

고연령층 청력 선별 검사로서 귓속말 검사의 유용성 및 신뢰도

가톨릭대학교 의과대학 이비인후과학교실,¹ 예방의학교실,⁵ 아주대학교 의과대학 이비인후과학교실,² 성균관대학교 의과대학 이비인후과학교실,³ 서울대학교 의과대학 이비인후과학교실⁴

박시내¹ · 정연훈² · 정원호³ · 오승하⁴ · 박용문⁵ · 이준호⁴
박현이² · 최성준² · 여상원¹ · 이원철⁵ · 박기현²

Validity and Reliability of Whispered Voice Test as a Screening Test for Hearing in the Elderly People

Shi-Nae Park, MD, PhD¹, Yun-Hoon Choung, DDS, MD, PhD², Won-Ho Chung, MD³, Seung-Ha Oh, MD, PhD⁴, Yong-Moon Park, MD, PhD⁵, Jun Ho Lee, MD, PhD⁴, Hun Yi Park, MD², Seong Jun Choi, MD², Sang Won Yeo, MD, PhD¹, Won Chul Lee, MD, Dr. Med Sci⁵ and Keehyun Park, MD, PhD²

¹Department of Otolaryngology-HNS; ⁵Preventive Medicine, The Catholic University of Korea College of Medicine, Seoul; and

²Department of Otolaryngology-HNS, Ajou University School of Medicine, Suwon; and

³Department of Otorhinolaryngology-HNS, Sungkyunkwan University School of Medicine, Seoul; and

⁴Department of Otorhinolaryngology, Seoul National University College of Medicine, Seoul, Korea

ABSTRACT

Background and Objectives : The purpose of this study is to evaluate the validity and reliability of the whispered voice test, which has been suggested as a screening test tool for hearing in the elderly group. **Subjects and Method** : A prospective multicenter trial has been designed. One hundred and seventy two elderly subjects from four university hospitals were included in this study. A whispered voice test using a number list and number/word list was performed by two testers for each elderly subject. To evaluate their test-retest reliability, a second whispered voice test was done 1–2 weeks later. The result of the whispered voice test was analyzed with the criteria of 26 dB and 40 dB hearing level, checked by pure tone audiometry. Sensitivity, specificity, and cut-off point of whispered voice test at each gold standard, inter-observer reliability and test-retest reliability have been investigated. **Results** : There were significant differences with respect to correctly answered numbers in the whispered voice test between the normal and impaired hearing group at both criteria of 26 dBHL and 40 dBHL. With the 26 dB criteria of hearing level which showed higher inter-observer reliability of the number/word list, the whispered voice test has 86.0% sensitivity and 58.9% specificity with the cut-off point of 3 out of six on the number set. The test-retest reliability of whispered voice test using the number list was as high as 0.9 of correlation coefficient. However, the inter-observer reliability of whispered voice test was less than 0.6 (Cohen's kappa value). **Conclusion** : As a screening test for hearing in the elderly group, the whispered voice test has quite high sensitivity and test-retest reliability. The environmental sound level and test method for each tester, however, should be monitored to get better specificity and inter-observer reliability. (Korean J Otorhinolaryngol-Head Neck Surg 2008;51:609–16)

KEY WORDS : Hearing test screening · Elderly · Validity · Reliability.

서 론

난청은 60세 이상 인구의 40%, 80세 이상 인구의 90%에서 발생하는 매우 높은 발병률을 보이는 질환이다.¹⁻³⁾ 노인에서 난청은 의사소통의 어려움으로 인한 사회적 고립, 우

논문접수일 : 2008년 1월 8일 / 심사완료일 : 2008년 5월 6일

교신저자 : 박시내, 137-701 서울 서초구 반포동 505

가톨릭대학교 의과대학 이비인후과학교실

전화 : (02) 590-2215 · 전송 : (02) 595-1354

E-mail : snparkmd@catholic.ac.kr

울증, 활동 감소나 동반된 정신과적 문제의 악화 등과 밀접한 연관성이 있으며, 난청의 선별 검사를 통한 조기 발견과 적절한 재활치료는 노인들의 건강과 삶의 질에 큰 도움을 줄 수 있다고 알려져 있다.⁴⁻⁶⁾

난청의 선별 검사는 빠르고 쉬우면서 비용 효과와 신뢰성이 높아야 한다. 현재까지 알려진 노인성 난청의 선별 검사로는 설문지를 활용한 환자 자가 보고법, 음차 검사법, 손가락 마찰 검사(Finger rubbing test), 휴대용 청력기를 이용한 검사 및 귓속말 검사법 등이 알려져 있으며 이들 검

사는 비교적 간단하나 정확도에 대한 보고는 일정치 않다. 특히 설문지를 이용한 자가 보고법은 여러 선별 검사법 중 71%의 가장 낮은 민감도를 보인 것으로 보고되고 있고, 음차 검사법도 60~90%의 민감도와 검사자의 경험에 따라 정확도가 많은 차이를 보이고 있으며, 한 주파수의 음차를 이용한 음차 검사법은 다양한 난청의 유형을 보이는 노인성 난청 환자의 선별 검사로써 한계가 있다고 지적된 바 있다.⁷⁻⁹⁾ 국내에서는 한국어로 번역한 HHIE-S(Hearing Handicap Inventory for the Elderly-Screening Version)와 SHS(Smith Hearing Screening) 설문을 이용한 난청의 선별 검사시 민감도와 특이도가 70~80% 수준으로 보고된 바 있는데 이들 연구에서는 휴대용 청력 검사기를 이용한 청력 검사 결과를 기준으로 하여 선별 검사 결과를 분석하였는데 이 경우 청력기준치(criteria) 자체의 신뢰도에 문제가 있을 수 있고, 사회생활 적응 수준인 순음청력도의 40 dBHL만을 기준으로 결과를 보고하여 난청의 시작 시점인 순음청력역치 26 dB를 기준으로 한 진정한 의미의 난청 선별 검사로서의 가치를 살피지 못했던 점, 한 기관에서만 연구가 시행되어 선별 검사시 발생할 수 있는 다양한 환경과 검사자에 대한 다각도에서의 분석이 이루어지지 못한 한계점이 있었다.^{10,11)}

난청 선별 검사로서 손가락 마찰 검사(Finger rubbing test)는 그 유용성과 신뢰도에 대한 연구가 부족하며 한 연구에서 약 80%의 민감도를 보인다고 보고된 바 있다.¹²⁾ 휴대용 청력 검사기를 이용한 난청 선별 검사 선행 연구결과에서도 부적절한 기술, 검사 장소의 소음, 대상자의 고의적 보고 오류 등이 문제점으로 지적되면서 선별 검사로서 비교적 큰 오차가 보고된 바 있다.¹³⁾

귓속말 검사는 소아와 노인에서 난청의 선별 검사로 가장 많은 연구가 이루어진 검사법으로 87~96%의 민감도와 70~90%의 특이도가 보고되고 있으며, 영국과 호주를 포함한 많은 구미 국가에서 노인의 난청 선별을 위한 국가간 강지침에 의해 추천되는 검사로 알려져 있다.^{8,14,15)} 그러나 아직까지 국내에서는 다양한 검사 환경과 요소를 감안한 다기관 평가를 통해 난청 선별 검사로서 귓속말 검사의 유용성이 연구된 바 없었다.

저자들은 2006년 국가 사업으로 지정된 생애전환기 건강검진사업에서 66세 전국민을 대상으로 시행하게 될 난청의 선별 검사로서 국가 기관에서 제안된 귓속말 검사의 검사 방법과 그 유용성 검증을 보건복지부로부터 요청받았기에 대한정각학회 산하 노인성난청선별검사 준비위원회를 구성한 후 귓속말 검사 방법을 제작하고 다기관 연구를 통해 본 검사법의 유용성과 적절성을 검증하고자 연구를 시행하였다.

대상 및 방법

대 상

3차 의료 기관인 네 곳의 의과대학 부속병원에서 2006년 1월부터 4월까지 이비인후과 외래를 내원한 환자들 중 귀 증상을 호소하여 이과 진료를 받은 60세 이상의 노인들 중 난청의 선별 검사로 시행하게 될 귓속말 검사 연구에 서면 동의한 환자들만을 대상으로 하였다. 본 연구는 임상시험윤리위원회의 승인을 받고 시행하였다.

방 법

검사 환경과 검사자의 특성 분석

본 연구가 지닌 다기관 연구의 특성을 감안하여 결과분석과 보정에 활용하고자 연구 시작 전에 한 명의 대표 연구자가 각 병원을 방문하여 검사를 시행하게 될 검사실 혹은 외래의 소음 정도와 검사자 8명의 귓속말 검사시 내게 되는 소리의 크기를 소리수준측정기(Sound level meter ; Quest Technologies, Model 2700, Impulse Sound Level Meter, Oconomowoc, Wisconsin)를 이용하여 측정하였다. 검사실과 외래는 가능한 향후 귓속말 검사를 시행하게 될 환경과 동일한 조건을 갖추도록 한 후 5초 간격으로 7차례 측정된 환경음의 소리 크기를 평균하여 검사 장소의 환경음으로 하였다. 귓속말의 소리 크기는 검사자들에게 미리 나누어준 귓속말 검사법 매뉴얼(Fig. 1)을 읽고 숙지하게 한 다음 각 검사자에게 숫자 조합 리스트 한 세트와 숫자-글자 조합 리스트 한 조합을 반복하여 읽게 하면서 검사자의 소리가 들리는 순간에 국한하여 7차례 5초 간격으로 귓속말 소리 크기를 측정하고 그 값을 평균한 후, 미리 구한 환경음의 평균 값을 뺀 값을 귓속말 검사시의 소리 크기로 하였다.

귓속말 검사법 매뉴얼의 개발

난수표를 활용하여 숫자 0~9, 글자 가-자까지를 무작위 추출하는 방식으로 3개의 숫자 조합 리스트와 숫자-글자-숫자 조합 리스트(Fig. 2) 세트를 제작하였다. 두 종류의 검사 리스트 중 보다 나은 난청 선별 검사 리스트를 골라 제안하기 위해 본 연구에서는 두 리스트 모두를 활용하여 귓속말 검사를 시행하되 숫자 조합 리스트를 이용한 귓속말 검사를 먼저 시행하여 결과를 기록하고, 이후 숫자-글자-숫자 조합 리스트를 이용한 귓속말 검사를 다시 시행하여 기록하는 방식을 선택하였다. 검사 방법은 국외 선행연구에서 제안된 바 있는 방식을 따랐다.^{8,14,15)} 구체적인

방법을 살펴보면 우선 잘 들리는 귀를 먼저 검사하되 한쪽 귀를 검사할 때 반대편 귀는 검사자가 환자의 외이도를 막고 손가락으로 이주(tragus)를 둥글게 문지름으로써 차폐

하도록 하였다. 검사 대상자가 검사자의 입술을 보지 못하도록 환자의 60 cm(한 팔 간격) 뒤에 서서 검사를 시행하며 검사자는 최대한 완전히 숨을 내 쉰 후 자신이 낼 수 있는 최소 크기의 속삭이는 귓속말 목소리로 검사 리스트 중 한 세트(한 줄)를 골라 3개씩 나누어 들려주고 이를 대상자에게 정확히 따라하게 하며 한쪽 귀당 6개 숫자 혹은 숫자-글자-숫자 조합 중 맞춘 개수를 기록하게 하였고, 반대편 귀도 동일한 방법으로 검사하고 기록하게 하였다.^{16,17)}

*** 귓속말 검사법**

1. 검사자가 환자의 60 cm 뒤에 서서 입모양을 보여주지 않는 상태에서 환자에게 물어보아 잘 들리는 귀를 먼저 검사한다.
2. 검사하는 동안 반대편 귀는 검사자가 외이도를 손가락으로 막고 이주(tragus)를 둥글게 문지름으로써 차폐하도록 한다.
3. (숫자 조합 검사)
최대한 속삭이는 조용한 소리를 내기 위해 검사자는 완전히 숨을 다 내친 후 낼 수 있는 최대한 조용한 속삭이는 목소리로 환자에게 숫자 조합 리스트에 있는 3개의 숫자 묶음 10세트(10줄)의 조합 중 한세트(한 줄 : 12개의 숫자 조합임)를 골라 먼저 숫자 3개를 연속적으로 불러주고 환자에게 따라하도록 지시한다. 다음 숫자 3개를 들려주고 따라하도록 지시한다. 숫자 조합에서 검사한 6개 숫자 중 환자가 정확히 맞춘 숫자 개수를 숫자 조합 맞춘 갯수란에 기록한다. 반대편 귀도 동일하게 검사한다.
4. (숫자-글자-숫자 조합 검사)
위 검사와 동일하게 하되 숫자 글자 조합 리스트에서 10세트 중 한 세트(한 줄)를 골라 양측 귀에 각각 들려준 6개 숫자-글자 중 맞춘 개수를 각각 숫자-글자 조합 검사 맞춘 갯수란에 기록한다.

Fig. 1. Manual for whispered voice test method.

귓속말 검사의 선별 검사로서의 유용성 검증

귓속말 검사는 비교적 조용한 이비인후과 외래 혹은 청력 검사실에서 시행하였고 모든 검사 대상자에게 숫자 조합 리스트와 숫자-글자-숫자 조합 리스트를 이용하여 각각 검사를 하였다. 검사 대상자들은 순음청력검사를 하였고 회화음역인 500, 1,000, 2,000, 4,000 Hz에서의 기도 순음청력역치 산술평균값을 구하였다.

순음청력평균값에서 경도 난청 기준인 26 dBHL와 사회적응청력(socially serviceable hearing level) 수준인 40 dBHL 기준으로 한 정상 청력귀와 난청귀를 나누고 귓속말 검사시 6개의 숫자 조합 혹은 숫자-글자-숫자 조합에서 정확히 따라한 정답 수를 난청 기준에 따라 비교해 보았고, 각 청력 기준치에서 난청 선별 검사로서의 민감도와 특이도를 구하여 난청 선별 검사로서의 적절한 기준점(cut-off point)를 제시하고자 하였다.

귓속말 검사의 선별 검사의 신뢰도 검증

검사자간 신뢰도(Inter-observer reliability)

귓속말 검사의 선별 검사로서 신뢰도 검증을 위해 동일 검사 대상자에게 의사/간호사 혹은 청각사로 구성된 두 명의 검사자가 같은 날 검사를 시행하였고 6개의 숫자 혹은 숫자-글자 조합에서 정확히 따라한 수를 비교 분석함으로써 검사자간 신뢰도를 검증하였다. 검사자간 신뢰도는 Cohen's k 값을 측정하여 검증하였다.

검사-재검사 신뢰도 검증(Test-retest reliability)

검사 시기에 따른 환자 및 각 검사자 내 검사 반응 혹은 검사 방법의 차이가 귓속말 검사에 미치는 영향을 확인하기 위해 검사-재검사 신뢰도 검증을 하였다. 1차 검사가 끝난 환자에게 1~2주일 후 다시 내원하게 한 후 두 명의 검사자가 다시 귓속말 검사를 시행하였으며 1~2주 전 시행했던 검사결과와 비교 분석함으로써 검사-재검사 신뢰도를 검증하였다.

귓속말 검사 리스트
(대한청각학회 노인성난청선별검사준비위원회-2007)
* 숫자는 “일, 이, 삼, ...” 등으로 읽어주시기 바랍니다.

(1) 숫자 조합 리스트

왼쪽 귀	오른쪽 귀
268 157	124 397
642 871	439 572
596 348	643 975
751 893	461 852
753 921	472 856
521 438	235 681
164 532	418 956
473 625	538 962
761 942	423 819
279 135	178 356

(2) 숫자-글자-숫자 조합 리스트

왼쪽 귀	오른쪽 귀
4자5 9라7	3가1 2바6
6자5 7가4	9나1 3마2
6자8 3마0	9사7 4가2
5다4 2마8	3사6 9나7
4사2 9바8	5가7 1마6
5자6 4나8	9바2 3사7
7바4 5가6	8다2 9마0
1다6 2마4	5자3 7바9
1라6 3가7	2나4 8사5
9다2 4마3	8바5 7마6

Fig. 2. Whispered voice test list made by randomization table was used in this study. Each lane consists of six different number list or number-word list.

통계 분석

SAS(ver 8.2)와 MedCalc(ver 8.1)을 이용하여 one-way ANOVA, Student t-test, Descriptive statistics, Cohen's kappa statistics, Sperman's 상관 분석을 시행하였으며, 유의수준 0.05 이하를 의미있게 판단하였다.

결 과

검사 환경과 검사자의 특성

각 검사가 이루어진 곳인 4개 센터 병원의 청각 검사실의 평균 소음 수준은 49.7 dBSPL이었으며, 그 외 귓속말 검사가 이루어질 장소인 외래 등의 공간에서 측정된 소음 수준은 평균 50.1 dBSPL로 두 공간의 평균 소음 수준은 통계적으로 유의한 차이를 보이고 있었다. 귓속말 검사를 하게 될 검사자들의 평균 귓속말 크기를 각 검사자의 평균 귓속말 소리 크기에서 각 검사 장소에서의 평균 환경 소음 수준을 뺀 값으로 구한 결과 평균 귓속말 크기는 1.2 dB였으며, 검사자간 귓속말 크기는 통계적으로 유의한 차이를 보이고 있었다($p < 0.01$, Table 1).

연구 대상자의 임상적 특성

연구 참여에 서면 동의한 연구 대상자는 총 172명으로 남자 76명, 여자 96명이었으며, 평균 연령은 67.8세였다. 대상자들이 이비인후과 외래를 찾은 주된 호소 증상은 난청이 가장 많아 87명(50.6%)였으며, 이명이 55명(31.9%)에서 관찰되었고, 그 외 어지럼과 이루 등을 호소하였다(Table 2). 2차 검사까지 완료된 환자는 144명이었고 이들의 1, 2차 검사결과를 활용해 검사-재검사 신뢰도 측정이 가능하였다.

귓속말 검사의 난청 선별 검사로서의 유용성

순음청력검사 회화음역에서의 평균 기도청력역치를 기준으로 1) 26 dB 이상의 청력소실이 있을 경우를 청력 기준으로 하였을 때와 2) 40 dB 이상의 청력소실이 있을 경우를 청력 기준으로 한 두 가지 경우에 대해 정상 청력귀와 난청이 있는 귀 사이의 귓속말 검사시 정확히 따라한 숫자 혹은 글자의 평균 수를 비교해 보았다.

그 결과 청력역치 26 dB 기준시 정상 청력을 보인 귀(95귀)에서 귓속말 검사시 숫자 조합의 평균 정답수는 3.9개, 난청을 보인 귀(193귀)에서 평균 정답수는 1.3개로 정상 청력귀와 난청 귀 사이에 정답을 맞춘 숫자는 통계적으로 유의한 차이를 보이고 있었다($p < 0.01$). 청력 역치 40 dB를 기준으로 살펴본 결과에서도 정상 청력귀(172귀)에서

Table 1. Mean sound levels of each environment and whispered voice test

	N	Sound levels (Mean±SD)	Statistics
Environment			
Audiologic test room ^a	4 (28)	49.7±4.6	
Others (OPD) ^b	4 (28)	50.1±0.6	^{a,b} $p < 0.0001$, t-test
WVT	8 (56)	1.2±1.0	$p < 0.0001$ 1 way ANOVA

WVT : whispered voice test

Table 2. Clinical characteristics of subjects (N=172)

Age (Mean±SD)	67.8±6.4 years
Sex M : F	76 : 96
Chief complaints	N (%)
Hearing impairment	87 (50.6)
Tinnitus	55 (31.9)
Dizziness	23 (13.4)
Otorrhea	6 (3.5)
Others	1 (0.6)

귓속말 검사시 맞춘 정답수가 숫자 조합에서 평균 3.2개, 난청귀(116귀)에서는 평균 0.6개, 숫자-글자 조합에서는 각각 평균 2.8개, 0.5개로 통계적으로 유의한 차이를 보였다($p < 0.01$, Table 3).

순음청력검사상 기도청력 26 dBHL과 40 dBHL 두 기준에 대해 귓속말 검사의 숫자 조합 리스트에서 정확히 따라한 각각의 정답 수에 따른 민감도, 특이도, 유사비(likelihood ratio)를 살펴본 결과 26 dBHL 기준치에서는 정답 수 3개에서 민감도, 특이도가 각각 86.0%, 58.9%로 난청 선별의 적절한 기준점(cut-off point)이라 판단되었고, 40 dBHL 기준치에서는 정답 수 1개에서 민감도 84.5%, 특이도 74.4%로 가장 적합한 기준점으로 판단되었다(Table 4).

순음청력검사상 기도청력 26 dBHL과 40 dBHL 두 기준에 대해 귓속말 검사의 숫자-글자 조합 리스트에서 정확히 따라한 각각의 정답 수에 따른 민감도, 특이도, 유사비를 살펴본 결과에서, 26 dBHL 기준치에서는 정답수 2개에서 민감도, 특이도가 각각 77.7%, 70.5%로 난청 선별의 적절한 기준점이라 판단되었고, 40 dBHL 기준치에서는 정답수 1개가 민감도 84.5%, 특이도 74.4%로 적합한 기준점으로 판단되었다(Table 5).

귓속말 검사의 선별 검사로서의 신뢰도 검증

검사자간 신뢰도

귓속말 검사의 신뢰도 검증을 위해 시행한 동일 환자 대상자로 두 명의 검사자가 시행한 귓속말 검사의 검사자간

Table 3. Comparison of correctly repeated test scores during whispered voice test between normal and impaired hearing subjects

	Normal hearing Mean \pm SD	Impaired hearing Mean \pm SD	p value
Gold standard 26 dBHL			
Number list	3.9 \pm 2.0 (n=95)	1.3 \pm 1.8 (n=193)	<.0001
Number/word list	3.4 \pm 1.9 (n=95)	1.2 \pm 1.7 (n=193)	<.0001
Gold standard 40 dBHL			
Number list	3.2 \pm 2.1 (n=172)	0.6 \pm 1.3 (n=116)	<.0001
Number/word list	2.8 \pm 1.9 (n=172)	0.5 \pm 1.2 (n=116)	<.0001

Table 4. Validity of whispered voice test with number list at each cut-off point

Criterion	Sens. (95% C.I.)	Spec. (95% C.I.)	+LR	-LR
Gold standard 26 dBHL				
<0	0.0 (0.0- 1.9)	100.0 (96.2-100.0)		1.00
≤ 1	64.2 (57.0- 71.0)	81.1 (71.7- 88.4)	3.39	0.44
≤ 2	77.7 (71.2- 83.4)	70.5 (60.3- 79.4)	2.64	0.32
$\leq 3^*$	86.0 (80.3- 90.6)	58.9 (48.4- 68.9)	2.10	0.24
≤ 4	92.2 (87.5- 95.6)	38.9 (29.1- 49.5)	1.51	0.20
≤ 5	99.0 (96.3- 99.8)	14.7 (8.3- 23.5)	1.16	0.07
≤ 6	100.0 (98.1-100.0)	0.0 (0.0- 3.8)	1.00	
Gold standard 40 dBHL				
<0	0.0 (0.0- 3.2)	100.0 (97.9-100.0)		1.00
$\leq 1^*$	84.5 (76.6- 90.5)	74.4 (67.2- 80.8)	3.30	0.21
≤ 2	91.4 (84.7- 95.8)	58.1 (50.4- 65.6)	2.18	0.15
≤ 3	94.8 (89.1- 98.1)	44.8 (37.2- 52.5)	1.72	0.12
≤ 4	98.3 (93.9- 99.7)	29.1 (22.4- 36.5)	1.39	0.06
≤ 5	100.0 (96.8-100.0)	9.3 (5.4- 14.7)	1.10	0.00
≤ 6	100.0 (96.8-100.0)	0.0 (0.0- 2.1)	1.00	

*recommended cut-off point for hearing screening. Sens. : sensitivity, Spec. : specificity, +LR : positive likelihood ratio, -LR : negative likelihood ratio

신뢰도 검증에서 숫자 조합 리스트의 신뢰 계수가 0.57, 숫자-글자 조합 리스트의 신뢰 계수는 0.51로 숫자 조합의 신뢰도가 더 높게 관찰되었다(Table 6).

검사-재검사 신뢰도 검증

1~2주 간격으로 두 차례 시행한 귓속말 검사의 검사-재검사 신뢰도를 검증한 결과 검사자 및 검사 환경에 관계 없이 두 번의 검사간 0.8 이상의 높은 상관 계수를 보이고 있었다. 비교적 조용한 환경인 청력 검사실에서 청각사에 의해 이루어진 귓속말 검사 결과가 소음 수준이 높은 외래 환경에서 이루어진 검사에 비해 검사-재검사 신뢰도가 좀 더 높음을 알 수 있었다(Table 7).

고 찰

인구의 노령화로 인해 노인 연령층이 전체 인구의 7% 이상을 차지하는 고령화 사회에서 노인 건강에 대한 관심이 높아지고 있고, 삶의 질을 고려한 건강관리 프로그램에 대

한 관심과 질환의 조기 발견을 위한 건강검진 프로그램이 국민건강증진을 목적으로 전 국민을 대상으로 시행되고 있다. 특히 2007년 보건복지부에서 생애전환기 건강검진 프로그램의 일환으로 66세 전국민에게 시행하게 될 건강검진 프로그램에서 높은 유병률을 가진 노인성 난청의 선별 검사가 포함되게 되었다. 난청의 조기 발견과 적절한 재활을 통한 치료는 노인 연령층의 보다 나은 삶의 질을 추구하는 수준 높은 국민건강관리의 일환이라 볼 수 있다. 생애전환기 국민건강진단 66세 청력 검사 항목으로 보건복지부에서 이미 채택한 귓속말 검사의 적절한 검사 방법 개발과 유용성 및 적절성에 대한 평가를 위해 대한이비인후과학회의 지원으로 대한청각학회 산하 노인성난청선별검사 준비위원회가 구성되었고 귓속말 검사의 유용성 검증을 위한 본 연구가 시행되었다.

노인성난청선별검사 준비위원회는 4개 의과대학 3차 의료기관인 대학 병원에서 다기관 연구를 계획하였고, 66세 전국민을 대상으로 검진을 시행할 경우 순음청력기를 소지하고 있지 않는 검진 기관에서 시행할 수 있는 적절한 노인

Table 5. Validity of whispered voice test with number-word list at each cut-off point

Criterion	Sens. (95% C.I.)	Spec. (95% C.I.)	+LR	-LR
Gold standard 26 dBHL				
<0	0.0 (0.0- 1.9)	100.0 (96.2-100.0)		1.00
≤1	64.2 (57.0- 71.0)	81.1 (71.7- 88.4)	3.39	0.44
≤2*	77.7 (71.2- 83.4)	70.5 (60.3- 79.4)	2.64	0.32
≤3	86.0 (80.3- 90.6)	58.9 (48.4- 68.9)	2.10	0.24
≤4	92.2 (87.5- 95.6)	38.9 (29.1- 49.5)	1.51	0.20
≤5	99.0 (96.3- 99.8)	14.7 (8.3- 23.5)	1.16	0.07
≤6	100.0 (98.1-100.0)	0.0 (0.0- 3.8)	1.00	
Gold standard 40 dBHL				
<0	0.0 (0.0- 3.2)	100.0 (97.9-100.0)		1.00
≤1*	84.5 (76.6- 90.5)	74.4 (67.2- 80.8)	3.30	0.21
≤2	91.4 (84.7- 95.8)	58.1 (50.4- 65.6)	2.18	0.15
≤3	94.8 (89.1- 98.1)	44.8 (37.2- 52.5)	1.72	0.12
≤4	98.3 (93.9- 99.7)	29.1 (22.4- 36.5)	1.39	0.06
≤5	100.0 (96.8-100.0)	9.3 (5.4- 14.7)	1.10	0.00
≤6	100.0 (96.8-100.0)	0.0 (0.0- 2.1)	1.00	

*recommended cut-off point for hearing screening. Sens. : sensitivity, Spec. : specificity, +LR : positive likelihood ratio, -LR : negative likelihood ratio

Table 6. Inter-observer reliability test of whispered voice test

	Cohen's kappa	95% CI
Number list	0.57	0.51, 0.64
Number/word list	0.51	0.45, 0.57

CI : confidence interval

Table 7. Test-retest reliability

Tester & Environment	Correlation coefficient (r)
Doctor/nurse at OPD	
Number list	0.88
Number/word list	0.84
Audiologist at audiologic test room	
Number list	0.90
Number/word list	0.90

OPD : out-patient department

성난청선별검사법으로 귓속말 검사가 유용한지 확인하기 위한 연구를 계획하였다. 구미 각국에서 난청 선별 검사로 활용됨이 보고된 바 있는 귓속말 검사의 구체적인 시행 방법과 검사 도구가 국내에서는 보고된 바 없어¹⁵⁻¹⁷⁾ 한국말로 검사를 시행할 경우의 유용성에 대한 검증이 없는 실정이다. 저자들은 본 연구를 통해 귓속말 검사법의 적절한 도구를 개발하고 3차 의료기관 이비인후과 청력실의 표준화된 순음청력도결과를 기준으로 연구대상자들의 난청 여부를 진단하여 보다 객관적이고 정확하게 연구결과를 분석하여 귓속말 검사의 유용성을 확인해보고자 하였다. 또한 선별 검사가 지닐 수 있는 장소나 검사자의 다양성을 고려하여 다기관 연구를 통해 다양한 장소와 검사자들을 포함시킴으로써 귓속말 검사의 난청 선별 검사로서의 유용성을

검증해 보고자 하였다.

우선 본 연구에서는 선별 검사로 시행하게 될 귓속말 검사의 자세한 검사법이 수록된 프로토콜과 검사시 검사자가 불러주게 검사 도구(test material)는 선행 연구에서 제안된 바 있는 6개의 숫자-알파벳-숫자 두 묶음¹⁵⁻¹⁷⁾ 대신 검사 대상자들이 이해하기 쉽게 6개의 숫자 조합 리스트와 숫자-글자-숫자 조합 두 묶음에서, 글자는 자음 대신 '가나 다라마바사아자차'로 구성된 글자를 활용하여 연구하기로 하였고, 난수표를 이용하여 무작위 추출한 숫자와 글자의 조합 혹은 숫자만의 조합으로 구성된 검사 도구(test material) 리스트 두 가지를 제안하게 되었다. 귓속말 검사와 관련한 선행 연구는 모두 국외에서 이루어졌기에 영문으로 된 숫자나 알파벳 조합을 그대로 인용할 경우 검사 대상자들이 정확히 따라하기 어려울 수 있고, 글자 이용시 한글의 자음 혹은 모음 하나를 읽을 경우 역시 검사 대상자들이 따라하기 어려울 수 있다는 것이 본 연구에서 제안하게 된 검사 리스트를 만든 근거라 할 수 있다. 한글의 특성을 감안하여 보편성이 있는 글자로 '가'에서 '차'까지만을 선택하여 숫자-글자 조합 리스트를 작성하게 되었고, 숫자만을 불러줄 경우 검사 대상자들이 더욱 따라하기 쉬울 것이라는 가정하에 선행 연구와는 달리 6개의 숫자만으로 된 숫자 조합 리스트를 따로 구성하여 이 두 종류 리스트 사이에 검사 도구로서의 적절성을 평가함으로써 난청 선별 검사로 귓속말 검사를 활용할 경우 더 나은 검사 도구를 제안하고자 하였다. 본 연구에서는 검사 리스트 제작시 숫자 조합과 숫자-글자-숫자 조합 리스트의 음문학적 연구를 시행하지 못하였

는데 향후 본 연구에서 제안된 검사 리스트의 국어 음문학적 연구를 통한 특성 분석이 추가된다면 보다 과학적이고 객관적인 우리말 귓속말 검사 리스트 제시가 가능할 것이다.

귓속말 검사가 이루어질 검진 기관의 검사 환경의 소음과 검사자의 특성이 다양할 수 있음을 감안하여 다기관 연구를 시행하게 되었다.

우선 연구에 참여한 4개 병원에서 의사/간호사 혹은 청각사로 검사자를 다양하게 구성하여 검사가 이루어질 외래 혹은 청력 검사실의 환경 소음과 귓속말 검사 매뉴얼을 나누어 준 후 귓속말 검사를 시행할 때 내게 될 소리의 크기를 측정해 보았다. 검사 환경은 청력 검사실이 일반 외래에 비해 환경 소음이 의미있게 낮았고, 검사자간 귓속말 검사 시 내는 소리의 크기도 평균 1.2 ± 1.0 dB로 통계적으로는 서로 다른 양상을 보이고 있었으나 그 차이가 3 dB 미만으로 귓속말을 정확히 낼 경우 검사자간 오차는 크지 않아 귓속말 검사의 선별 검사로서의 결과에 큰 차이를 보이지 않을 것임을 예측할 수 있었다.

본 연구에서 대상이 된 3차 병원 이비인후과를 찾는 60세 이상 노인 환자 중 귀 증상으로 내원하였던 환자들의 가장 흔한 주증상은 난청과 이명임을 알 수 있었다. 검사에 동의한 60세 이상 연령의 대상자 중 순음청력검사를 통해 26 dBHL 혹은 40 dBHL의 평균 회화음역기도청력역치를 보이는 두 기준으로 정상청력군과 난청군을 나누었을 때 두 경우 모두 두 구간 충분한 개체수를 보여 통계적 분석시 오류를 배제할 수 있었다.

귓속말 검사의 난청 선별 검사로서의 유용성을 검증하기 위해 일반적인 경도 난청의 시작 기준인 회화음역기도청력역치 26 dB와 사회생활에 필요한 최소청력역치인 40 dB 기준 두 경우로 나누어 정상청력군과 난청군을 나눈 후 각각의 경우에서 귓속말 검사시 정확히 따라한 정답수를 비교한 결과 두 구간 귓속말 검사시 정답수는 통계적으로 유의한 차이를 보이고 있었다. 각 난청 기준에서 귓속말 검사의 도구에 따라 숫자 조합일 경우와 숫자-글자 조합일 경우 높은 민감도와 특이도를 보이는 기준점은 각각 다른 양상을 보였는데, 일반적으로 난청의 시작 기준인 26 dB 기준에서 숫자 조합 리스트는 6개 중 3개 이상을 맞춘 경우 통과 기준으로 삼을 수 있음을 관찰할 수 있었고, 이는 선행 연구결과와 비교적 유사한 결과라 할 수 있었다. 그러나 숫자-글자 조합의 경우는 6개 중 2개 이상을 맞춘 경우가 난청 선별 검사에서 통과 기준점으로 보여 저 숫자 조합에서 3개를 기준으로 할 수 있었던 경우에 비해 낮은 값을 알 수 있었는데 이는 본 연구에서 사용한 우리말의 숫자-글자 조합의 경우 숫자 조합에 비해 정확

히 따라하기가 더 어려운 것이 이유가 될 것으로 판단되었다. 선행 연구에서는 귓속말 검사시 6개 숫자-알파벳 조합 중 3개 이상을 맞출 경우 통과 기준으로 삼았기에¹⁴⁻¹⁷ 검사의 편의성과 선행 연구와의 동질성 등을 감안할 때 본 연구 결과를 통해 제안할 수 있는 선별 검사로서의 귓속말 검사의 통과 기준은 난청 기준 26 dB일 때 귓속말 검사 숫자 조합 리스트를 활용한 검사에서 6개 중 3개 이상을 정확히 따라하는 경우로 지정할 수 있을 것이고 그렇지 못할 경우 난청 여부에 대해 순음청력검사를 포함한 보다 정밀한 검사가 필요하므로 '의뢰'하는 것이 바람직할 것으로 제안하는 바이다.

각 검사 리스트에 대해 귓속말 검사의 난청 선별 검사로서의 신뢰도 검증을 위한 검사자간 신뢰도 검증에서는 동일 환자를 대상으로 두 명의 검사자가 같은 날 검사를 시행하여 그 결과를 비교하였는데 Cohen's kappa 값을 이용한 신뢰도는 숫자 조합에서 0.57, 숫자-글자 조합에서 0.5로 숫자 조합에서 더 높아 숫자 조합을 활용하는 것이 더 바람직할 것으로 판단되었다. 신뢰도 수치는 선행 연구에서도 0.16~1.0까지 다양하게 보고되어 검사자의 귓속말 검사시 내는 소리의 크기 차이가 검사자간 신뢰도를 결정짓는 요소가 될 수 있다고 제안된 바 있는데,¹²⁾ 본 연구에서는 검사 매뉴얼을 나누어 주고 검사자들에게 이를 숙지시킨 후 귓속말 검사가 이루어졌음에도 신뢰도가 크게 높지 않았기에 본 검사법의 보다 정확한 실행을 위해서는 시청각 시스템을 활용한 검사법과 같은 상세한 교육 프로그램의 개발 등이 추가로 필요할 것으로 판단되었다.

반면 검사 시기를 달리하여 동일 환자에서 두 차례 검사가 이루어졌을 때 검사-재검사 신뢰도를 측정한 경우에는 숫자 조합 및 숫자-글자 조합 모두 두 차례의 검사간 상관계수가 0.8 이상으로 높게 관찰되었고, 특히 비교적 조용한 환경인 청력 검사실에서 귓속말 검사가 이루어진 경우 상관계수가 0.9로 매우 높아 검사의 재현성은 높음을 알 수 있었다.

본 연구에서 귓속말 검사시 난청 환자와 정상 청력인 사이에서 그 검사 결과가 뚜렷하게 의미있는 차이를 보여 적절한 기준점을 활용할 경우 난청 환자 선별에 활용할 수는 있을 것이다. 그러나 검사자의 검사법과 환경 소음 수준에 영향을 받을 수 있는 검사임을 본 연구를 통해 다시 한번 확인하였고 따라서 검사자간 신뢰도가 매우 높은 검사라 할 수는 없었기에 저자들은 귓속말 검사만을 채택하여 노인성 난청 선별검사를 건국 검진기관에서 시행하는 것보다는 순음청력검사가 불가능한 건강검진기관에서 본 검사법이 차신책으로 활용되는 것이 보다 바람직할 것으로 제안하는 바이다. 또한 귓속말 검사가 선별 검사로서 활용되기 위해서는

정확한 검사법에 대한 교육법이 개발되고 모니터링되어야 할 것으로 생각되며 향후 다양한 검사 환경 소음 수준에 따른 선별 검사로서의 정확성 분석이 추가되어 그 결과에 따른 검사 환경의 소음 수준 관리가 필요할 것으로 판단된다.

결론

생애전환기 국민건강진단 66세 청력 검사로 채택된 귓속말 검사의 검사법 개발과 난청 선별 검사로서의 유용성 및 신뢰도 검증을 위해 본 연구를 계획하였다. 검사시 활용할 수 있도록 개발된 숫자 조합 리스트와 숫자-글자 조합 리스트 중 유용성과 신뢰도 모두에서 숫자 조합 리스트가 숫자-글자 조합 리스트에 비해 높은 값을 보여 귓속말 검사를 활용한 선별 검사시에는 숫자 조합 리스트를 활용하는 것이 보다 유용할 것으로 판단되었다.

그러나 검사 장소에 따라 환경 소음의 차이가 클 수 있고 특이도와 검사자간 신뢰도가 높지 못한 점, 또 검사 장소에 따라 검사-재검사 신뢰도가 영향을 받을 수 있는 점 등과 같은 문제점을 본 연구를 통해 발견할 수 있었기에 귓속말 검사를 통한 난청 선별은 순음청력검사가 불가능한 검진기관에서 차선택으로 활용하는 것이 바람직할 것이며, 이 경우에도 검사법의 철저한 교육과 검사 환경의 관리가 필요할 것으로 판단된다.

중심 단어 : 청력 선별 검사 · 고연령층 · 유용성 · 신뢰도.

■ 감사문

본 연구를 위해 데이터 수집에 많은 수고를 해주신 각 병원의 청각사, 연구 간호사, 전임의 및 전공의 선생님들과 통계 분석에 많은 도움을 주신 가톨릭대학교 예방의학교실의 연구원들, 연구 과정에 대해 관심과 지원을 보내 주신 대한이비인후과학회, 대한청각학회와 보건복지부에 감사를 드립니다.

본 연구는 2007년 6월 제44차 대한청각학회 학술대회에서 사업 보고 형태로 구연 발표되었음.

REFERENCES

- 1) Abutan BB, Hoes AW, Van Dalsen CL, Vershuure J, Prins A. *Prevalence of hearing impairment and hearing complaints in older adult: A study in general practice. Fam Pract* 1993;10 (4):391-5.
- 2) Sindhusake D, Mitchell P, Smith W, Colding M, Newall P, Hartley D, et al. *Validation of self-reported hearing loss. The blue mountains hearing study. Int J Epidemiol* 2001;30 (6):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 (9):879-86.
- 4) Bess FH, Lichtenstein MJ, Logan SA, Berger MC, Nelson E. *Hearing impairment as a determinant of function in the elderly. J Am Geriatr Soc* 1989;37 (2):123-8.
- 5) Strawbridge WJ, Wallhagen MI, Shema SJ, Kaplan GA. *Negative consequences of hearing impairment in old age: A longitudinal analysis. Gerontologist* 2000;40 (3):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 (3):188-94.
- 7) Nondahl DM, Cruickshanks KJ, Wiley TL, Tweed TS, Klein R, Klein BE. *Accuracy of self-reported hearing loss. Audiology* 1998;37 (5):295-301.
- 8) Burkey JM, Lippy WH, Schuring AG, Rizer FM. *Clinical utility of the 512 Hz Rinne tuning fork test. Am J Otol* 1998;19 (1):59-62.
- 9) Mulrow C. *Screening for hearing impairment in the elderly. Hosp Pract* 1991;26 (2A):79-86.
- 10) Kim YS, Won CW, Kim BS, Choi HR, Kim SH, Kim JC, et al. *Usability of HHIE-S as a screening test of hearing impairment in the Korean elderly. J Korean Acad Fam Med* 2001;22 (6):878-85.
- 11) Song SW, Kang SY, Yeum KS, Choi WS, Shin HC, Park ES. *Smith Hearing Screening Test for Identifying Hearing-Impaired Patients. J Korean Acad Fam Med* 1994;15:465-75.
- 12) Uhlmann RF, 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 (2):90-6.
- 13) Mulrow CD, Lichtenstein MJ. *Screening for hearing impairment in the elderly: Rationale and strategy. J Gen Intern Med* 1991;6 (3):249-58.
- 14) Pirozzo S, Papinczak T, Glasziou P. *Whispered voice test for screening for hearing impairment in adults and children: Systematic review. BMJ* 2003;327 (7421):967.
- 15) Eckhof 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 (409):473-4.
- 16) Swan IR, Browning GG. *The whispered voice as a screening test for hearing impairment. J R Coll Gen Pract* 1985;35 (273):197.
- 17) Macphree GJ, Crowther JA, McAlpine CH. *A simple screening test for hearing impairment in elderly patients. Age Ageing* 1988;17 (5):347-51.