

Effects of Oral Physostigmine in Alzheimer's Disease

Yaakov Stern, PhD,*† Mary Sano, PhD,* and Richard Mayeux, MD*†

Previous studies of oral physostigmine in the treatment of Alzheimer's disease have: (1) assumed physostigmine is effective only in mildly affected patients; (2) relied on an initial "dose-finding" phase to determine the most effective dose and excluded nonresponders; and (3) primarily assessed memory. We examined the response of 22 patients to six different daily dosages of oral physostigmine, using selective reminding tests that were administered twice daily. Nine patients had a "best" dose/day (mode = 13 mg/day), which was used in a subsequent double-blind crossover study. The other 13 were given the highest tolerated dose. The selective reminding test and a full neuropsychological battery were given during the drug and placebo periods. As a group, the 22 patients improved significantly on the Wechsler Adult Intelligence Scale-Revised Digit Symbol subtest and a shape cancellation task ($p < 0.05$). Nine patients showed improved performance on the selective reminding test during physostigmine treatment, and 9 showed no response; 4 patients performed better during placebo treatment. Dose finding did not help in predicting response in the crossover study; only 2 of the 9 who showed improvement had a best dose. Dementia severity did not predict crossover response. This suggests that: (1) physostigmine as administered had no pronounced effect on memory in Alzheimer's disease; (2) oral physostigmine produces no greater benefits on memory in mildly than in moderately demented patients; (3) response in a dose-finding phase does not predict response in double-blind crossover; and (4) Digit Symbol and cancellation tasks may be more sensitive than memory tests to the effects of oral physostigmine.

Stern Y, Sano M, Mayeux R: Effects of oral physostigmine in Alzheimer's disease. *Ann Neurol* 22:306-310, 1987

Physostigmine has been considered a promising agent for the treatment of Alzheimer's disease (AD). Investigators have reported a small, but statistically significant, improvement in patient performance when oral physostigmine (OP) was administered [16, 22]. However, negative findings have also been reported [5]. We attempted to address three major procedural issues in the study of this medication:

1. The effective dose of physostigmine has been defined by an inverted U-shaped function, such that lower or higher dosages are not as effective as those at the median level [3, 7, 18, 22]. Many studies have applied this observation by using a titration or dose-finding phase to determine the optimal dosage for use during a crossover with placebo (e.g. [22]). Patients who did not show the expected dose-response curve have often been excluded from crossover trials [16, 22]. This procedure could bias the results of the crossover study, increasing the probability of finding some drug effect. We investigated the presence of an inverted U-shaped function in patients' response to OP. However, all patients were included in a subsequent

double-blind crossover study, comparing performance on OP to performance on placebo.

2. It could be assumed that physostigmine is more effective in patients with less severe dementia [22]. This is based on its presumed mode of action: it inhibits the breakdown of synaptically released acetylcholine as opposed to increasing the amount of acetylcholine available. We assessed this concept by including patients who were variably affected by dementia and comparing their relative response to OP.

3. Most studies have concentrated on assessing physostigmine's effect on measures of memory, but other cognitive functions might also benefit [8, 17]. We included tests of a wide range of intellectual function to investigate this possibility.

Methods

Subjects

Twenty-two patients with probable AD were selected and gave informed consent for this study. Average patient age was 67.1 years (± 8.4); age at onset, 64.0 (± 8.5); education, 15.0 (± 3.7). The average score on the modified Mini-Mental State Examination (mMMS) was 41.0 (± 8.3). In our

From the Departments of *Neurology and †Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY.

Received Sept 16, 1986, and in revised form Dec 9, 1986, and Jan 5, 1987. Accepted for publication Jan 5, 1987.

Address correspondence to Dr Stern, Neurological Institute, 710 W 168th St, New York, NY 10032.

experience, the average mMMS score in nondemented elderly people is 52.3 (± 4.3). All patients were required to meet the criteria for primary degenerative dementia of the American Psychiatric Association's Diagnostic and Statistical Manual [1] as well as NINCDS-ADRDA criteria for "probable" AD [15]. In addition, patients were screened for their ability to recall a minimum of three words on several trials of the selective reminding test, described below. The overall reaction of patients to testing was also assessed because it can be stressful. Those who could not tolerate this procedure were not entered into the study. Twelve patients were eliminated for these reasons.

Selective Reminding Test

The primary measure of memory in this study was a selective reminding test (SRT) [4]. In this list-learning task, the subject attempts to recall a list of 12 words after they have been read to him. After each recall attempt, the words not recalled are repeated and the subject is asked to attempt again to recall the entire list. This procedure is repeated for a total of twelve recall trials.

Three measures of SRT performance were derived: (1) total recall—the total of all words recalled on all trials (maximum score = 144, 12 words \times 12 trials); (2) long-term retrieval—words recalled on two successive trials (without an intervening reminder) were considered to have been retrieved from long-term storage on these two and all following trials; and (3) intrusions—wrong words were corrected the first time the patient made the error but were scored as intrusion errors for each use thereafter.

Because this study involved repeated use of the SRT, 17 different versions were used. Each consisted of words derived from unrelated categories. Ten versions have been described previously, and their relative equivalence has been established [12]. The remaining forms were constructed to contain words of equivalent frequency of usage [13] and were tested to ensure that they were comparable in repeated-measure reliability studies.

Procedure

DOSE FINDING. After an initial day of baseline testing, OP dosage was increased daily for 5 days. The following total daily doses were all administered on a dosage schedule of every 2 hours, starting at 9:00 A.M.: 2 mg ($\times 6$), 2.5 mg ($\times 6$), 3 mg ($\times 5$), 3.5 mg ($\times 4$), and 4 mg ($\times 4$). Patients were monitored daily for potential side effects. The SRT was administered twice daily at standard times, one-half hour after administration of OP. The SRT performance of each patient was assessed to determine if the means of the three SRT performance measures were consistently better on one particular daily dosage. If a patient had a "best" daily dosage, then that was used for the subsequent double-blind crossover study. Otherwise, the highest dose tolerated without side effects was used.

DOUBLE-BLIND CROSSOVER STUDY. All patients from the dose-finding phase participated in the double-blind crossover study. After a 1-day washout period, patients received either OP or placebo for 3 days. This was followed by a 1-day washout period, and then patients were crossed over to the

alternate drug or placebo for the second 3-day period. The SRT was again administered twice daily. On the third day of each crossover period, the following battery of neuropsychological tests was administered: the Wechsler Adult Intelligence Scale-Revised (WAIS-R) Digit Symbol subtest [25], Wechsler Memory Scale Logical Memory (immediate and 10-minute delay), Visual Recall and Associate Learning subtests [24], Rosen Drawing Test [19], mMMS [9, 14], controlled word association test [2], and category-naming test [10]. In addition, two types of cancellation tasks were administered. One involved the detection of a specific shape within a shape array while another used letter triads as a target for detection [20]. This test battery was chosen to assess a range of functions that are affected in probable AD. Alternate versions of the Wechsler Memory, controlled word association, and category-naming tests were used; the order of administration of alternate forms was held constant because drug/placebo order was randomized across patients. Alternate versions of the other tests were not used because, in our experience, performance is not affected by repeated testing in patients with probable AD.

Results

Adverse Effects

The most common side effects were dizziness and nausea. These were sufficient to terminate the dose-finding phase, before the maximum dosage was reached, in 8 patients. In all instances, reducing the dosage eliminated these effects.

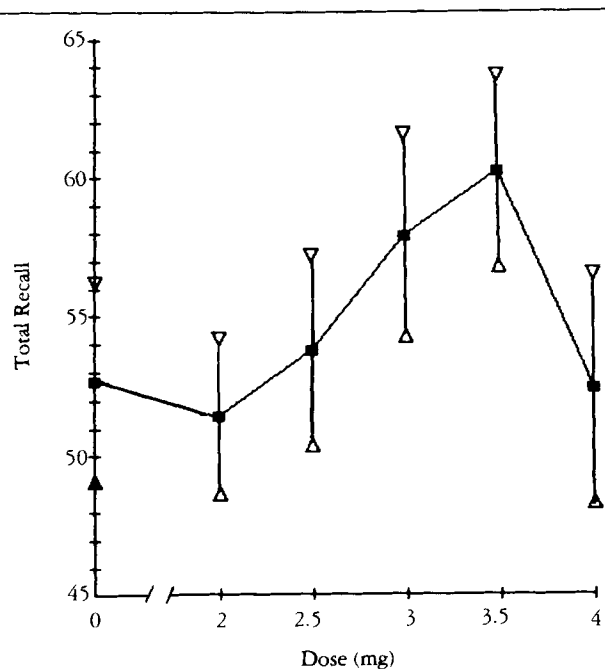
Dose-finding Phase

Correlations between the scores obtained from the two SRTs administered at each dosage during the dose-finding phase were calculated to investigate the stability of SRT performance. At baseline (before drug administration), the correlation of total score and of long-term recall on the two tests was significant ($r = 0.91$ for both, $p < 0.01$), but the correlation of intrusion scores did not reach significance. On other daily dosages, all SRT values correlated significantly.

In grouped data, mean SRT performance at different dosages of OP appeared to have the inverted U-shaped function, with maximal performance at 3 to 3.5 mg every 2 hours (Figure). Each patient's performance was evaluated by two independent raters to determine whether this U-shaped function was present. Nine patients showed the pattern. In the remaining 13, there was no regular dose-response relationship. In 3 of these 13 patients there was still a particular daily dosage that yielded the best SRT performance. In the other 10, there was no best dosage, and the highest dose tolerated without side effects was used for the subsequent double-blind crossover trial.

Double-Blind Crossover Study

SRT MEASURES. Possible patient improvement on SRT measures during the drug period was investigated



Relationship of total recall scores on the selective reminding test to individual doses received during the dose-finding phase. Triangles depict standard errors of the mean.

Table 1. Selective Reminding Test Scores During the Drug and Placebo Phases of the Double-Blind Crossover Study in 22 Patients

Selective Reminding Test	Placebo		Drug	
	Mean	SD	Mean	SD
Total recall	56.6	14.6	57.6	14.6
Long-term retrieval	22.2	12.3	22.5	14.4
Intrusions	5.8	6.9	6.1	6.1

using Student's *t* tests, comparing the difference between SRT scores derived from the first and second crossover periods in the two drug-placebo sequence groups (i.e., those who received drug first and those who received placebo first) [11, 23]. Paired *t* tests for evaluating changes in performance between the drug and placebo trials, as well as repeated-measure analysis of variance (ANOVAs), were also employed. There were no significant differences between the two conditions (Table 1). Tests for carryover effects did not show significance. To control for possible carryover effects, the same analyses were repeated using only SRT performance on day 3 of drug and placebo conditions; again, there were no significant differences between the conditions. No difference in performance as a function of the order of administration of drug and placebo was noted. To investigate the relative efficiency of OP treatment for patients with mild versus moderate dementia, patients were stratified into

Table 2. Neuropsychological Test Scores in the Drug and Placebo Phases of the Double-Blind Crossover Study in 22 Patients

Test	Placebo		Drug	
	Mean	SD	Mean	SD
mMMS	41.0	8.3	40.8	8.3
WMS (immediate recall)				
Logical Memory	3.3	1.2	3.8	1.4
Visual Memory	3.3	3.1	3.3	2.3
Paired Associates	6.7	2.4	7.3	2.2
WMS (delayed recall)				
Logical Memory	0.7	0.9	0.8	0.8
Visual Memory	0.4	1.0	1.0	1.6
Paired Associates	2.2	1.1	1.6	0.9
WAIS-R Digit Symbol (age-scaled score)	6.0	3.8	6.6	3.7 ^a
CFL (mean)	11.8	4.3	10.3	4.9
Category naming (mean)	11.5	4.9	11.6	4.4
Rosen Drawing Test	10.5	3.1	10.3	2.7
Cancelations (sec)				
Letters	123.2	39.6	131.2	68.4
Shapes	116.5	109.0	103.0	65.4 ^a

^aDrug-placebo means were significantly different at $p < 0.05$.

mMMS = modified mini-mental state; WAIS-R = Wechsler Adult Intelligence Scale-Revised; CFL = Controlled Word Association; WMS = Wechsler Memory Scale.

two groups, according to their performance on the mMMS. There was no difference in the relative improvement between drug and placebo conditions in these two groups.

The performance of individual patients was evaluated by two independent raters who were blinded to drug status, to determine whether all three SRT measures were consistently superior in one of the two conditions. Nine patients showed improved performance with OP treatment, 9 had no consistent difference in performance between the two conditions, and 4 had better performance on placebo.

NEUROPSYCHOLOGICAL BATTERY. Changes in performance on the neuropsychological battery were investigated using ANOVA for repeated measures (Table 2). Performance on the WAIS-R Digit Symbol subtest and the shape cancelation test was significantly better during the drug condition. No differences were noted in performance on other tests.

RELATIONSHIP BETWEEN DOSE-FINDING AND DOUBLE-BLIND CROSSOVER PERFORMANCE. The utility of the dose-finding phase for predicting improvement during double-blind crossover was investigated. Only 3 of the patients who showed a response to OP had been judged to have a U-shaped dose-response pattern, and only 1 additional patient had had a best daily dosage during dose finding.

DISEASE SEVERITY AND OP RESPONSE. Baseline neuropsychological values of responders and nonresponders were compared to determine if these two groups differed in the severity of their dementia. Mean values for all neuropsychological test scores were comparable in the two groups.

Discussion

This study suggests that OP may slightly improve intellectual function in patients with probable AD. On the primary dependent measure used in this study, the SRT, there was no significant difference between performance on placebo and drug. Although the number of patients in this study was relatively small, the number was sufficient to detect a change as small as one-third a standard deviation (or approximately 5%) in SRT scores at an alpha level of 0.05 with a power of 80% [6]. Several patients were classified as OP responders on the basis of consistent improvement on SRT measures, but changes in scores were often slight and it is not clear that they imply improvement in "real life" function.

WAIS-R Digit Symbol performance and performance on a shape cancellation task were significantly better on OP than placebo. Both of these tests require speeded performance and tap aspects of attention and visuospatial ability. Poor performance on these tasks is not indicative of a specific cognitive or anatomic deficit. Still, these findings suggest that it is important to assess functions other than memory when evaluating the efficacy of OP. Improvement in constructional tasks and on a scale designed to assess the severity of probable AD have also been reported with physostigmine [16, 17], and attention improved with cholinomimetics [18].

Other investigators have relied on a dose-finding phase to determine an optimal dosage of physostigmine, and in some studies, patients who did not have an "optimal" dose were not included in subsequent placebo-controlled investigations (e.g., [22]). This selective elimination of patients clearly can bias subsequent studies. We included patients in the double-blind crossover phase regardless of their performance on a dose-finding phase and found that the dose-finding phase was not predictive of later performance. We also used higher doses of OP than have been reported previously. It might be more useful simply to determine the highest dose that each patient can tolerate without side effects and then use that dosage for placebo-controlled studies.

One source of variability in assessing patients' optimal drug doses might have been differences in drug absorption after administration. All patients were hospitalized throughout the entire study and ate at standard times each day; dietary effects were probably minimal. However, central nervous system effects of

OP can vary between patients [22] and this was not assessed.

It has been assumed that physostigmine does not increase the synthesis of acetylcholine but instead maintains whatever acetylcholine is present in the synaptic cleft. Given this presumed mode of action, the drug should be most effective in patients with more intact cholinergic systems, that is, patients with relatively mild dementia. We studied patients suffering from varying degrees of dementia and found no difference in OP response between mildly and moderately affected patients. In fact, a few patients who began with baseline neuropsychological scores in the lower range showed the most marked improvement. This probably cannot be attributed to a "ceiling" effect in the patients with more intact mental functioning, since all patients were defective on the SRT in comparison to normative data derived from age-matched healthy control subjects. Still, the more demented patients did have more room for improvement on these measures. Another study, using intramuscularly administered physostigmine, reported greater improvement in more severely demented patients [21].

This study suggests that: (1) OP as administered has no substantial effect on memory; (2) OP has no greater beneficial effect on memory in mildly than in moderately demented patients; (3) response in a dose-finding phase does not predict response in double-blind crossover; and (4) Digit Symbol or cancellation tasks may be more sensitive than memory tests to the effects of OP.

Supported by the Charles S. Robertson Memorial Gift for Alzheimer's Disease Research and by Federal Grants AGO2802 and AGO5433. It was completed in a Clinical Research Center, and data were analyzed on a CLINFO system (both RR00645).

We thank the nurses and staff of the Columbia University Clinical Research Center for their dedicated care of the patients in this study, and Mr Donald P. McMahon for his assistance with data management. Physostigmine and placebo were generously provided by Forest Pharmaceutical, Maryland Heights, MO.

Presented in part at the 38th Annual Meeting of the American Neurological Association, Boston, MA, October 1986.

References

1. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, ed 3. Washington DC, APA, 1980
2. Benton A: FAS Test. In Spreen O, Benton A (eds): Neurosensory Center Comprehensive Examination for Aphasia. Victoria, British Columbia, University of Victoria, 1967
3. Berger PA, Davis KL, Hollister LE: Cholinomimetics in mania, schizophrenia, and memory disorders. In Barbeau A, Growden JH, Wurtman RF (eds): Nutrition and the Brain, Vol 15. New York, Raven, 1979, pp 425-441
4. Buschke H, Fuld PA: Evaluating storage, retention, and retrieval in disordered memory and learning. *Neurology* 24:1019-1025, 1974
5. Caltagirone C, Gainotti C, Masullo O: Oral administration of

- chronic physostigmine does not improve cognitive or mnemonic performances in Alzheimer's presenile dementia. *Int J Neurosci* 16:247-249, 1982
6. Cohen J: *Statistical Power Analysis for the Behavioral Sciences*. New York, Academic, 1977
 7. Davis KL: Multiple dose intravenous physostigmine in Alzheimer's disease. International study group on the pharmacology of memory disorders associated with aging. Zurich, Switzerland, April 1981
 8. Davous P, Lamour Y: Bethanacol decreases reaction time in senile dementia of the Alzheimer type. *J Neurol Neurosurg Psychiatry* 48:1297-1299, 1985
 9. Folstein MF, Folstein SE, McHugh PR: "Mini-mental state," a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12:189-198, 1975
 10. Goodglass H, Kaplan E: *The Assessment of Aphasia and Related Disorders*. Philadelphia, Lea and Febiger, 1972
 11. Grizzle JE: The two-period change-over design and its use in clinical trials. *Biometrics* 21:467-480, 1965
 12. Hannay HJ, Levin HS: The selective reminding test: an examination of the equivalence of four forms. *J Clin Exp Neuropsychol* 7:251-264, 1985
 13. Lorge I, Thorndike EL: *A Semantic Count of English Words*. New York, Columbia University Press, 1938
 14. Mayeux R, Stern Y, Rosen J, Leventhal J: Depression, intellectual impairment and Parkinson's disease. *Neurology* 31:645-650, 1981
 15. McKhann G, Drachman D, Folstein M, et al: Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 34:939-944, 1984
 16. Mohs RC, Davis BM, Johns CA, et al: Oral physostigmine treatment of patients with Alzheimer's disease. *Am J Psychiatr* 142:28-33, 1985
 17. Muramoto O, Sugishita M, Kaxuya A, et al: Cholinergic system and constructional praxis: a further study of physostigmine in Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 47:484-489, 1984
 18. Peters BH, Levin HS: Memory enhancement after physostigmine treatment in the amnesic syndrome. *Arch Neurol* 34:215-219, 1977
 19. Rosen W: *The Rosen Drawing Test*. Bronx, NY, Veterans Administration Center, 1981
 20. Sano M, Rosen W, Mayeux R: Attention deficits in Alzheimer's disease. Presented at the 92nd Annual Meeting of the American Psychological Association, Toronto, Ontario, August, 1984
 21. Schwartz AS, Kohlstaedt EV: Physostigmine effects in Alzheimer's disease: relationship to dementia severity. *Life Sci* 38:1021-1028, 1986
 22. Thal LJ, Fuld PA, Masur DM, Sharpless NS: Oral physostigmine and lecithin improve memory in Alzheimer's disease. *Ann Neurol* 13:491-496, 1983
 23. Wallenstein S, Fisher AC: The analysis of two-period repeated measurements crossover designs with application to clinical trials. *Biometrics* 33:261-269, 1977
 24. Wechsler D: *A standardized memory scale for clinical use*. *J Psychol* 19:87-95, 1945
 25. Wechsler D: *Wechsler Adult Intelligence Scale-Revised*. New York, Psychological Corporation, 1981