Deriving Criteria for Hearing Impairment in the Elderly: A Functional Approach

Michael J. Lichtenstein*
Fred H. Bess†
Susan A. Logan†
M. Candice Burger‡

Abstract

We describe a method for deriving criteria for hearing impairment in the elderly based on self-reported handicap. Using the Sickness Impact Profile (SIP) and Hearing Handicap Inventory for the Elderly - Screening (HHIE-S) version as functional measures of handicap, the analysis proceeded in five steps:

1. Audiometric thresholds at various frequencies were inter-correlated. This was done both within and between ears.
2. Better and poorer ear thresholds were determined for each frequency, and these were correlated with the HHIE-S and SIP scores.
3. Using the HHIE-S and SIP scores as dependent variables, stepwise multiple linear regressions were used to select the frequencies that explained the most variance in the functional scales.
4. Using the HHIE-S and SIP as standards, receiver operating curves were constructed for each frequency to select the threshold level that provided the best test accuracy.
5. The newly-derived criteria were then compared against four other "traditional" criteria of hearing impairment.

In general, the newly-derived criteria combined a relatively low frequency with a relatively high frequency, with the low frequency being functionally more important. Depending on the functional scale used, the threshold level was in the 25 to 35-dB range for the lower frequencies and 40 to 45-dB for the higher frequencies. These features provide a suitable compromise to the current debate over which threshold levels comprise the best discrimination of aged persons who are hearing-impaired. Future research should focus on developing consensus standards for functional hearing impairment and handicap in the elderly.

Key Words: Hearing disorders, aged

A commonly accepted criterion for defining hearing impairment is an average pure tone level of 25 dB HL or more in the better ear at frequencies 500, 1000, and 2000 Hz (Clark, 1981; ASHA, 1981; Goldstein, 1984). The assumption underlying this criterion is that a hearing level in excess of 25 dB HL can impose a handicapping condition; that is, the hearing loss can negatively affect a person's ability to perform routine activities of daily life (Solomon, 1986). The wisdom of using a single intensity level to describe a hearing handicap has long been questioned, especially for the elderly population (Weinstein, 1984; Solomon, 1986; Bess, 1985). For example, it has been argued that the classical pure tone criterion is an inadequate fence, since speech understanding is heavily dependent on hearing above 2000 Hz,

*Division of Geriatrics and Gerontology, Geriatrics Research Education and Clinical Center, Audie L. Murphy Memorial Veterans Hospital, San Antonio, Texas

†Division of Hearing and Speech Sciences, Vanderbilt University School of Medicine and Bill Wilkerson Hearing and Speech Center, Nashville, Tennessee

‡Division of General Internal Medicine, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee

Reprint requests: Fred H. Bess, Division of Hearing and Speech Sciences, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

especially when listening under noise conditions. Others have suggested that any criterion based on pure tone data alone is inappropriate since these measures do not correlate well with perceived impact on daily function (Weinstein, 1984; Rosen, 1978; Ventry and Weinstein, 1983). There are large individual differences in how the elderly with similar degrees and configurations of hearing loss respond to a hearing impairment. Because of the imperfect relationship between audiometric data and perceived handicap, many clinicians have begun supplementing the audiologic evaluation with self-assessment scales (Weinstein, 1984; Bess et al, 1989a; Noble, 1978; Ward et al, 1977; Solomon, 1986).

The need for determining the most appropriate criterion for classifying a hearing impairment has become more evident with the increased interest in the hearing-impaired elderly. Ventry and Weinstein (1983) recommended a 40 dB fence at 1000 and 2000 Hz as the pass-fail criterion for the elderly population. This criterion appears to have achieved general acceptance among audiologists and physicians. Several investigators, however, have suggested that a lower cut-off level would identify more individuals with milder forms of hearing loss who are in need of intervention. To help resolve this controversy, Bess and co-workers (1989b) examined the association between four different commonly accepted hearing impairment criteria and functional health status. The outcome measures included the Sickness Impact Profile (SIP, a measure of global function) and the Hearing Handicap Inventory for the Elderly-Screening version (HHIE–S, a communications-specific measure of functional impairment). Considerable overlap among the various criteria was found. In addition, it was noted that the criterion of hearing impairment chosen depends on the functional standard used. When a measure of global function was utilized, the speech frequency average (0.5 kHz, 1 kHz, 2 kHz-^{0} 25 dB HL) was the most discriminating criterion. If the HHIE–S was used the criterion of Ventry and Weinstein (1983) was considered the most appropriate for determining presence of handicap. A next step in this research is to determine whether a criterion of hearing impairment in the elderly can be derived that best describes the extent of the functional handicapping condition. This report represents a continuation of the work of Bess and co-workers (1989b) and seeks to develop specific criteria of hearing impairment for the elderly population using the two functional assessment scales, the SIP and the HHIE–S, as external standards.

METHOD

Recruitment of Subjects

Seven primary-care internists working in six practices agreed to participate. Each practice agreed to screen a consecutive sample of 50 patients over 65 years of age for hearing impairment with the Hearing Handicap Inventory for the Elderly–Screening version (HHIE–S; Ventry and Weinstein, 1982) and an audioscope (Lichtenstein et al, 1988). Regardless of the outcome of the screens, subjects were then referred to the Bill Wilkerson Hearing and Speech Center for an audiologic assessment.

Persons were not to be screened if they were bedfast or chairfast, acutely ill, had an active sinus or upper respiratory infection, or had documented strokes with a residual paresis or aphasia.

Testing at the Hearing Center

Subjects were not aware of their audiologic status prior to completion of the functional assessment forms. Moreover, the audiologists had no knowledge of the subjects' functional status at the time of the audiologic assessments.

Audiometry

Audiologic assessments were obtained with one of several diagnostic audiometers (GSI-16[3], Madsen OB822), which were calibrated periodically in accordance with ANSI (1972) specifications. Biologic checks were performed weekly. All patients were evaluated in custom sound-treated rooms that had an acoustic environment suitable for threshold measurement (ANSI, 1977). Pure tone thresholds were established using the modified Hughson-Westlake approach (Carhart and Jerger, 1959). Air conduction thresholds were obtained at octave intervals ranging from 250 to 8000 Hz. The speech recognition threshold (SRT) was obtained via monitored live voice using spondee materials. The SRT was measured by the Tillman and Olsen (1973) method.

Global Functional Scale–The Sickness Impact Profile (SIP)

The SIP is a 136-statement standardized questionnaire that assesses function in a behavioral context (Gilson et al, 1975; Bergner et al, 1976; Bergner et al, 1981). The statements are weighted and grouped into 12 subscales: ambulation, mobility, body care/movement, social interaction, communication, alertness, emotional, sleep/rest, eating, work, home management, and recreation/pastimes. In addition there are three main scales: physical (combining ambulation, mobility, body care/movement), psychosocial (combining social interaction, communication, alertness, emotional), and overall (combining all 12 subscales). The higher the SIP score, the greater the perceived functional impairment. The SIP is a valid, reliable tool that has been applied in a number of areas to measure sickness-related dysfunction (Carter et al, 1976; Pollard et al, 1976). As examples, the SIP has been used in studies of end-stage renal disease (Hart and Evans, 1987), survivors of myocardial infarction (Bergner et al, 1985), chronic obstructive pulmonary disease (McSweeney et al, 1982), and rheumatoid arthritis (Deyo and Inui, 1984). In addition, SIP scores have been associated with level of hearing impairment; as hearing worsens in the elderly, SIP scores increase (Bess et al, 1989). In the present study, the SIP was self-administered. The few subjects who could not complete the SIP by themselves had the questions read to them.

Hearing Impairment Functional Scale–Hearing Handicap Inventory for the Elderly–Screening Version (HHIE–S)

The HHIE–S is a self-administered 10-item questionnaire designed to detect emotional and social problems associated with impaired hearing (Ventry and Weinstein, 1982; 1983). Subjects respond to questions about circumstances related to hearing by stating whether the situation presents a problem. A "no" response scores 0, "sometimes" scores 2, and "yes" scores 4. Total HHIE–S scores range from 0 to 40. The HHIE–S results from the Hearing Center were compared against the audiometric results obtained on the same visit.

Data Analysis

We sought to derive criteria for hearing impairment that best classified individuals against functional scales of handicap. The data analysis proceeded in five discrete steps:

1. Pearson product-moment correlations of threshold levels between frequencies within ears and between ears were determined (Armitage, 1971). Since the threshold levels at various frequencies are *not* independent, it is important to know the correlations; for example, if one knows the threshold at 0.5 kHz in the left ear, one can make a reasonable estimate of what the threshold is at 1 kHz in the same ear or 0.5 kHz in the opposite ear.
2. Better and poorer individual ear thresholds were determined for each frequency, and these were correlated with the functional scales.
3. Using the functional scale scores as dependent variables and the frequencies as independent variables, stepwise multiple linear regressions were conducted to select the frequencies that explained the most variance in the functional scales (Armitage, 1971).
4. Using the functional scales as standards, receiver operating curves were constructed for each frequency to select the threshold level that provided the best overall accuracy for categorizing persons as impaired or unimpaired. (Sackett et al, 1985).
5. The newly-derived criteria were then compared against four other criteria of hearing impairment. Again, using stepwise multiple linear regression, the functional scales were continuous dependent variables. Subjects were categorized by each criterion of hearing impairment, and the regression was utilized to select the criterion that explained the most variance in the functional scale. The four criteria of hearing impairment chosen to compare against the newly derived criteria were:
   a) The criterion of Ventry and Weinstein (Ventry and Weinstein, 1982; Ventry and Weinstein, 1983). A subject considered hearing impaired if (a) there was a 40-dB loss at either the 1 *or* 2 kHz frequencies in both ears or (b) there was a 40-dB loss at 1 *and* 2 kHz in one ear.
   b) Speech Frequency Pure Tone average

(SFPTA) (Clark, 1981; Carhart, 1971). A subject was considered hearing impaired if the average hearing loss at 0.5, 1, and 2 kHz was 25 dB or greater in the better ear.

c) High Frequency Pure Tone Average (HFPTA) (Goldstein, 1984; Solomon, 1986). A subject was considered hearing impaired if the average hearing loss at 1, 2, and 4 kHz was 25 dB or greater in the better ear.

d) Speech Recognition Threshold (SRT) (Clark, 1981; Carhart, 1971). A subject was considered hearing impaired if the SRT was 25 dB or greater in the better ear.

These criteria were selected because of traditional usage (b, d), special standardization with the elderly (a), and representation of high frequency hearing status (d).

RESULTS

Recruitment and Characteristics of Study Group

Three hundred and four patients were screened and referred by the six practices. Of these, 178 (59 percent) kept appointments for testing at the Hearing Center. The mean age of responders was 74.2 (SD=6.4) years, 78 percent were Caucasian, and 63 percent were women. The responders had an average of 2.8 medical conditions and were taking 2.6 medications. The persons who kept appointments did not differ substantially from nonrespondents except for having higher HHIE-S scores. Respondents had mean HHIE-S score of 10.0 (SD=10.1) compared to 7.6 (SD=10.3) for nonrespondents (Lichtenstein et al, 1988).

Of the 178 responders, all completed the HHIE-S a second time at the Hearing Center, while 153 (86 percent) completed the SIP. The reasons subjects failed to complete the SIP included fatigue or the perception that questions were either too personal or offensive. Compared to the total sample of respondents, the 25 persons not completing the SIP were older (78 versus 74 years), more likely to be black (32 percent versus 22 percent), and more likely to be female (80 percent versus 63 percent). Persons not completing the SIP did not differ from those who did in terms of number of medical conditions or medications (Bess et al, 1989).

Of the 178 subjects, one person had missing data at 4 kHz. Therefore, when the HHIE-S is the functional scale the number of subjects are 178 or 177 depending on whether 4 kHz information is used. When the SIP is the functional scale the number of subjects are 153 or 152, depending on whether 4 kHz information is used.

Correlation of Frequencies Within and Between Ears

The Pearson product-moment correlations for frequencies *within* ears are displayed in Table 1. Note that within each ear the correlations are typically high, on the order of 0.8 to 0.9 for frequencies that are proximate to each other, e.g., 0.5 and 1 kHz, or 1 kHz and 2 kHz. The correlations between frequencies decrease as the frequency difference widens. Thus low frequencies are not as highly correlated with high frequencies. The implication of these findings is that hearing impairment criteria ought to combine a low frequency with a high frequency since they are more likely to provide independent information.

The Pearson product-moment correlations for frequencies *between* ears are given in Table 2. Note here too that the frequencies are substantially correlated. Comparing the same frequencies between ears the correlations range from 0.70 to 0.78, implying that once a threshold is known in one ear, 50 percent of the variance in the threshold at the same frequency in the opposite ear has been accounted for. This lack of independence between ears is underscored in Table 3, which gives the differences in decibel

Table 1 Within-Ear Pearson Product-Moment Correlations for Threshold Levels at Each Frequency (N=177)*

	250 Hz	*500 Hz*	*1000 Hz*	*2000 Hz*
Right ear				
500 Hz	0.94			
1000 Hz	0.83	0.90		
2000 Hz	0.69	0.74	0.84	
4000 Hz	0.49	0.53	0.61	0.72
Left ear				
500 Hz	0.94			
1000 Hz	0.82	0.89		
2000 Hz	0.64	0.72	0.84	
4000 Hz	0.50	0.53	0.62	0.74

*All correlation coefficients statistically significant at p = <0.001

threshold between ears for each frequency. The mean difference between ears was normally distributed and ranged from 0 dB for 0.5 kHz to 4 dB for 4 kHz. The median and modal dB differences between ears were 0 dB at all frequencies. There were a few subjects with substantial between-ear differences ranging up to 95 dB at 2 kHz. However, in general, the picture of presbyacusis in a screened population was one of bilateral losses of roughly equal magnitude in each ear.

Correlation of Frequencies with Functional Scales

Pearson product-moment correlations for the threshold levels within each frequency and the functional scales are given for each ear in Table 4. Note that although statistically significant, the correlations were modest. They ranged from 0.42 at 4 kHz in the left ear to 0.57 at 1 kHz in the right ear for the HHIE–S, the hearing handicap specific scale. The correlations were lower for the SIP, the global functional scales, ranging from 0.24 to 0.48. Comparing the main SIP subscales, the correlations tended to be higher for the Physical SIP.

Table 2 Between-Ear Pearson Product-Moment Correlations for Threshold Levels at Each Frequency (N=177)*

Right ear	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
Left ear					
250 Hz	0.72	0.71	0.64	0.47	0.37
500 Hz		0.75	0.67	0.54	0.43
1000 Hz			0.70	0.60	0.47
2000 Hz				0.70	0.59
4000 Hz					0.78

*All correlation coefficients statistically significant at p = <0.001

Table 3 Differences in Threshold Level Between Ears by Frequency

	Differences			
Frequency	Mean (SD) dB	Median dB	Mode dB	Maximum dB
250 Hz	2 (14)	0	0	90
500 Hz	0 (13)	0	0	90
1000 Hz	1 (15)	0	0	90
2000 Hz	2 (16)	0	0	95
4000 Hz	4 (15)	0	0	65

Table 4 Pearson Product-Moment Correlation Coefficients for Threshold Levels at Each Frequency in Each Ear and the SIP (N=152) and HHIE-S (N=177)*

	SIP Scales			
	Psychosocial	Physical	Overall	HHIE-S
Right ear				
250 Hz	0.35	0.48	0.43	0.51
500 Hz	0.31	0.46	0.39	0.53
1000 Hz	0.27	0.41	0.34	0.57
2000 Hz	0.28	0.40	0.33	0.49
4000 Hz	0.26	0.30	0.27	0.44
Left ear				
250 Hz	0.26	0.34	0.31	0.48
500 Hz	0.24	0.38	0.32	0.53
1000 Hz	0.25	0.32	0.30	0.53
2000 Hz	0.31	0.33	0.32	0.48
4000 Hz	0.27	0.25	0.26	0.42

*All correlation coefficients statistically significant at p = <0.001

Next, each of the dB thresholds at each frequency was compared for each ear and categorized as the better or poorer ear threshold. These thresholds were then correlated with the functional scales for each frequency; the correlations are presented in Table 5. By sorting thresholds in this way the correlations for the better ear threshold are observed to be generally lower than the correlations for the poorer ear thresholds. We infer from these results that functionally derived criteria for determination of whether a person has a hearing handicap will

Table 5 Pearson Product-Moment Correlation Coefficients for Better and Poorer Ear Threshold Levels at each Frequency and the SIP (N=152) and HHIE-S (N=177)*

	SIP Scales			HHIE-S
	Psychosocial	Physical	Overall	
Poorer ear				
250 Hz	0.36	0.48	0.44	0.50
500 Hz	0.34	0.50	0.44	0.55
1000 Hz	0.33	0.43	0.40	0.56
2000 Hz	0.35	0.44	0.40	0.49
4000 Hz	0.30	0.28	0.30	0.42
Better ear				
250 Hz	0.28	0.39	0.34	0.52
500 Hz	0.24	0.37	0.30	0.54
1000 Hz	0.23	0.35	0.28	0.58
2000 Hz	0.27	0.33	0.29	0.52
4000 Hz	0.26	0.28	0.26	0.46

*All correlation coefficients statistically significant at p = <0.001

be better if poorer ear thresholds are used (independent of the ear in which they occur). Once a subject is considered hearing impaired then the question of which ear to fit with an aid becomes a separate issue. Therefore, in the remainder of the analyses, only poorer ear thresholds are considered.

Selecting Frequencies Based on Functional Scales

Stepwise multiple linear regression was used to select the frequencies that explained the most variance in the functional scales. The functional scales were used as continuous dependent variables and the frequencies as continuous independent variables. The results of the regressions are given in Table 6. In each regression the standardized residuals were normally distributed and did not differ significantly from zero. The table lists the frequencies selected by the regression for each functional scale, the regression coefficient, and the percentage variance attributed to that frequency as it entered the model.

For example, with the HHIE-S as the functional scale, the 0.5 kHz was selected first, accounting for 31 percent of the variance in the scale. The 4 kHz was selected second and accounted for a further 2 percent of the variance.

With the Physical SIP as the functional scale the 0.5 kHz was selected first, accounting for 25 percent of the variance. The 2 kHz was selected second, accounting for 3 percent of the variance. The 1 kHz was selected third by the regression, accounting for 2 percent of the variance. Interestingly, however, after adjustment for the 0.5 and 2 kHz frequencies, the 1 kHz was negatively associated with the Physical SIP score. A 10 dB increase in 1 kHz poorer ear threshold results in a decrease of 2.4 points in Physical SIP score.

With the Pyschosocial SIP the regressions selected the 0.25 kHz and 2 kHz frequencies accounting for 12 percent and 3 percent of the variance respectively. For the Overall SIP the selected frequencies were 0.5 and 2 kHz accounting for 19 percent and 3 percent of the variance respectively.

Table 6 Regression Equations for Selecting Frequencies for Each of Functional Standards

	Regression Equation			
		Regression Coefficient‡		
*Functional Standard**	*Frequency†*	*Beta*	*(S.E.)*	*R^2§*
HHIE-S (N=177)	500	2.9	(0.4)	31%
	4000	1.0	(0.3)	2%
Physical SIP (N=152)	500	3.8	(0.9)	25%
	2000	2.1	(0.7)	3%
	1000	−2.4	(1.2)	2%
Psychosocial SIP (N=152)	250	1.8	(0.7)	12%
	2000	1.5	(0.6)	3%
Overall SIP (N=152)	500	2.2	(0.7)	19%
	2000	1.4	(0.6)	3%

*Functional standards as continuous dependent variables

†Frequencies as continuous independent variables. Each frequency contains the poorer ear threshold for either ear.

‡Change in functional standard score for a 10-dB change in threshold within each frequency.

§Percent variance attributed to each frequency as it entered the model. Total variance of the model attained by adding the separate variances.

Selecting Threshold Levels Within Frequencies

Receiver operating curves were constructed to select the threshold levels within each frequency that gave the best discrimination between those subjects with and those without handicapping functional scores. To do this, cut points for what was considered an abnormal or handicapping level were chosen for each functional scale. For the HHIE-S, a score of 10 or greater was indicative of handicap (N=67). For the SIP scales, the upper third of the score distributions were considered handicapping levels. For the Psychosocial SIP this was a score of 5.5 or greater (N=52), for the Physical SIP a score of 4.4 or greater (N=52), and for the Overall SIP a score of 9.1 or greater (N=52).

Next, sensitivities (the proportion of people with handicap correctly identified by the threshold criterion), specificities (the proportion of persons without handicap correctly identified by the threshold criterion), and test accuracy (the proportion of persons with or without handicap correctly identified by the threshold) were calculated for each threshold criterion within each frequency. These data are presented for each functional scale in the detailed Tables A–D in the Appendix. In these tables the sensitivities, specificities, and accuracies are given at levels greater than or equal to the stated threshold levels. The threshold level providing the best test accuracy was taken as the threshold criterion to be used within the selected frequencies for the derived criteria. If two thresholds provided similar test accuracies, then the lower threshold was chosen (e.g., 72 percent accuracy was calculated at both 30 and 35 dB at 1 kHz for the Overall SIP; the 30-dB level would be selected as the threshold).

The receiver operating curves for the frequencies selected by the regressions for each of the functional standards are plotted in Figures 1 to 4. Each point represents the sensitivity and 1-specificity for a threshold level at a given frequency for the particular functional scale. On each curve, from right to left, the points plotted are for the 15, 20, 25, 30, 35, 40, and 45-dB threshold levels. For each plot, the circled point represents the threshold level that gave the best test accuracy at that frequency for that functional standard.

The results of the threshold levels selected for each frequency are summarized in Table 7. For example when the HHIE-S is the functional scale, a 35-dB threshold criterion at 0.5 kHz provided the best test accuracies. To decide how best to combine threshold levels we looked at two possibilities. In the first, the subject had to meet criteria at all frequencies (e.g., greater than or equal to a 35-dB level at 0.5 kHz *AND* greater than or equal to a 45-dB level at 4 kHz). In the second, the subject had to meet criteria at only one frequency (e.g., greater than or equal to a 35-dB level at 0.5 kHz *OR* greater than or equal to a 45-dB level at 4 kHz). To determine which combination to use we looked at the prevalence of impairment defined by the two methods. These are also listed in Table 7. For the *AND* combination the prevalence of impairment was low, ranging from 10 to 13 percent across the

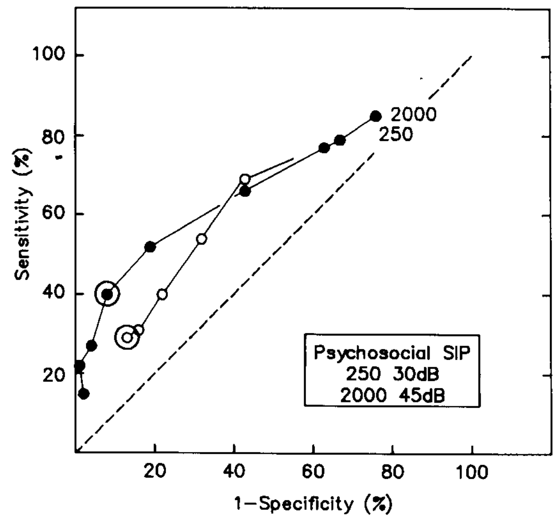

Figure 2 Receiver operating curves for threshold levels within 0.25 and 2 kHz-frequencies with the Psychosocial SIP as the functional standard. Impairment defined as a Psychosocial SIP score of 5.5 or greater.

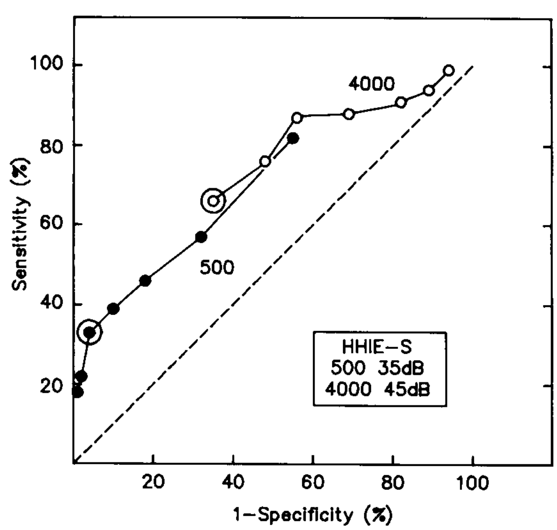

Figure 1 Receiver operating curves for threshold levels within 0.5 and 4-kHz frequencies with the HHIE-S as the functional standard. Impairment defined as HHIE-S score of 10 or greater.

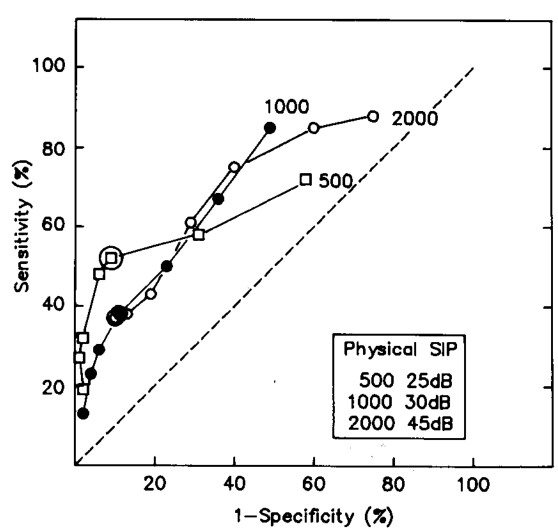

Figure 3 Receiver operating curves for threshold levels within 0.5, 1, and 2 kHz-frequencies with the Physical SIP as the functional standard. Impairment defined as a Physical SIP score of 4.4 or greater.

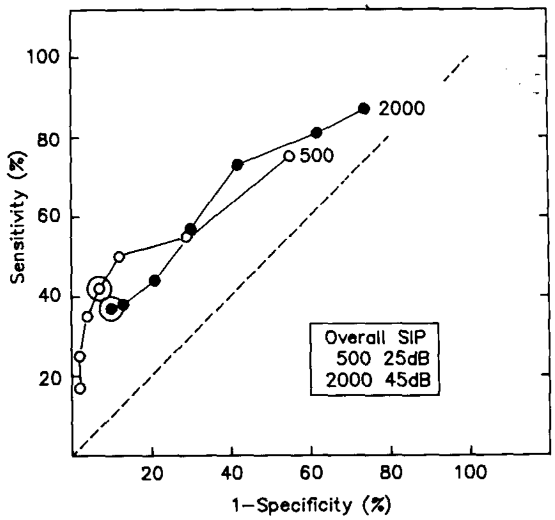

Figure 4 Receiver operating curves for threshold levels within 0.5 and 2 kHz-frequencies with the Overall SIP as the functional standard. Impairment defined as an Overall SIP score of 9.1 or greater.

Table 7 Prevalence of Hearing Loss by Derived Criteria. Differences Between Combining Frequencies With "AND" Function or "OR" Function*

Functional Standard	Frequencies	Threshold Level	N	Prevalence (%) "AND"	Prevalence (%) "OR"
HHIE-S	500 4000	35 45	177	13	49
Psychosocial SIP	250 2000	30 45	153	10	28
Physical SIP	500 2000 1000	25 45 30†	153	11	35
Overall SIP	500 2000	25 45	153	11	33

*Person considered impaired if observed threshold greater than or equal to proposed threshold at either frequency (e.g., for the HHIE-S, consider impaired if observed threshold GE 35 dB at 500 Hz *or* GE 45 dB at 4000 Hz).

†After adjustment for 500 Hz and 2000 Hz, an increasing threshold at 1000 Hz was negatively correlated with Physical SIP (see text).

differing functional scales. For the *OR* combination the prevalence of impairment ranged from 28 to 49 percent across the differing functional scales. As the reported prevalence of hearing impairment in the aged is 25 to 40 percent (Herbst, 1983; Moss and Parson, 1986), we believe that the *OR* function is the optimal way to combine the thresholds recorded at each frequency. So, for example, with the Overall SIP as the functional standard, the criteria for hearing impairment would be a greater than or equal to 25-dB loss at 0.5 kHz *OR* a greater than or equal to 45-dB loss at 2 kHz using poorer ear thresholds.

Comparing Functionally Derived Criteria For Hearing Impairment With Other Criteria

We then sought to compare the functionally derived criteria for hearing impairment with other criteria, the HFPTA, SFPTA, SRT, and criterion of Ventry and Weinstein. The Kappa statistics for the overlap between the hearing impairment criteria are given in Table 8. Kappa values represent the proportion of agreement between two tests that is not due to chance. A value of zero implies that all agreement is due to chance, and a value of one implies that there is perfect agreement between criteria. From Table 8 it is seen that the HHIE-S derived criteria agrees best with the HFPTA (Kappa = 0.58), with lower levels of agreement for the other "traditional criteria." For the SIP derived criteria, the agreement is poorest with the HFPTA (Kappa range 0.34 to 0.42) but is quite high for the other criteria. Kappa values here ranged from 0.65 for the Psychosocial SIP criteria and SRT up to 0.84 for the Physical SIP and SFPTA. Overall, the functionally derived criteria overlap a great deal with older, more "traditional," criteria.

Table 8 Kappa Statistics for Agreement Beyond That Due to Chance for the Derived Criteria for Hearing Loss and Other Criteria for Hearing Loss

Other Criteria	Criteria Derived from Functional Standards: HHIE-S (N=177)	Physical SIP (N=153)	Psychosocial SIP (N=153)	Overall SIP (N=153)
HFPTA	0.58	0.34	0.42	0.38
SFPTA	0.46	0.84	0.74	0.80
SRT	0.37	0.70	0.65	0.68
Ventry and Weinstein	0.48	0.70	0.74	0.74

*All kappa values statistically significant at p = <0.01

The functionally derived criteria were then compared with the "traditional" criteria in stepwise multiple linear regressions. In these regressions the functional scales served as continuous dependent variables. The subjects were categorized by whether they passed or failed each of the criteria, and these were used as the independent variables. The results are summarized in Table 9. Again, for each of the regressions, the residuals were normally distributed and did not differ substantially from zero. When the HHIE-S (a hearing specific functional scale) was the dependent variable, the regression selected the criteria of Ventry and Weinstein, accounting for 27 percent of the variance, and the HFPTA, accounting for 3 percent of the variance. The HHIE-S derived criteria did not enter the model.

With the SIP (a more global measure of function) scales as the dependent variables, the newly derived criteria were the only criteria to enter the models in the regression analyses. For the Physical, Psychosocial, and Overall SIP scales, the SIP derived criteria accounted for 19 percent, 12 percent, and 19 percent of the variance respectively.

Table 9 Regression Equations for Selecting Criteria for Hearing Loss. Functional Criteria Compared with Other Criteria

	Regression Equation			
*Functional Standard**	*Hearing Loss Criteria Chosen†*	*Regression Coefficient‡ Beta*	*(S.E.)*	*R^2§*
HHIE-S (N=177)	Ventry and Weinstein	9.4	(1.5)	27%
	HFPTA	4.1	(1.5)	3%
Physical SIP (N=153)	Physical SIP	9.1	(1.5)	19%
Psychosocial SIP (N=153)	Psychosocial SIP	9.1	(2.0)	12%
Overall SIP (N=153)	Overall SIP	10.5	(1.8)	19%

*Functional standards as continuous dependent variables
†Hearing criteria as pass/fail variables
‡Change in functional standard score for a failing criteria compared to passing the criteria
§Percent variance attributed to each criteria as it entered the model. Total variance of the model attained by adding the separate variances.

DISCUSSION

This paper has described a method for using measures of functional status, the HHIE-S and the SIP, to derive hearing impairment criteria for the elderly. The analyses were performed in such a way as to maximize the accuracy of the criteria in categorizing individuals as impaired or unimpaired as judged by these functional standards. The analyses proceeded in five steps.

First, the between and within ear threshold levels were correlated among frequencies. This indicated a high degree of dependence between adjacent frequencies and between ears for this population of aged individuals. The implication from this finding was that a combination of a relatively low frequency and high frequency were most likely to be part of a functional criterion.

Second, better and poorer ear thresholds were correlated with the functional scales. Because poorer ear thresholds were more closely correlated with the functional scales, we used poorer ear thresholds as a means of increasing accuracy of deciding whether a person was hearing-impaired.

Third, stepwise multiple linear regressions were used to select the frequencies that accounted for the most variance in the functional scales.

Fourth, receiver operating curves (sensitivity, specificity, and test accuracy) were generated to determine the threshold criterion within each frequency that gave the best discrimination between impaired and unimpaired individuals.

Fifth, the functionally derived criteria were compared to other criteria of hearing impairment through examination of agreement between criteria and multiple linear regressions with the functional scales as the dependent variables.

The findings of this study should be interpreted within its limitations. First, the criteria derived for hearing impairment depend on the functional scale chosen to define impairment. With the HHIE-S as the functional standard, combining the criteria of Ventry and Weinstein and the HFPTA provided the best estimate of impairment; the criteria derived through our approach did not enter the regression model. As the correlation between pure tone thresholds and the HHIE-S is on the order of 0.50, there may still be room for developing a functional scale of hearing impairment that is more closely associated with hearing loss. When the SIP scales served as functional standards, the criteria were slightly different depending on which SIP scale was used. Comparing the criteria derived from SIP data with more "traditional" criteria, the functionally derived criteria accounted for more variance in the SIP scales and were the only criteria to enter the regression models.

A second limitation is that the threshold criteria will change if the definition of impairment within each functional scale changes. If more stringent cut points for impairment are set (e.g., an HHIE-S score of 24 or the upper quintile of SIP scores), then it is likely that higher threshold levels would provide the better discrimination between impaired and unimpaired. The cut point chosen for the HHIE-S is the published guideline for impairment. For the SIP scales, no normative data exist to firmly establish what constitutes impairment, as the SIP measures a continuum of sickness-related dysfunction (W. Carter, personal communication, 1988). We chose the upper third of the SIP distributions as indicative of impairment to have an unbiased cutpoint and to preserve statistical power in our analyses.

A third limitation in selecting threshold level criteria comes in selecting the point of best test accuracy ("fence"). Stratifying the groups by threshold strata resulted in small numbers of subjects in some strata. Thus, test accuracies of 70 percent at one threshold versus 72 percent at another threshold may be statistically and clinically insignificant. Although our sample is of moderate size, it is not robust enough to clearly differentiate the single best threshold "fence" for each frequency. Our larger purpose however is to demonstrate a method for deriving criteria, and to that end the results are valid.

We do not suggest that the functionally derived criteria presented in this paper should be accepted as standard criteria of hearing impairment in the aged. To set such standards would require independent derivation in other samples of aged individuals. These samples should be representative of a defined group of aged persons, screened without respect to whether or not they have perceived hearing impairment (avoidance of a volunteer bias), and be substantially larger than the present study to be able to establish fences with more precision. We welcome independent derivation and validation of our results in independent samples of the elderly.

Based on our experience with this study, if criteria for hearing impairment in the elderly are functionally derived, we believe that when replicated, criteria derived by others will contain the following features:

1. The criteria will combine a relatively low frequency with a relatively high frequency. The lower frequency will be functionally more important than the higher frequency, but the higher frequency will contribute some functional impairment information.
2. The threshold criterion for the lower frequency will be in the 25- to 35-dB range. The threshold criterion for the higher frequency will be in the 40- to 45-dB range.

These features provide a suitable compromise to the current debate over which threshold levels comprise the best clinical identification of aged persons who are hearing-impaired.

In spite of the study limitations we believe this functional approach for determining criteria for hearing impairment is crucial to settling debates over which combinations of frequencies and threshold levels are the best discriminators of hearing impairment. It may be that simply dichotomizing individuals into impaired and unimpaired based on hearing level criteria is too crude to justly serve their needs. An alternative approach may be to measure hearing level and then decide whether to rehabilitate the individual with a hearing aid based on their perceived handicap (as assessed by a functional scale). Thus some persons with minimal hearing loss, but perceived handicap, would be candidates for rehabilitation; alternatively, a person with moderate hearing loss, but no perceived handicap, might not be a candidate for rehabilitation. Practicing audiologists may consider using the criteria put forward here and comparing their accuracy against more traditional impairment criteria in identifying candidates for aural rehabilitation. Future research should focus on developing consensus standards for functional impairment in the aged and then independently testing the efficacy of hearing aid rehabilitation on reducing impairment at differing levels of hearing loss.

Acknowledgments. We thank Drs. Anderson Spickard, Frank Gluck, Taylor Hays, Craig Heim, Mark Houston, John Griscom, and F. Tremaine Billings for taking time to participate in this study; Mary Sue Fino and Andrea Hedly Williams for help with data collection and entry; Dottie Adams for providing expert manuscript preparation; and Barbara Weinstein and William Schaffner for providing critiques of earlier versions of the manuscript.

This work was supported in part by grants from The Retirement Research Foundation and The Oticon Foundation

REFERENCES

American National Standards Institute. (1972). Specifications for audiometers. *ANSI* S3.6-1969. New York:ANSI.

American National Standards Institute. (1977). Criteria for permissable ambient noise during audiometric testing. *ANSI* S3.1-1977. New York:ANSI.

Armitage P. (1971). Multiple regression and multivariate analysis. In: *Statistical methods in medical research.* Oxford: Blackwell Scientific Publ, 302-348.

ASHA (American Speech-Language-Hearing Association). (1981). Task force on when hearing impairment becomes a hearing handicap. On the criterion of hearing handicap. (Committee report.) *ASHA* 23:293-297.

Bergner M, Bobbitt RA, Carter WB, Gilson BS. (1981). The Sickness Impact Profile: developments and final revision of a health status measure. *Med Care* 19:787-805.

Bergner M, Bobbitt RA, Pollard WE, Martin DP, Gilson BS. (1976). The Sickness Impact Profile: validation of a health status measure. *Med Care* 14:57-67.

Bergner L, Hallstrom AP, Bergner M, Eisenberg MS, Cobb LA. (1985). Health status of survivors of cardiac arrest and of myocardial infarction controls. *Am J Public Health* 75:1321-1323.

Bess FH. (1985). The minimally hearing-impaired child. *Ear Hear* 6:43-47.

Bess FH, Lichtenstein MJ, Logan SA, Burger MC, Nelson EC. (1989a). Hearing impairment as determinant of function in the elderly. *J Am Geriatr Soc* 37:123-128.

Bess FH, Lichtenstein MJ, Logan SA, Burger MC. Comparing criteria of hearing impairment in the elderly: a functional approach. *J Speech Hear Res* 1989b 32:795-802.

Carhart R. (1971). Observations on relations between thresholds for pure tones and for speech. *J Speech Hear Disord* 36:476-483.

Carhart R, Jerger JF. (1959). Preferred method for clinical determination of pure tone threshold. *J Speech Hear Disord* 24:330-345.

Carter WB, Bobbitt RA, Bergner M, Gilson BS. (1976). Validation of an internal scaling: The Sickness Impact Profile. *Health Serv Res* 74:516-528.

Clark JG. (1981). Uses and abuses of hearing loss classification. *ASHA* 23:493-500.

Deyo RA, Inui TS. (1984). Toward clinical applications of health status measures: sensitivity of scales to clinically important changes. *Health Serv Res* 19:275-289.

Gilson BS, Gilson JS, Bergner M, Bobbitt RA, Kressel S, Pollard WE, Vesselago M. (1975). The Sickness Impact Profile: development of an outcome measure of health care. *Am J Public Health* 65:1304-1310.

Goldstein D. (1984). Hearing impairment, hearing aids and audiology. *ASHA* 26:24-35, 38.

Hart LG, Evans RW. (1987). The functional status of ESRD patients as measured by the Sickness Impact Profile. *J Chronic Dis* 40:117S-130S.

Herbst KRG. (1983). Psychological consequences of disorder of hearing in the elderly. In: Hinchcliffe R, ed. *Hearing and balance in the elderly.* New York: Churchill Livingstone, 174-200.

Lichtenstein MJ, Bess FH, Logan SA. (1988). Validation of screening tools for identifying hearing-impaired elderly in primary care. *J Am Med Assoc* 259:2875-2878.

McSweeney A, Grant I, Heaton RK, Adams KM, Tims RM. (1982). Life quality of patients with chronic obstructive pulmonary disease. *Arch Intern Med* 142:473-478.

Moss AJ, Parsons VL. (1986). Current estimates from the National Health Interview Survey, 1985. *Vital Health Stat* [10], Number 160.

Noble WG. (1978). *Test manual for the hearing measurement scale.* Armidele N, S. W. Australia: University of New England.

Pollard WE, Bobbitt RA, Bergner M, Martin DP, Gilson BS. (1976). The Sickness Impact Profile: reliability of a health status measure. *Med Care* 14:146-155.

Rosen J. (1978). The evaluation of handicap secondary to acquired hearing impairment. *J Acad Rehab Audiol* 11:2-9.

Sackett DL, Haynes RB, Tugwell P. (1985). The clinical examination. In: *Clinical epidemiology: a basic science for clinical medicine.* Boston, MA/Toronto, Canada: Little, Brown, 17-46.

Solomon G. (1986). Hearing problems in the elderly. *Dan Med Bull,* Special Supplement Series, 3:1-22.

Tillman TW, Olsen WO. (1973). Speech audiometry. In: Jerger J, ed. *Modern developments of audiology.* 2nd ed. New York: Academic Press, 37-74.

Ventry IM, Weinstein B. (1982). The hearing handicap inventory for the elderly: a new tool. *Ear Hear* 3:128-134.

Ventry IM, Weinstein G. (1983). Identification of elderly people with hearing problems. *ASHA* 25:37-42.

Ward PR, Tucker AM, Tudor CA, Morgan DC. (1977). Self-assessment of hearing impairment: test of the expanded hearing ability scale questionnaire on H-I adults in England. *Br J Audiol* 11:33-39.

Weinstein BE. (1984). A review of hearing handicap scales. *Audiology* 9:91-109.

APPENDIX

Tabulation of sensitivity, specificity, and test accuracy of each dB threshold for each frequency:
Table A. HHIE-S as functional standard
Table B. Psychosocial SIP as functional standard
Table C. Physical SIP as functional standard
Table D. Overall SIP as functional standard

Table A Sensitivity, Specificity, and Test Accuracy Stratified by Frequency and Threshold Level. HHIE-S as Functional Standard (N=177)*

	Frequency (Hz)				
Threshold Level	*250*	*500*	*1000*	*2000*	*4000*
Sensitivity (%):					
15 dB	81	82	88	93	99
20 dB	67	57	76	88	94
25 dB	51	46	58	78	91
30 dB	39	39	43	67	88
35 dB	28	33	30	58	87
40 dB	21	22	25	45	75
45 dB	16	18	13	40	66
Specificity (%):					
15 dB	30	45	50	25	6
20 dB	52	68	59	40	10
25 dB	77	82	81	60	18
30 dB	87	90	89	72	31
35 dB	95	96	94	86	43
40 dB	97	98	96	89	52
45 dB	98	99	98	92	65
Accuracy (%):					
15 dB	49	59	65	51	41
20 dB	58	63	71	58	42
25 dB	67	69	72	67	46
30 dB	69	71	70	71	56
35 dB	70	72	70	75	59
40 dB	69	70	69	72	60
45 dB	67	69	66	72	66

*HHIE-S: Not handicapped = 0-8 (N = 111), Handicapped = 10-40 (N = 67)

Table B Sensitivity, Specificity, and Test Accuracy Stratified by Frequency and Threshold Level. Psychosocial SIP as Functional Standard (N=152)*

	Frequency (Hz)				
Threshold Level	*250*	*500*	*1000*	*2000*	*4000*
Sensitivity (%):					
15 dB	79	77	77	85	94
20 dB	67	50	63	77	92
25 dB	46	44	44	69	88
30 dB	40	35	29	54	88
35 dB	27	29	23	40	83
40 dB	21	21	19	31	63
45 dB	15	15	10	29	52
Specificity (%):					
15 dB	33	45	46	23	4
20 dB	57	69	62	36	9
25 dB	81	85	74	57	17
30 dB	92	89	84	68	29
35 dB	96	93	91	78	38
40 dB	98	96	94	84	43
45 dB	98	97	96	87	58
Accuracy (%):					
15 dB	49	56	57	44	35
20 dB	60	61	63	50	38
25 dB	69	71	64	61	42
30 dB	74	70	65	63	50
35 dB	72	71	66	65	54
40 dB	72	70	68	66	50
45 dB	70	69	66	67	56

*Psychosocial SIP: Unimpaired = 0.0-5.4 (N = 100), Impaired = 5.5- (N = 52)

Table C Sensitivity, Specificity, and Test Accuracy Stratified by Frequency and Threshold Level. Physical SIP as Functional Standard (N=152)*

	Frequency (Hz)				
Threshold Level	*250*	*500*	*1000*	*2000*	*4000*
Sensitivity (%):					
15 dB	77	73	85	88	96
20 dB	63	58	67	85	96
25 dB	50	52	50	75	90
30 dB	46	44	38	61	90
35 dB	31	37	29	48	85
40 dB	23	27	23	38	63
45 dB	17	19	13	37	54
Specificity (%):					
15 dB	32	44	50	25	5
20 dB	56	73	64	40	11
25 dB	84	89	77	59	18
30 dB	96	94	89	71	30
35 dB	98	97	94	81	39
40 dB	99	99	96	87	43
45 dB	99	99	98	90	58
Accuracy (%):					
15 dB	47	50	61	46	36
20 dB	58	68	65	55	40
25 dB	72	76	68	65	43
30 dB	78	77	72	68	51
35 dB	75	76	72	70	55
40 dB	73	75	71	71	50
45 dB	71	72	69	72	57

*Physical SIP: Not handicapped = 0.0-4.3 (N = 100), Handicapped = 4.4- (N = 52)

Table D Sensitivity, Specificity, and Test Accuracy Stratified by Frequency and Threshold Level. Overall SIP as Functional Standard (N=152)*

	Frequency (Hz)				
Threshold Level	*250*	*500*	*1000*	*2000*	*4000*
Sensitivity (%):					
15 dB	81	73	83	87	96
20 dB	67	54	63	81	96
25 dB	50	50	48	73	90
30 dB	44	42	38	58	88
35 dB	29	35	29	44	83
40 dB	21	25	23	38	71
45 dB	15	17	13	37	60
Specificity (%):					
15 dB	34	45	48	26	5
20 dB	57	71	62	38	11
25 dB	83	88	76	58	18
30 dB	94	93	89	69	29
35 dB	97	96	94	79	38
40 dB	98	98	96	87	47
45 dB	98	98	98	90	62
Accuracy (%):					
15 dB	50	55	60	48	36
20 dB	61	65	63	53	40
25 dB	72	75	66	63	43
30 dB	77	76	72	65	50
35 dB	74	75	72	67	54
40 dB	72	73	71	70	56
45 dB	70	70	69	72	61

*Overall SIP: Unimpaired = 0.0-9.0 (N = 100), Impaired = 9.1- (N = 52)