Effectiveness of Three Central Auditory Processing (CAP) Tests in Identifying Cerebral Lesions

Raymond M. Hurley* 
Frank E. Musiek'

Abstract

The effectiveness of three central auditory processing (CAP) tests was evaluated using clinical decision analysis (CDA) procedures. The tests under study were the Dichotic Digits Test (DDT), the Auditory Duration Patterns Test (ADPT), and the P300 evoked potential test (P300). Subjects with normal hearing, sensorineural hearing loss (SNHL), and central lesions provided the data for CDA analyses. To identify the most effective test, we used the conventional CDA parameters of hit rate (sensitivity), false-positive rate (1 - false-positive rate = specificity), and A' (test performance at specific hit/false-positive combinations). Further, we illustrated the extension of the conventional CDA parameters to posterior probability determination, which incorporates disorder prevalence to compute the likelihood of a patient having a disorder when a test result is positive (Pr[D/+]) or not having a disorder when the test result is negative (Pr[N/-]). Last, we used the CDA parameter of hit rate and disorder prevalence to determine cost effectiveness.

Key Words: Auditory Duration Patterns Test, central auditory system, central auditory test, clinical decision analyses, Dichotic Digits Test, P300

Abbreviations: ADPT = Auditory Duration Pattern Test; CANS = central auditory nervous system; CAP = central auditory processing; CDA = clinical decision analysis; DDT = Dichotic Digit Test; NH = normal hearing; NL = nonlesion subject; LS = lesion subject; P300 = event related evoked potential; ROC = receiver operating characteristic; SNHL = sensorineural hearing loss; HR = hit rate, sensitivity, true-positive rate; FPR = false-positive rate, false alarm rate; A' = measure of overall test performance; Pr[D/+] = probability of being correct with a positive test result; Pr[N/-] = probability of being correct with a negative test result

The identification of central auditory processing (CAP) disorders has received considerable attention over the past 3 decades. The germinal work of Bocca et al (1954) arose from the observation that patients with brain lesions and complaints of hearing problems had normal pure-tone audiograms. CAP tests were initially used to assess the integrity of the central auditory nervous system (CANS) for adult patients with a variety of neurologic disorders including neoplasms, degenerative disease, vascular disorder, or trauma (Bocca et al, 1954). Previous research has demonstrated the utility of various assessment measures to identify a CANS dysfunction, often with qualified conclusions (Lynn and Gilroy, 1984; Mueller et al, 1987). Sources that may account for these outcomes include the limitations of the current test procedures and the heterogeneous nature of CANS lesions. Superimposed on these issues have been the limitations resulting from inadequate data analyses and small subject samples.

The purpose of the present retrospective investigation was to critically re-examine the Dichotic Digits Test (DDT) (Musiek, 1983), the Auditory Duration Patterns Test (ADPT) (Musiek et al, 1990), and the P300 evoked potential test (P300) (Musiek et al, 1992). While we have reported previously on these three tests and
demonstrated that CANS dysfunction subjects performed significantly poorer than non-CANS dysfunction subjects on each of these measures (Musiek et al, 1990, 1991, 1992), we wanted to show how determining test efficacy could be improved by using clinical decision analyses (CDAs) (Schultz, 1972; Jerger, 1983; Hyde et al, 1991; Turner, 1991). CDA has been used to determine the effectiveness of retrocochlear tests (Turner et al, 1984; Musiek et al, 1996) and the effectiveness of otoacoustic emissions (Gorga et al, 1993; Hurley and Musiek, 1994); however, CDA has not been used to determine the effectiveness of CAP tests. CDA, which has grown out of signal detection theory (Swets, 1988), examines a sample by determining relationships between presence and absence of a disorder and whether or not test results are positive or negative. In addition to considering hit rate (sensitivity), CDA takes into account false-positive rate (1 - false-positive rate = specificity) and A', which is a measure of overall test performance at specific hit/false-positive combinations. A' values range from 0.50 for a test of no diagnostic value to 1.00 for the perfect diagnostic test. In addition to CDAs, we provide an estimation of the cost effectiveness based on the CDAs of the three tests.

**METHOD**

**Subjects**

Three separate groups of subjects provide the data for the CDAs: subjects with normal hearing (NH), subjects with sensorineural hearing loss (SNHL), and subjects with neurologically, radiologically, and/or surgically confirmed CANS lesions (LS). Briefly, the subjects with NH had HLLs of ≤25 dB for the octave frequencies of 500 to 4000 Hz while the subjects with SNHL were diverse in both the configuration and the degree of their hearing loss. HLLs ranged from 0 to 95 dB across the octave frequencies of 500 to 4000 Hz. However, all of the subjects with SNHL had a HL of >25 dB for at least two frequencies in the 500- to 4000-Hz range in both ears. While diverse in etiology and hemisphere involvement (Table 1), all subjects in the LS group had a lesion involving the auditory areas of the cerebrum (Galaburda and Sanides, 1980; Musiek, 1986).

For the ADPT, the data from 50 NH subjects, 24 SNHL subjects, and 21 LS subjects were used (Musiek et al, 1990). In the NH group, there were 42 females and 8 males with an age range of 19 to 39 years. The SNHL group consisted of 8 females and 16 males with an age range of 22 to 73 years. Figure 1a displays the average HLLs (±1 SD) for the SNHL subjects. The LS group had 13 females and 8 males with an age range of 16 to 58 years and neurologically, radiologically, and/or surgically confirmed CANS lesions. Specific lesions in the LS subjects included seven with epileptic foci, seven with mass lesions, two with vascular lesions, two with trauma, one with inflammation and infection, and one with undefined neuropathy. Each LS subject had HLLs of ≤25 dB for the octave frequencies of 500 to 4000 Hz.

The data for the DDT were obtained from 40 NH subjects, 30 SNHL subjects, and 26 LS subjects (Musiek et al, 1991). Figure 1b displays the average HLLs for the SNHL subjects. The NH group consisted of 28 females and 12 males with an age range of 21 to 33 years. In the SNHL group, there were 9 females and 21 males with an age range of 14 to 64 years. Figure 1b displays the average HLLs (±1 SD) for the SNHL subjects. The LS group consisted of 17 females and 9 males ranging in age from 13 to 73 years with neurologically, neurosurgically, or radiologically confirmed lesions of the CANS. Specific lesions in the LS subjects included 11 with epileptic foci, 7 with mass lesions, 3 with vascular lesions, and 5 with trauma. Each LS subject had HLLs of ≤25 dB for the octave frequencies of 500 to 4000 Hz.

The data for the P300 analyses were obtained from 20 LS and 20 control subjects matched for age and HL at 1000 and 2000 Hz.
Effectiveness of Central Auditory Processing Tests/Hurley and Musiek

(Musiek et al, 1992). In the LS group, there were 9 females and 11 males ranging in age from 18 to 65 years with neurologically, neurosurgically, or radiologically confirmed lesions of the CANS. Specific lesions in the LS subjects included two with epileptic foci, nine with mass lesions, six with vascular lesions, two with trauma, and one with cortical atrophy.

Procedures

The ADPT (Musiek et al, 1991) is a sequence of three consecutive 1000-Hz tones with one differing by being either longer (L), 500 msec, or shorter (S), 200 msec, in duration than the other two tones in the sequence. The tones have a rise/fall time of 10 msec and an interstimulus interval of 300 msec. Six different sequences, LLS, LSL, LSS, SLS, SLL, and SSL, are used in the test. A total of 30 three-tone sequences were presented monaurally at 50 dB re: speech reception threshold (SRT). Previous to testing, several practice items were provided to each subject to ensure their understanding of the task. The subjects respond to each stimulus presentation with a verbal description of the sequence heard.

The DDT (Musiek, 1983) is composed of nine naturally spoken digits from 1 through 10, excluding the digit 7. Two digits are presented to each ear dichotically at 50 dB re: SRT. Forty presentations of the dichotic digit pairs constituted a complete test. Three practice dichotic digit pairs preceded the 40 test presentations to ensure that each subject understood the task. A free recall response mode was used.

For the P300 event related potential measures, the “oddball” paradigm was used with the frequent tone, 1000 Hz, occurring 85 percent of the time and the rare tone, 2000 Hz, appearing 15 percent of the time. The stimuli were presented at 75 dB SPL and at a rate of 0.8/sec. The subjects were instructed to count the number of “rare” tones they heard and to report this number after each block of trials. Each subject received an instructional/practice period to ensure their understanding of the task. In P300 latency determination, if the peak was bifid, then the largest of the two peaks was selected. Amplitude was computed by measuring the trough-to-peak of the front and back sides of the P300 wave and taking the average of these two measurements (Gooding et al, 1986). Similarly, the amplitude of the P2 wave was determined to compute the P300/P2 amplitude ratio. The latency and amplitude measures were the average of two trials. The P300 was considered to be absent when there was no replicable activity in the 250- to 600-msec latency range. Agreement between two reviewers familiar with these recordings on the latency and amplitude values was necessary for the measurements to be included in this study.

As the goal of this retrospective investigation was to determine the performance of the three CAP tests in a clinically relevant manner, the data for NH and SNHL subjects were collapsed with the belief that a CANS lesion can exist concurrently with NH or SNHL and thus a CAP test should be independent of peripheral hearing sensitivity. In short, interpretation of a CAP test result should not be qualified by a mild to moderate SNHL (Miltenberger et al, 1978). Thus, the NH and SNHL subjects were combined to form our nonlesion (NL) subject group.

To determine the ability of the three tests to separate the two subject groups, the data were analyzed using CDA procedures (Schultz, 1972; Jerger, 1983; Hyde et al, 1991; Turner, 1991). Recall that the primary measures used in CDA are the hit rate, false-positive rate, and A'. For the DDT, ADPT, and the P300 amplitude and amplitude ratio parameters, the scores of the LS and NL groups were analyzed to provide the percentage of subjects correctly identified as
having a test score less than a criterion score, a hit in the LS group, and the percentage of subjects incorrectly identified by having a test score less than a criterion score, a false positive in the NL group. For the P300 latency parameter, the latencies of the LS and NL groups were analyzed to provide the percentage of subjects correctly identified as having a latency value greater than a criterion latency, a hit in the LS group, and the percentage of subjects incorrectly identified by having a latency value greater than a criterion latency, a false positive in the NL group. In CDA, the criterion score is systematically varied resulting in corresponding changes in the hit rate and false-positive rate, thus generating a receiver operating characteristic (ROC) curve. While a detailed explanation of CDA is beyond the scope of this paper, an expanded treatment of CDA is provided by Hyde et al. (1991) and Turner (1991).

**RESULTS**

The results of these CDAs are portrayed in the ROC curves depicted in the upper segment of Figure 2. The data were analyzed using a $2 \times 2$ matrix with the categories of LS or NL constituting the columns and the test score constituting the rows. In addition, the $A'$ value was computed as a measure of overall test performance (Turner, 1991). The ability of each test score to separate subjects is demonstrated by the degree of ROC curve displacement toward the upper left corner where the hit rate is high and the false-positive rate is low. An inspection of the ROC curves in Figure 2 reveals that the CAP tests were variable in their ability to separate the LS and NL groups. The ADPT had a hit rate of 83 percent and a false-positive rate of 8 percent while the DDT had a hit rate of 75 percent and a false-positive rate of 9 percent. The P300 latency parameter had a hit rate of 57 percent and a false-positive rate of 10 percent, the P300 amplitude parameter had a hit rate of 59 percent and a false-positive rate of 12 percent, and the P300 amplitude ratio (P300/P2) parameter had a hit rate of 31 percent and a false-positive rate of 9 percent. These hit rate/false-positive rate data are depicted in Figure 2b. Test scores (values) that correspond to the above hit rate/false-positive rate combinations were <83 percent on the ADPT and <90 percent on the DDT. The test values for the P300 parameters were >340 msec for latency, <7.0 $\mu$V for amplitude, and <1.0 for amplitude ratio.

The hit rates and $A'$ values for the approximate 10 percent false-positive rate are displayed in Table 2. In addition to the hit rate, false-positive rate, and $A'$ value, we computed the posterior probability of an LS given a positive test result (Pr[D/+]) and the posterior probability of NL given a negative test result (Pr[N/-]). The posterior probabilities are a function of hit rate, false-positive rate, and the prevalence of the disorder. While the true prevalence of CANS lesions is unknown, we used an assumed prevalence of 5 percent in the calculation of the posterior probabilities. Based on the $A'$ value, the ADPT was the best overall test, with an $A'$ of 0.93 followed closely by the DDT with an $A'$ of 0.90 and the P300 latency parameter with an $A'$ of 0.84. The ADPT demonstrated the best posterior...
Table 2  Summary of the Clinical Decision Analyses (CDAs) for the Three Tests

<table>
<thead>
<tr>
<th>Test</th>
<th>HR (%)</th>
<th>FPR (%)</th>
<th>(Pr[D+/])</th>
<th>(Pr[N/-])</th>
<th>A' (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADPT</td>
<td>83</td>
<td>08</td>
<td>0.93</td>
<td>35</td>
<td>99</td>
</tr>
<tr>
<td>DDT</td>
<td>75</td>
<td>09</td>
<td>0.90</td>
<td>30</td>
<td>98</td>
</tr>
<tr>
<td>P300 latency</td>
<td>57</td>
<td>10</td>
<td>0.84</td>
<td>23</td>
<td>98</td>
</tr>
<tr>
<td>P300 amplitude</td>
<td>59</td>
<td>12</td>
<td>0.83</td>
<td>21</td>
<td>98</td>
</tr>
<tr>
<td>P300 amplitude ratio</td>
<td>31</td>
<td>09</td>
<td>0.74</td>
<td>15</td>
<td>96</td>
</tr>
</tbody>
</table>

While we have demonstrated the utility of applying CDA to CAP tests, we want to caution against the literal interpretation (generalization) of the CDA results because of the retrospective nature of the study and the fact that not all of our subjects were administered all three CAP tests. In addition, the three LS groups were not matched for lesion etiology and hemisphere involvement; thus, the degree of lesion involvement among the three LS groups was not controlled. Again, the primary purpose of this investigation was to illustrate how determining CPA test efficacy could be improved by using CDA and not to provide a ranking of the three CAP tests.

Recall that to evaluate further test efficacy, we considered posterior probability, which indicates the likelihood of a patient having a CANS disorder when a test result is positive (Pr[D+/]) or not having a CANS disorder when the test result is negative (Pr[N/-]). In interpreting the posterior probability values, one must keep in mind that the positive and negative posterior probabilities are not equally affected by the prevalence of a disorder. For instance, if the hit rate and false-positive rate for the auditory brainstem response to identify a retrocochlear lesion are 95 percent and 11 percent, respectively, and the prevalence of a retrocochlear lesion is 2 percent, the positive posterior probability (Pr[D+/]) is 15 percent. With a prevalence of 5 percent or 10 percent, the positive posterior probabilities (Pr[N+/]) are 31 percent and 49 percent, respectively. For the negative posterior probability (Pr[N/-]), disorder prevalence of 2 percent, 5 percent, and 10 percent produce negative posterior probabilities (Pr[N/-]) of >99 percent, >99 percent, and 99 percent, respectively. It is only when the prevalence of a disorder is high, such as 50 percent, that the negative posterior probabilities (Pr[N/-]) are affected by the hit rate and false-positive rate.

Another factor in determining test usage should be cost effectiveness. In our computations, we based our cost-effectiveness estimation on 100 administrations of a test and a CANS prevalence of 5 percent although the true prevalence of CANS disorders is unknown. For the ADPT and DDT, we assumed a cost per test of $40.00, which is the median fee charged for a CAP test at the audiology facilities we sampled. The cost of testing 100 individuals with either the ADPT or DDT would be the number tested times the cost per test ($40 \times 100$) or $4000. The ADPT with a hit rate of 83 percent and with a CANS prevalence of 5 percent, five individuals will be designated an LS with (5 \times 0.83) 4.15 identified by the ADPT. The cost per LS identified is the total cost of testing divided by the number of hits. Since there were 4.15 hits for the ADPT, the cost per identification is $963.85. The DDT with a hit rate of 75 percent and with a CANS prevalence of 5 percent, five individuals will be designated an LS with (5 \times 0.75) 3.75 identified by the DDT. The cost per LS identified is the total cost of testing divided by the number of hits. Since there were 3.75 hits for the DDT, the cost per identification is $1066.67. For the P300, we assumed a cost per test of $200.00, which is the median fee charged for evoked potential testing at the audiology facilities we sampled. With 100 administrations of the test, the cost of testing would be the number tested times the cost per test ($200 \times 100$) or $20,000. The P300 with a hit rate of 57 percent for the latency parameter and with a prevalence of 5 percent, five individuals will be designated with (5 \times 0.57) 2.85 LS identified by the P300. The cost per LS identified is the total cost of testing divided by the number of hits. Since there were 2.85 hits for the P300, the cost per identification is $7017.54.

In summary, to critically examine the CAP tests, we used the conventional CDA parameters of hit rate (sensitivity), false-positive rate (1 - false positive rate = specificity), and A' (test performance at specific hit/false-positive combinations). We illustrated the extension of the conventional CDA parameters to posterior probability determination, which incorporates disorder prevalence to compute the likelihood of a patient having a disorder when a test result is positive (Pr[D+/]) or not having a disorder when the test result is negative (Pr[N/-]). Further,
we cautioned that the positive and negative posterior probabilities values are not equally affected by the prevalence of a disorder. Last, we used the hit rate and disorder prevalence to determine cost effectiveness for each of the three tests.

Acknowledgment. Preliminary data were presented at the Annual Convention of the American Academy of Audiology, Nashville, TN in 1992. We thank our colleagues Karen M. Gollegly, M.A., Karen S. Kibbe, M.A., and Suzanne B. Lenz, M.A. for their help in collection of these data. We also thank Anne Marie Tharpe, Ph.D. and the three anonymous reviewers for comments on a previous version of this paper.

REFERENCES


