

Efficacy of language assessment in Alzheimer's disease: comparing in-person examination and telemedicine

Lindsey Vestal¹,
Laura Smith-Olinde¹,
Gretchen Hicks¹,
Terri Hutton¹,
John Hart Jr¹⁻³

¹Department of Speech Pathology,
University of Arkansas at Little Rock;
²GRECC of CAVHS, Little Rock, AR,
USA; ³Center for BrainHealth,
University of Texas at Dallas, Dallas,
TX, USA

Background: With the large number of aging individuals requiring screening of cognitive functions for dementing illnesses, there is a necessity for innovative evaluation approaches. One domain that should allow for online, at a distance, examination is speech and language dysfunction, if the auditory and visual transmission is of sufficient quality to allow adequate patient participation and reliable, valid interpretation of signs and symptoms (Duffy et al 1997).

Objective: Examine the effectiveness of language assessment in mild Alzheimer's patients using telemedicine (TM) compared with traditional in-person (IP) assessment.

Design: Ten patients with mild Alzheimer's disease, enrolled at a Geriatric Memory Clinic received a battery of standard language tests under two conditions: face-to-face and via satellite TM.

Results: Comparison of TM and IP testing conditions were assessed within each for scores on each test in the two conditions. On each of the five language tasks, the Wilcoxon signed ranks test indicated no significant difference on performance between the TM and IP conditions for each participant. Overall acceptance of the TM evaluation in an elderly population was rated at a high level except for one individual.

Conclusion: Telemedicine can improve access to speech and language evaluation services which is relevant to both dementia and other neurological diseases of the elderly. In particular, this specific assessment tool can be used to provide evaluations in under-served rural areas.

Introduction

Alzheimer's disease (AD) is a progressive decline in cognitive abilities associated with formation of senile plaques and neurofibrillary tangles resulting in neuronal loss. The most prominent cognitive impairments in AD in addition to episodic memory loss include impairments in attention/concentration, orientation, judgment, visuospatial skills, executive function, and language. As the disease progresses, cognitive symptoms accelerate in both frequency and severity, resulting in the loss of functional abilities (Kuhn 2002). A gradual decline in speech and language skills is also associated with AD, including difficulties in verbal fluency, naming, comprehension, reading, and writing (Cohen 1999). Language deficits occur in 8%–10% of individuals in the early stages of AD, and become more severe and numerous during its later stages (Groves-Wright et al 2004).

The number of individuals in the US with Alzheimer's disease and/or dementia of some variety is presently estimated to be 4 million individuals, with this number estimated to double by the year 2012 (ADERC 2004). The diagnosis of speech and language disorders, including speech and language impairments of patients with AD, traditionally involves IP verbal interaction between clinician and patient (Duffy et al 1997). With this increasing number of individuals with or at-risk for dementia and the present emphasis on early detection for future treatment interventions, it will become an increasing public health issue to provide effective screening that is feasible to as many elderly individuals at risk. However, due to the traditional IP interaction that is used in cognitive/neuropsychological evaluations, many individuals, particularly those

Correspondence: John Hart Jr.
2200 W. Mockingbird Lane, Dallas, TX
75235, USA
Tel +1 214 905 3007
Fax +1 214 905 3026
Email jhart@utdallas.edu

in rural communities with limited healthcare access or those who depend on caregivers may be deprived of this resource without the development of evaluative tools to overcome these difficulties.

Telemedicine is the emerging use of telecommunication technology to provide medical information and services, particularly those located in rural communities. Telemedicine can be as simple as two health professionals discussing a case over the telephone, or as sophisticated as using satellite technology to broadcast a consultation between providers at two different facilities. In this study, we have applied the use of “tele-cognitive testing” to determine and evaluate language and speech functions that are typically affected in early AD. The following describes our findings and the implications for diagnostic and therapeutic issues in mild Alzheimer’s disease.

Methods

Subjects

Eligible participants ($n = 15$) were identified from the CAVHS Geriatric Memory Clinic and all were male. These 15 were screened using the Mini Mental Status Examination (MMSE) (Folstein et al 1975). Only those individuals with scores > 20 on the MMSE were included as participants. The MMSE identified 11 of the 15 participants who met this inclusion criterion. These 11 participants also received a hearing screen to ensure adequate registration. One participant was excluded due to a severe hearing loss that resulted in his inability to hear and understand speech through the TM equipment, even when wearing his hearing aids (cut-off 20db). Ten of the 15 participants who were screened consented to participate in the TM study. The average age of the 10 participants was 73.9 years (standard deviation [SD] = 3.7) with ages ranging from 68 to 78 years, and the average MMSE score for the group was 26.1 (SD = 1.4).

Stimuli

The following tests were chosen due to sensitivity for the functions impaired in early AD and common use in evaluations of dementia in the elderly. The language battery in this study consisted of the following:

1. Picture Description (auditory response version) (Boston Diagnostic Aphasia Examination [BDAE]) (Goodglass et al 2001).
2. Boston Naming Test (BNT) (Goodglass et al 2001).
3. Token Test (Multilingual Aphasia Examination) (Benton et al 2000).

4. Aural Comprehension of Words and Phrases (Benton et al 2000).
5. Controlled Oral Word Association Test (Benton et al 2000).

Procedures

Each language test had two forms (A and B) that consisted of equivalent items. Each participant was given a battery of language tests using the traditional IP assessment and via satellite TM. Half of the participants were randomly assigned to the IP assessment group followed by assessment via satellite (IP/TM) and the other half received assessment via satellite followed by the IP assessment (TM/IP). The order of the administration of each individual test was also randomized within the two testing situations in an effort to reduce fatigue and sequencing/learning effects.

During the IP assessment, verbal instructions were given to each participant. Any clarification was also given as needed. The researcher–participant interaction consisted of normal gestures and proximity. During the TM assessment, verbal instructions and any clarification were also given; however, the researcher–participant interaction was limited due to distance. The participant was in a room on the third floor of the North Little Rock Veterans Affairs Hospital (NLRVA), while the researcher was in a room located on the first floor at the NLRVA. Any necessary materials, such as the Token Test protocol, were physically present with the participant. A research assistant attended the participant in the videoconferencing room to ensure adequate volume and picture. The instructions, administration, and scoring of the speech and language tests were performed by a licensed speech and language pathologist with extensive experience in performing similar evaluations in demented and healthy aging control patients.

The presence of background noise (primarily heating and air conditioning ventilation) rendered the TM testing situation less than optimal in terms of patient responses. To compensate for this, a tape recorder was placed on the table in front of the participant when the TM equipment was being used for assessment. Recording the participant’s answers enabled the clinician to later verify responses transcribed during testing. In addition, a clinician-generated Token Test palate (Appendix B) was used to aid the participants in the correct re-formation of the tokens while being administered the Token Test via TM.

Telemedicine equipment

The two video-conferencing sites were at the NLRVA. Each site had a television monitor and a microphone. The two

locations were connected using a Polycom 384K broadband line that connected each site to allow for real time interaction. Dial up connection was made using a TCP/IP (Transmission Control Protocol/Internet Protocol) with the h323VOP standard. This connection allowed adequate bandwidth for the assessment of speech and language abilities.

Results

Figures 1A and 1B display the raw scores for each participant across the two testing conditions. All 10 participants completed

the BNT, Token Test, and Picture Description tasks; only nine completed Aural Comprehension of Words and Phrases and Controlled Oral Word Association Test due to time constraints. On visual inspection, the test scores were similar across the TM and the IP testing conditions, in most cases.

Statistical analysis

To assess the similarity of the TM and in person testing conditions, the scores obtained by each participant on each test in the two conditions were compared using the

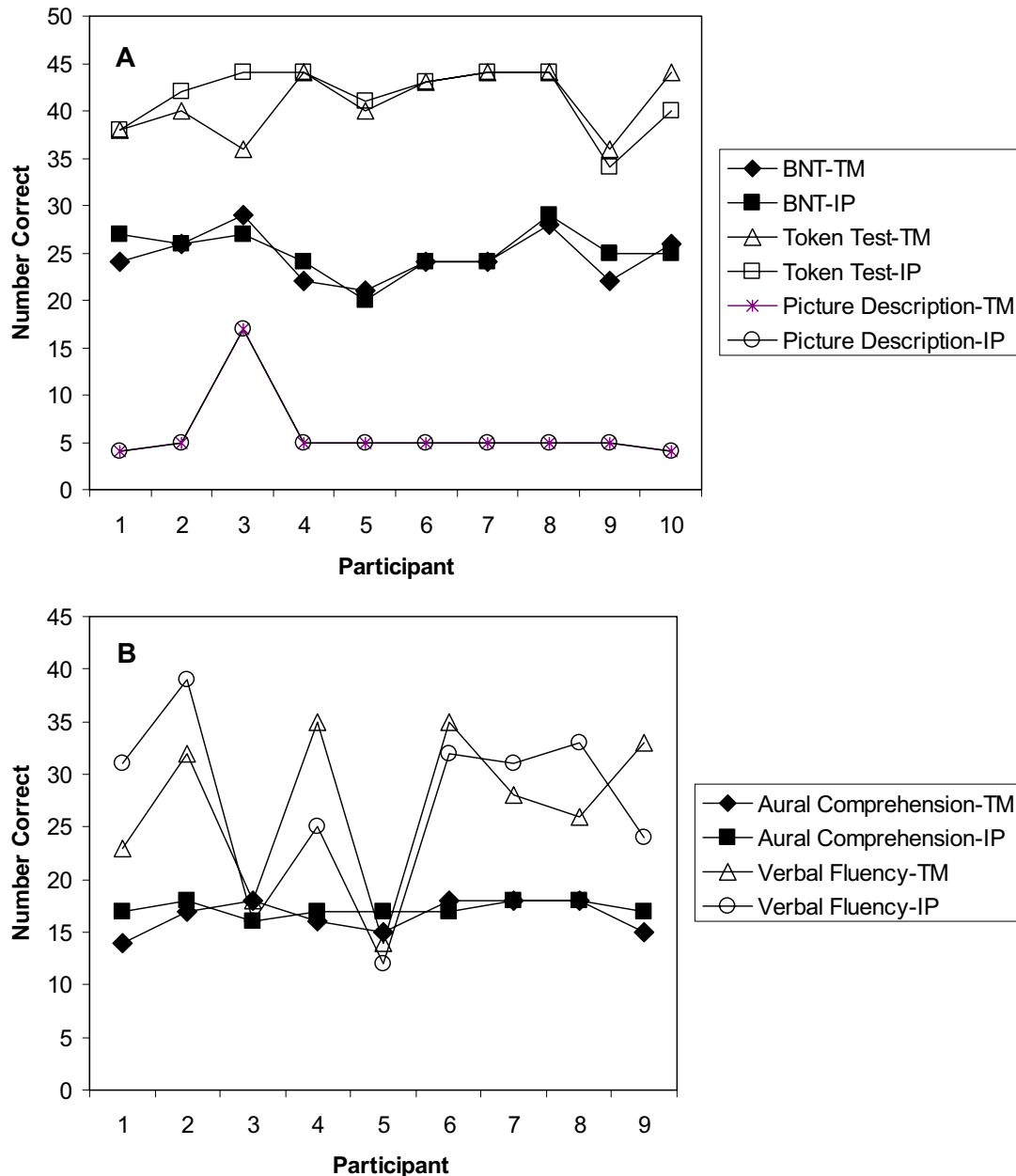


Figure 1 Individual scores on the in-person and telemedicine administration methods. **Panel A** Boston Naming Test, Token Test and Picture Description scores.

Panel B Aural Comprehension and Verbal Fluency scores.

Abbreviations: BNT, Boston Naming Test; IP, in-person; TM, telemedicine.

nonparametric Wilcoxon signed-ranks test. This statistical test was chosen instead of a paired t-test because the assumption of a normal distribution for the data was violated in several instances. A statistical significance value of $p < 0.05$ was chosen a priori, and a 2-tailed test was used. For each comparison, the score obtained in the TM condition was subtracted from the score measured in the IP condition. Therefore, when a rank was negative, the TM score was lower than the IP score; when a rank was positive, the TM score was higher than the IP score; no difference between the TM score and the IP score was considered a tie. Appendix E presents the frequencies and ranks for the participants' language test scores in the two conditions. On each of the five language tasks, the Wilcoxon signed ranks test indicated no significant difference on performance between the TM and IP conditions for each participant ($p > 0.05$). Table 1 presents the z scores and p -values for each test.

Subjects' acceptance

An additional question of interest was subjects' acceptance of the TM method of language assessment. Participants completed a 5-question paper-and-pencil survey. The questions, mean scores and SDs are listed in Table 2. The mean scores indicate a high degree of acceptance with the TM examination. Eight of the ten participants strongly agreed they could understand equally well what the healthcare provider told them via satellite or IP. One participant was neutral to this statement and one strongly disagreed. All participants either agreed or strongly agreed that the TM visit was private enough to ask questions. Six participants agreed or strongly agreed that the TM technology saved time, while the other participants were neutral to this statement. Nine participants agreed or strongly agreed to choose a TM visit again and would rather use TM than travel to the clinic, while one participant was neutral to this statement.

Discussion

The aim of this study was to determine whether patients with mild Alzheimer's disease can be effectively evaluated for their language dysfunction equally well via remote station-

to-station TM as traditional in person evaluation. Thus far, there have been few studies comparing the efficacy of language evaluations in TM, particularly focusing on dementia assessments, due to the belief that speech and language assessments challenge the logistical and technical aspects of tele-assessment more than most clinical medical applications. The present study demonstrates the efficacy of the use of TM in conducting language screening in elderly individuals at risk for dementia.

The present study is consistent with a prior investigation which reported excellent patient satisfaction with TM evaluation (Shores et al 2004). The overall acceptance of the TM equipment in the current study was rated at a high level except for one individual, with this participant expressing frustration understanding the instructions during the language assessment in both testing conditions.

It is also important to note that all testing was administered at one facility for the purposes of assessing the efficacy of performing these language evaluations. No participant saved time in terms of travel and the majority of participants rated the question concerning time saved as neutral. It is the intended purpose of this study to demonstrate the efficacy of language assessments in the elderly by TM and then to be able to extend this to distant station-to-station evaluations of patients in underserved regions where access to relevant personnel is minimal. In addition, this type of study raises the possibility that one evaluator could administer one set of instructions to a number of patients located at multiple stations, with that evaluator potentially recording all responses or staff at the distant stations recording/scoring responses.

These findings have particular clinical relevance given that the tests chosen have been shown to be sensitive to detecting the language impairments typical of early dementia (Croisile et al 1996; Groves-Wright et al 2004). Studies have shown that the COWAT and Boston Naming Test are particularly sensitive to detecting early language dysfunction in Alzheimer's disease (Groves-Wright et al 2004). Other tests that are found to be typically disrupted early in AD—written description of a complex scene—are difficult to assess in the

Table 1 Comparison statistics of the in-person and telemedicine methods of test administration

	Boston Naming Test	Token Test	Picture Descriptions*	Aural Comprehension of Words and Phrases	Controlled Oral Word Association
z statistic	-0.171	-1.084	0.000	-1.200	-1.316
p -value (2-tailed)	0.864	0.279	1.000	0.230	0.188

Note: *Cookie theft and picnic scene.

Table 2 Telemedicine participant survey

Please rate the following statements about your telemedicine experience: (N = 9)	*Rating Mean (SD)
1. I understand what the healthcare-provider told me as well as if it had been in person.	4.3 (1.4)
2. The telemedicine visit was private enough for me to ask the questions I wanted to ask.	4.8 (0.4)
3. The telemedicine technology saved me time.	4.1 (0.9)
4. I would choose to have a telemedicine visit again.	4.4 (0.7)
5. I would rather use telemedicine than travel to the clinic.	4.4 (1.0)

Note: *5, strongly agree; 4, agree; 3, neutral; 2, disagree; 1, strongly disagree.

Abbreviations: SD, standard deviation.

TM situation. Telemedicine equipment and technology advances are now allowing medical specialists to offer improved access to patients who otherwise may not receive services or receive delayed services due to physical barriers such as distance or lack of transportation. The capability to use language tests that are sensitive to diagnosing early dementia allows for the expansion of TM technology usage to increase the number of individuals who can be screened for incipient dementia.

Although the findings of the present study indicate a consistent relationship between IP and TM assessments, several limitations need to be considered. First, the study included only a small number of Caucasian men. While on the surface we would not expect race or gender differences in this type of comparison, it is possible that different results would be found for a larger number of adults including women and other ethnic/racial groups. This concern is somewhat reduced since the findings of the present study were similar to studies examining the use of TM assessment in dementia populations where various ethnic and racial groups were represented (eg, Duffy et al 1997; Shores et al 2004). A second potential limitation of the study is that only mild Alzheimer's disease patients were tested. As cognitive and memory deficits progress, speech and language deficits become more evident and disabling (Groves-Wright et al 2004). It is also suggested that some patients with severe cognitive and memory impairments may have difficulty grasping the nature of the interactive process over a television monitor (Duffy et al 1997). Therefore, it is possible that TM evaluations would not be reliable in assessing individuals with severe cognitive or memory impairments. Although patients with severe cognitive impairments may have problems with interactive television, some obstacles such as technology confusion might be overcome with explicit explanation of the process and how the equipment works, as well as clear and simple directions from the tester.

The results from this study add new evidence to the literature about the use of TM in speech-language evaluations. This research demonstrates that assessment of language skills in mild AD patients can be accomplished via TM and achieve results that are not significantly different from an IP assessment. Additional research should investigate its efficacy in other populations, such as individuals with aphasia, dysarthria, and/or apraxia of speech in order to expand the generalizability of these findings.

References

- [ADERC] Alzheimer's Disease Education and Referral Center. 2004. Alzheimer's disease medication fact sheet [online]. Accessed on April 2, 2004. URL: <http://www.alzheimers.org>.
- [ATA] American Telemedicine Association. 1999. Telemedicine: a brief overview [online]. Accessed April 3, 2004. URL: <http://www.atmeda.org>.
- Benton A, Hamsher K, Sivan A. 2000. Multilingual aphasia examination. 3rd ed. Iowa City, IA: AJA Associates.
- Cohen E. 1999. Alzheimer's disease: prevention, intervention, and treatment. Lincolnwood, IL: NTC/Contemporary Publishing Group, Inc.
- Croisile B, Ska B, Brandant MJ, et al. 1996. Comparative study of oral and written picture description in patients with Alzheimer's disease. *Brain Lang*, 53:1-19.
- Duffy JR, Werven GW, Aronson AE. 1997. Telemedicine and the diagnosis of speech and language disorders. *Mayo Clinic Proceedings*, 72:1116-22.
- Folstein MF, Folstein SE, McHugh PR. 1975. Mini mental status: a practical method for grading the cognitive state of patients for the clinician. *J Psych Res*, 12:189-98.
- Goodglass H, Kaplan E, Barresi B. 2001. Boston Diagnostic Aphasia Examination, 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins.
- Groves-Wright K, Neils-Strunjas J, Burnett R, et al. 2004. A comparison of verbal and written language in Alzheimer's disease. *J Commun Disord*, 37:109-30.
- Kuhn D. 2002. Family caregiver alliance. California Department of Mental Health.
- Shores MM, Ryan-Dykes P, Williams RM, et al. 2004. Identifying undiagnosed dementia in residential care veterans: comparing telemedicine to in-person clinical examination. *Int J Geriatr Psych*, 19:101-8.