple tests conducted in the office.

To some extent the utility of the various hearing screening tests depends on the age of the patient and whether hearing loss is sensorineural or conductive. Sensorineural loss results from damage to neural structures and is most commonly due to degenerative hearing loss of ageing (presbycusis).<sup>10</sup> Conductive hearing loss results from interference with the conduction of sound vibrations and is most commonly caused by impacted cerumen, otitis media, or otosclerosis. Some of the simple tests that can be used to screen for hearing loss include patients' self report, tuning fork tests, a rubbing sound from the examiner's fingers, and the whispered voice test. All are relatively easy to perform, but their accuracy varies. Self report has the lowest sensitivity in detecting hearing impairment (71%) and is unlikely to be useful in young children.<sup>11</sup> Tuning fork tests are most effective in detecting conductive hearing loss, with a sensitivity of 60-90%, but their accuracy depends on the experience of the tester.<sup>12</sup> Because the tuning fork test evaluates hearing at a single low frequency, it is not appropriate for most elderly patients with presbycusis, who typically have lost the ability to hear high frequencies.<sup>13</sup> Although the finger rub test has not been evaluated extensively, one study found a sensitivity of 80% in elderly ambulatory patients.<sup>14</sup>

Of all the simple hearing tests, the whispered voice test (box 1) is the only one that has been studied in both children and adults. It can be used for detecting both types of hearing loss and its performance compares favourably with the portable audioscope, which has a sensitivity of 87-96% and a specificity of 70-90%.<sup>12</sup> Currently, general practitioners in many Western countries (including the United Kingdom and Australia) are advised by national health guidelines to screen for hearing impairment in the elderly population, and the

School of Population Health, University of Queensland, Royal Brisbane Hospital, Herston, Qld 4029, Australia

Sandi Pirozzo  
senior lecturer in epidemiology

Tracey Papinczak  
research officer

Paul Glasziou  
professor of evidence based practice

Correspondence to:  
s.pirozzo@sph.uq.edu.au

bmj.com 2003;327:967

**Box 1: Conducting the whispered voice test<sup>15-17</sup>**

- The examiner stands arm's length (0.6 m) behind the seated patient and whispers a combination of numbers and letters (for example, 4-K-2) and then asks the patient to repeat the sequence
- The examiner should quietly exhale before whispering to ensure as quiet a voice as possible
- If the patient responds correctly, hearing is considered normal; if the patient responds incorrectly, the test is repeated using a different number/letter combination
- The patient is considered to have passed the screening test if they repeat at least three out of a possible six numbers or letters correctly
- The examiner always stands behind the patient to prevent lip reading
- Each ear is tested individually, starting with the ear with better hearing, and during testing the non-test ear is masked by gently occluding the auditory canal with a finger and rubbing the tragus in a circular motion
- The other ear is assessed similarly with a different combination of numbers and letters

whispered voice test is one of the tests recommended. Its potential utility in both adults and children, particularly in developing countries that may have limited access to standard audiometric facilities, is promising. Several studies have shown that it is sufficiently accurate for detecting hearing impairment in adults,<sup>15-17</sup> but there is disagreement about the appropriate technique and the value of the test in children.<sup>18-19</sup> This systematic review synthesises the literature on the accuracy of the test in detecting hearing impairment.

## Methods

### Search strategy

We identified eligible studies by searching Medline, Embase, and Science Citation Index from the beginning of each database until June 2002. We also searched the web for unpublished theses and perused bibliographies of known primary and review articles to identify studies not found through electronic searching. In addition, several authors of relevant papers were contacted to inquire about possible unpublished studies and to clarify questions about the data contained in their published studies.

The search strategy (box 2) included terms for the index test, the reference test, the patient problem and a methodological filter. Both MeSH and text words were used.

### Study selection and data extraction

To be included, studies had to be cross sectional studies in which at least one of the index tests was the whispered voice test and the reference test, audiometry, was performed on at least 80% of the participants. The sensitivity and specificity of the test needed to be reported or calculable from the data provided.

One of the authors (SP) initially screened the titles and abstracts of the search results. Once full manuscripts of all relevant papers were obtained, two reviewers (SP, TP) independently reviewed each paper for inclusion according to the predefined inclusion criteria and extracted data by using a specially designed

data extraction form. In cases of duplicate publication we selected the most complete version of the study. Since there were no language restrictions on the search strategy, two colleagues familiar with diagnostic test methodology and the non-English language in question helped with study selection. Any differences between reviewers in relation to study selection or data extraction were resolved by the third reviewer (PG).

We extracted data on study characteristics, study quality, and accuracy of results from each selected paper. Study characteristics consisted of patients' characteristics and the procedures used to conduct the whispered voice test and audiometry. We divided participants into two groups: adults ( $\geq 17$  years) and children ( $< 17$  years). The primary outcome measure of interest was the accuracy of the test as reflected by its sensitivity and specificity. Wherever possible we used the raw data to construct  $2 \times 2$  tables and calculate sensitivity and specificity. If insufficient raw data were available to calculate measures of accuracy, we used the measures provided by the authors of the paper. In relation to the quality of the studies, we assessed whether the method of sampling was consecutive or random; whether the comparisons between the index test and the reference test were independent, and whether they were blinded; whether the adequacy of the test descriptions would allow replication; and whether there was at least 80% verification with the reference test.

## Results

### Literature identification and study quality

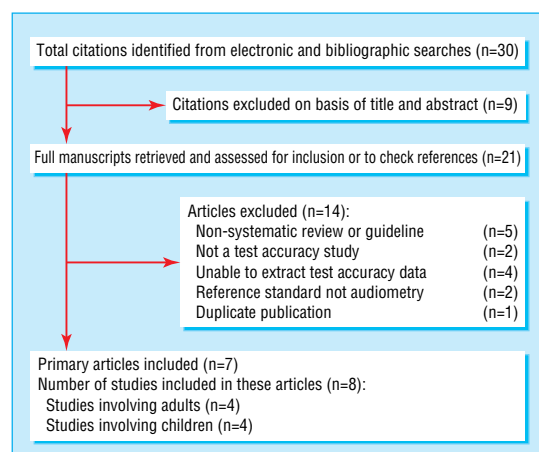
The literature search identified 17 primary studies (from 16 articles), four non-systematic reviews, and one guideline summary. Figure 1 summarises the process of study selection.

Of the studies identified, only eight English language studies (seven articles) met all of the predefined inclusion criteria. Of these, four studies included a total of 290 adults (ages 17-89)<sup>14-17</sup> and four studies included 716 children (aged 3-12 years).<sup>18-20</sup>

Overall, the methodological quality of the studies was modest, with many important elements not reported.<sup>21</sup> For example, in most studies it was either unclear or simply not stated whether comparison between the whispered voice test and audiometry had been blind and independent, and none of the childhood studies reported using consecutive or random sampling. On the basis of information

**Box 2: Search strategy**

1. whisper\*
2. explode 'Audiometry'/all subheadings in MIMe,MJME
3. explode 'Hearing-Tests'/methods in MIMe,MJME
4. explode 'Hearing-Loss-Partial'/diagnosis in MIMe,MJME
5. audiometr\*[tw]
6. hearing adj test[tw]
7. #2 or #3 or #4 or #5 or #6
8. explode 'Sensitivity-and-Specificity'/all subheadings in MIMe,MJME
9. sensitivity and specificity[tw]
10. #8 or #9
11. (#1 and #7) or (1 and #10)



**Fig 1** Selecting studies of whispered voice test for systematic review

provided in the articles, only one study fulfilled all five quality criteria.<sup>17</sup> The studies conducted in children were of poorer quality than studies in adults (table 1).

### Whispered voice test in adults

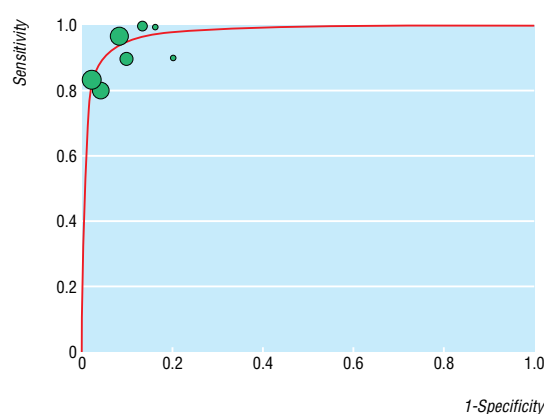
Table 2 shows the characteristics of the four studies examining the accuracy of the whispered voice test in adults. The participants were generally elderly; only one study included participants younger than 55 years.<sup>15</sup> The prevalence of hearing impairment ranged from 26% to 61%. Three studies used similar techniques for the whispered voice test and a 30 dB positivity threshold for hearing impairment by audiometry.<sup>15–17</sup> The fourth study used a different technique for the test and a 40 dB positivity threshold, and its results were reported in such a way that it was not possible to calculate an overall sensitivity and specificity, although specificities for sensitivities of 80% and 90% were provided.<sup>14</sup> In addition, the distance from the examiner to the person's ear was less than half the distance in the other studies (11 inches (28 cm) *v* 24 inches (61 cm)). In the three comparable studies the sensitivity of the whispered voice test was either 90% or 100% and specificity ranged from 80% to 87%. Positive likelihood ratios ranged from 4.6 to 7.7, showing that a positive test is moderately strong in ruling in hearing impairment. Negative likelihood ratios were zero or close to it, showing no hearing impairment when the test is negative.

### Whispered voice test in children

Table 3 shows the four studies examining the accuracy of the whispered voice test in children. The children were aged from 3 to 12 years and the prevalence of hearing impairment ranged from 9% to 31%. All of the studies used slightly different techniques to conduct

the whispered voice test and the threshold for hearing impairment by audiometry ranged from 20 dB to 35 dB. Only one study used digits and letters<sup>18</sup>; the other three studies used spondee words (two-syllable words with equal stress on the syllables—for example, baseball). In the two studies involving younger children, the technique for children under 6 years was slightly modified: rather than repeating the spondee word, the children were asked to point to a picture representing the word.<sup>19</sup>

Overall, the whispered voice test in children was less sensitive but more specific than in adults (sensitivity 80–96%, specificity 90–98%). As compared to the adult studies, the positive likelihood ratios were higher than in the studies on adults, so a positive test argues even more strongly for hearing impairment. However, the higher negative likelihood ratios were less convincing in ruling out disease. Figure 2 plots the individual study results and the receiver operator characteristics curve for all seven studies.



**Fig 2** Receiver operator characteristics curve for individual study results for the whispered voice test in predicting hearing impairment in adults and children. Size of the bubble is proportional to sample size; the four largest bubbles represent the four childhood studies

### Reliability of whispered voice test

Several studies also examined the reliability or reproducibility of the whispered voice test. Uhlmann compared the results of an otolaryngologist and an audiologist for 63% of the patients and found a correlation of 0.67,<sup>14</sup> and Macphée et al found concordance between a geriatrician and an otolaryngologist of 0.88.<sup>17</sup> However, in the study by Eekhof et al, where the results of six examiners were compared with those of the first examiner, the specificity ranged from 14% to 100% and the interobserver reliability (measured by Cohen's  $\kappa$ ) ranged from 0.16 to 1.0.<sup>16</sup> The authors attributed the

**Table 1** Methodological quality (\*="adequate" or better) of studies of whispered voice test according to five quality criteria

Criterion	Studies of adults				Studies of children			
	Macphée et al <sup>17</sup>	Uhlmann et al <sup>14</sup>	Swan and Browning <sup>15</sup>	Eekhof et al <sup>16</sup>	Groen <sup>20</sup>	Dempster and Mackenzie <sup>18</sup>	Prescott et al <sup>19</sup>	
Consecutive or random sampling	*Yes	*Yes	**"All patients"	**"All patients"	Not stated	Not stated	Not stated	Not stated
Independent assessment	*Claimed	*Claimed	Not stated	Not stated	*Claimed and named	Not stated	No	No
Blind assessment	*Yes	*Yes	Not stated	Not stated	Not stated	Not stated	No	No
Adequate test description	*Good	*Adequate	*Good	Poor	*Adequate	*Adequate	*Adequate	*Adequate
>80% verification with audiometry	*100%	*100%	*100%	*100%	*100%	*98%	*94%	*98%

**Table 2** Characteristics of participants, methods of testing and results of studies of use of whisper test in adults

Criterion	Swan and Browning <sup>15</sup>	Macphie et al <sup>17</sup>	Eekhof et al <sup>16</sup>	Uhlmann et al <sup>14</sup>
<b>Participants and setting</b>				
Number of participants (No of ears)	101 (202)	62 (124)	62 (124)	31 (62) (non-demented group only)
Age (years)	Mean 57, range 17-89	Mean 81, range 66-96	≥55	Mean 76, SD 5.8
Setting	Audiology Clinic	Geriatric Assessment Unit	ENT Outpatient Department	Otolaryngology Clinic
<b>Whispered voice technique</b>				
Audiometry cut-off for hearing impairment	30 dB	30 dB	30 dB	40 dB
Distance	2 feet (arm's length)	6 inches and 2 feet	"Standard method"*	Began at 6 inches and move away at 6 inch increments
Masking	Occlusion and rubbing of external auditory canal	Occlusion and rubbing of external auditory canal	"Standard method"*	Pushing tragus over external auditory canal
Letters/numbers/words	Triplets of letters and numbers, eg 5B6	3 random numbers	"Standard method"*	Spondee words†
Number of repeats	1	1 or 2	"Standard method"*	At each distance until unable to hear
Threshold for hearing impairment	<50% correct	≤50% correct	≥2 combinations incorrect	<50% correct
<b>Results of testing</b>				
Prevalence of hearing impairment (%)	43	61	59	26
Sensitivity (%; 95% CI)	100 (96-100)	100 (95-100)	90 (82-95)	90‡
Specificity (%; 95% CI)	87 (80-92)	84 (71-92)	80 (68-89)	70‡
Likelihood ratios:				
Positive	7.7	6.4	4.6	Not done (insufficient data)
Negative	0	0	0.12	Not done (insufficient data)

\*"The whispered voice test was performed by a standard method (slightly modified)"—reference to Swan, 1985

†Spondee words are two syllable words with equal stress on both syllables, for example "baseball" or "hotdog."

‡Specificities were given for two sensitivities (80% and 90%)—unable to calculate confidence interval or likelihood ratio.

broad variation between examiners' outcomes to the difference in loudness of the whispering, a supposition that was supported by patients who spontaneously complained about the quiet whispering of several examiners.

## Discussion

The whispered voice test is a simple and accurate test for detecting hearing impairment and compares favourably with the portable audioscope. Despite some

variations in the methodology of studies and the populations sampled, findings are relatively consistent.

One area of concern is the reproducibility of the whispered voice test. The results of the studies that measured reliability indicate that the test can be reliable if a standard procedure is used. At the moment there is considerable room for improvement in standardising the technique of conducting the test and in setting the threshold for hearing impairment by the whispered voice test.

**Table 3** Characteristics of participants, methods of testing, and results of studies of use of whisper voice test in children

Criterion	Groen <sup>20</sup>	Dempster and Mackenzie <sup>18</sup>	Prescott et al <sup>19</sup>	
			Study 1	Study 2
<b>Participants and setting</b>				
Number of participants (No of ears)	197 (394)	141 (282)	177 (354)	201 (402)
Age in years	Mean 5	Mean 7.2, range 5 to 12	Range 3-12	Range 3-7
Setting	Audiology centre	Otorhinolaryngology clinic	Hospital	School classroom
<b>Whispered voice technique</b>				
Audiometry cut-off for hearing impairment	25 dB	30 dB (also 20 dB and 25 dB)	20 dB	35 dB
Distance	3 m, 1 m, 0.3 m	6 inches, 2 feet	2 feet	
Masking	Finger occlusion of external auditory canal	Tragal rubbing	Tragal pressure but only if hearing loss was asymmetrical	
Letters, numbers, or words	Spondee words (10)	3 numbers or 3 numbers and letters	Spondee words	
			<6 yrs—point to picture	
			≥6 yrs—repeat word	
Threshold for hearing impairment	<50% correct at 3 m	<2/3 correct at 2 feet	<50% correct	
<b>Results of testing</b>				
Prevalence (%) of hearing impairment	14	13	31	9
Sensitivity (%; 95% CI)	96 (82-99)	90 (69-97)	80 (68-88)	83 (61-94)
Specificity (%; 95% CI)	92 (87-95)	90 (84-94)	96 (91-98)	98 (95-99)
Likelihood ratio:				
Positive	11.6	9.1	19.5	38.1
Negative	0.04	0.12	0.21	0.17



The most appropriate letters, numbers, or words for testing also needs further investigation. In the elderly population, where presbycusis is the most common type of hearing loss, difficulty in hearing sounds in the higher frequencies is common. As the consonants of speech are usually higher frequency sounds than the vowels,<sup>22</sup> using different consonants and vowels in testing could alter the results of the test considerably.

The greatest difficulty in standardising the test is the loudness of the whisper. However, only a few studies in the review mentioned that the whispered sequence occurred after a full expiration. This seems to be an important determinant of the loudness of the whisper.

Applying the findings from this review raises other concerns, particularly in children. With the test sensitivity much lower in children than adults, it might be argued that the test is of limited value in children, as it would fail to identify hearing impairment in a large proportion of children. Why this difference in sensitivity between the adult and childhood studies exists is unclear. Although the overall quality of the childhood studies was rated lower than that of the adult studies, this may have been due to lack of detail reported, rather than to less rigorous methods. Technique also differed in terms of the spoken sequences (spondee words versus letters or numbers), the distance between the examiner and the patient, and the threshold for hearing impairment in both the reference test and the whispered voice test. Which of these components of the testing procedure needs to be modified in order to optimise sensitivity is not known. Further studies are needed that compare the diagnostic accuracy of the whispered voice test when different methods are used in younger and older children.

In most Western countries, national health guidelines encourage general practitioners to screen elderly people for hearing loss. The whispered voice test is one test recommended for this screening, yet it has not been adequately evaluated in primary care settings. None of the studies in this review were conducted in primary care settings, and few of the clinicians performing the tests were general practitioners. Thus, future research into the utility of the whispered voice test should be conducted by general practitioners in primary care settings.

We thank Fred de Looze and Eva Pietrzak of the University of Queensland for translating the non-English papers.

Contributors: SP designed the search strategy, participated in the search, designed the data extraction form, reviewed all abstracts and papers, selected studies for inclusion, extracted data, contacted authors of papers, analysed and interpreted data, and wrote and revised the paper. TP conducted the literature search, retrieved full text papers, reviewed all abstracts and papers, selected studies for inclusion, extracted data, and reviewed the manuscript before submission. PG conceived of the study, modified the search strategy, acted as third reviewer in decisions relating to study inclusion and data extraction, helped analyse and interpret data, and modified the paper. SP is guarantor.

Funding: National Health and Medical Research Council programme grant 211205, which supports the Screening and Test Evaluation Program.

Competing interests: None declared.

Ethical approval: Not needed.

## What is already known on this topic

Screening for hearing impairment has been recommended by national health guidelines as an integral part of overall health assessment

The whispered voice test is one of the few simple screening tests that have been evaluated in both adults and children

## What this study adds

The whispered voice test is an accurate and simple test of hearing impairment that could be used by general practitioners but has not been adequately evaluated in primary care settings

Differences in accuracy among published studies could be explained by differences in conducting the test

The technique for conducting the test needs to be standardised to optimise sensitivity of the test, particularly in children

- 1 Abutan BB, Hoes AW, Van Dalsen CL, Verschuure J, Prins A. Prevalence of hearing impairment and hearing complaints in older adults: a study in general practice. *Fam Pract* 1993;10:391-5.
- 2 Sindhusake D, Mitchell P, Smith W, Golding M, Newall P, Hartley D, et al. Validation of self-reported hearing loss. The Blue Mountains hearing study. *Int J Epidemiol* 2001;30:1371-8.
- 3 Cruickshanks KJ, Wiley TL, Tweed TS, Klein BE, Klein R, Mares-Perlman JA, et al. Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The epidemiology of hearing loss study. *Am J Epidemiol* 1998;148:879-86.
- 4 Wallhagen MI, Strawbridge WJ, Shema SJ, Kurata J, Kaplan GA. Comparative impact of hearing and vision impairment on subsequent functioning. *J Am Geriatr Soc* 2001;49:1086-92.
- 5 Strawbridge WJ, Wallhagen MI, Shema SJ, Kaplan GA. Negative consequences of hearing impairment in old age: a longitudinal analysis. *Gerontologist* 2000;40:320-6.
- 6 Mulrow CD, Aguilar C, Endicott JE, Tuley MR, Velez R, Charlip WS, et al. Quality-of-life changes and hearing impairment. A randomized trial. *Ann Intern Med* 1990;113:188-94.
- 7 Fortnum HM, Summerfield AQ, Marshall DH, Davis AC, Bamford JM. Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment study. *BMJ* 2001;323:536-40.
- 8 Cross AW. Health screening in schools. Part I. *J Pediatr* 1985;107:487-94.
- 9 Davis A, Hind S. The impact of hearing impairment: a global health problem. *Int J Pediatr Otorhinolaryngol* 1999;49(suppl 1):S51-4.
- 10 Nadol JB Jr. Hearing loss. *N Engl J Med* 1993;329:1092-102.
- 11 Nondahl DM, Cruickshanks KJ, Wiley TL, Tweed TS, Klein R, Klein BE. Accuracy of self-reported hearing loss. *Audiology* 1998;37:295-301.
- 12 Burkey JM, Lippy WH, Schuring AG, Rizer FM. Clinical utility of the 512-Hz Rinne tuning fork test. *Am J Otol* 1998;19:59-62.
- 13 Mulrow C. Screening for hearing impairment in the elderly. *Hosp Pract* 1991;26:79, 83-6.
- 14 Uhlmann RE, Rees TS, Psaty BM, Duckert LG. Validity and reliability of auditory screening tests in demented and non-demented older adults. *J Gen Intern Med* 1989;4:90-6.
- 15 Swan IR, Browning GG. The whispered voice as a screening test for hearing impairment. *J R Coll Gen Pract* 1985;35:197.
- 16 Eekhof JA, de Bock GH, de Laat JA, Dap R, Schaapveld K, Springer MP. The whispered voice: the best test for screening for hearing impairment in general practice? *Br J Gen Pract* 1996;46:473-4.
- 17 Macphree GJ, Crowther JA, McAlpine CH. A simple screening test for hearing impairment in elderly patients. *Age Ageing* 1988;17:347-51.
- 18 Dempster JH, Mackenzie K. Clinical role of free-field voice tests in children. *Clin Otolaryngol* 1992;17:54-6.
- 19 Prescott CA, Omoding SS, Fermor J, Ogilvy D. An evaluation of the "voice test" as a method for assessing hearing in children with particular reference to the situation in developing countries. *Int J Pediatr Otorhinolaryngol* 1999;51:165-70.
- 20 Groen JJ. Pure tone audiometry and whispered voice test. Conformities and differences in tests results. *ORL J Otorhinolaryngol Relat Spec* 1973;35:65-70.
- 21 STAARD Group. Towards complete and accurate reporting of studies on diagnostic accuracy: The STAARD initiative. November 2001. [www.consort-statement.org/stardstatement.htm](http://www.consort-statement.org/stardstatement.htm) (accessed 26 Aug 2003).
- 22 Fleming KC, Evans JM, Weber DC, Chutka DS. Practical functional assessment of elderly persons: a primary-care approach. *Mayo Clin Proc* 1995;70:890-910.

(Accepted 5 August 2003)