*Movement Disorders*
Vol. 15, No. 3, 2000, pp. 467–473


# Early Detection of Probable Idiopathic Parkinson's Disease: I. Development of a Diagnostic Test Battery

*Erwin B. Montgomery, Jr., MD, ‡William C. Koller, MD, PhD, †Theodora J. K. LaMantia, §Mary C. Newman, PhD, ‖Elizabeth Swanson-Hyland, PhD, §Alfred W. Kaszniak, PhD, and ‡Kelly Lyons, PhD

**Departments of Neurology and †Neuroscience, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, Ohio, U.S.A.; ‡National Parkinson Foundation, University of Miami Medical Center, Miami, Florida, U.S.A.; §Department of Psychology, The University of Arizona, Tucson, Arizona, U.S.A.; and ‖Responsive Health Care, Overland Park, Kansas, U.S.A.*

**Summary:** We developed a test battery as an inexpensive and objective aid for the early diagnosis of idiopathic Parkinson's disease (iPD) and its differential diagnoses. The test battery incorporates tests of motor function, olfaction, and mood. In the motor task, a wrist flexion-and-extension task to different targets, movement velocities were recorded. Olfaction was tested with the University of Pennsylvania Smell Identification Test. Mood was assessed with the Beck Depression Inventory. An initial regression model was developed from the results of 19 normal control subjects and 18 patients with early, mild, probable iPD. Prospective application to an independent validation set of 122 normal control subjects and 103 patients resulted in an 88% specificity rate and 69% sensitivity rate, with an area under the Receiver Operator Characteristic curve of 0.87. **Key Words:** Parkinson's disease—Diagnosis—Movement—Olfaction—Mood.

Idiopathic Parkinson's disease (iPD) is a common chronic and disabling condition with an enormous impact on quality of life and the cost of health care. Both the prevalence and costs of iPD will increase markedly in the next decade as the "baby boom" generation enters the age of greatest risk. This increase will occur unless new treatments that slow or stop disease progression are developed and coupled with improved diagnostic measures. Although the effectiveness of selegiline to slow the progression of iPD is debated,[1–4] other agents, such as nerve growth factors[5] and gangliosides,[6] offer hope as possible protective measures.

The clinical diagnosis of iPD is difficult, particularly early in the disease when the symptoms and signs may be subtle. Yet, this early stage is the critical time for diagnoses if physicians hope to intervene in the progression of the disease. In addition, many patients wait too long after recognizing that something is wrong before seeking help. Changes in health care delivery, with disincentives for referral to neurologists, will place a greater burden on the primary care physician to make the diagnosis, particularly early in the course of the disease.

We describe here the initial development of a battery of tests for iPD (PD Battery) that may be useful in the early diagnosis of Parkinson's disease.

## METHODS

### Criteria for Developing the Test Battery

Wasson et al.[7] described a set of guidelines for the development of clinical tests. These include a clear definition of outcome and predictive findings, blinded assessment of outcome and clinical prediction, applicability and accuracy of clinical prediction rules, clinical prediction rules that have effects on patient care, and mathematical techniques to ensure that the results can be generalized to the whole population. We followed these guidelines in developing the PD Battery.

The PD Battery incorporates tests of motor function, olfaction, and mood (depression). Tests of motor performance were chosen because of the well-recognized symptoms and signs of bradykinesia. The tests of motor function using wrist flexion and extension movements to two types of targets were derived from neurophysiologi-

Received April 6, 1999; revisions received July 13, 1999, September 15, 1999, and March 1, 2000. Accepted April 5, 2000.

Address correspondence and reprint requests to Erwin B. Montgomery, Jr., MD, Department of Neurology, Desk S90, Cleveland Clinic Foundation, 9500 Euclid Ave., Cleveland, OH 44195, U.S.A.



*See also page 474*

cal studies in nonhuman primates before and after induction of parkinsonism with n-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) and in humans with probable iPD.[8,9]

The wrist task is complicated by the fact that bradykinesia is one of several inclusion criteria for patients with probable iPD. In this sense, the wrist task results may not be independent of the selection criteria. This potential lack of independence raises the concern that the wrist task would be abnormal in the patient group because the patients were selected to be slow in their movements. However, bradykinesia alone was only one of the criteria we used and was not a necessary or sufficient criterion for diagnosis. The diagnosis also depends on the presence of tremor, rigidity, and postural abnormalities.

Olfaction and mood were included because these characteristics are abnormal in a large percentage of patients with iPD.[10–15] Symptoms of olfactory dysfunction and depression often antedate the diagnosis of iPD.[16–19] Recently, several studies have shown that olfactory testing may distinguish patients with iPD from patients with progressive supranuclear palsy (PSP).[20] This observation increases the likelihood that the PD Battery may help differentiate patients with iPD from patients with PSP. Olfaction does not distinguish iPD from Alzheimer's disease.[14] However, the combination of olfaction, mood, and motor testing may do so.

Clinical predictive tests can be made stronger by combining different subtests, provided that the subtests are independent; that is, that they do not measure the same feature. It is reasonable to assume that the wrist movement task, the olfaction test, and mood assessment evaluate different and independent features of iPD. Thus, their combination should improve the predictive value of the test battery.

### Selection of Patients and Control Subjects

As yet, there is no antemortem method for establishing the diagnosis of iPD. Thus, to be accurate, patients studied here, and in most published studies, can be said only to have Parkinson's syndrome (PS) rather than iPD. This difficulty in differentiating iPD and other diagnoses in PS can be confusing because the terms seem to be used interchangeably. The selection criteria used in this study were designed to maximize the probability of iPD and to minimize other disorders in PS. We use the term "probable iPD" to reflect the lack of independent objective confirmatory data.

Patients with probable iPD were defined as having three of the following four symptoms or signs: (1) resting tremor of 4 to 6 Hz that attenuated with movement; (2) slowing of movement, the absence of associative movements, or both; (3) rigidity as measured by increased resistance to passive movement; and (4) flexed posture, impaired posture-righting reflexes, or both.

Patients were excluded if they met any one of the following criteria: (1) a history of exposure, within the previous 6 months, to drugs that either block dopamine receptors or deplete dopamine stores; (2) a lack of response to levodopa therapy; (3) clinically significant weakness; (4) changes in tendon reflexes or the presence of pathologic reflexes, such as a Babinski sign; (5) ataxia; (6) dementia; (7) impairment of vertical gaze; or (8) marked autonomic nervous system abnormalities, such as orthostatic hypotension and urinary bladder dysfunction. Although these exclusion criteria do not eliminate the possibility of including patients with conditions other than iPD, these criteria decrease the probability of such inclusion.

All patients were Hoehn and Yahr stage 2 or 2.5 and were mildly and bilaterally affected. Normal control subjects were required to meet the same exclusion criteria as described above for patients with probable iPD. In addition, control subjects and patients were excluded if they had a medical history that could account for diminished olfaction from other causes. For example, potential participants were excluded if they had a history of significant head injury, smoking, allergic rhinitis, or nasally administered medications.

Some newly diagnosed patients were taking medications, which could introduce some selection bias. Patients not yet receiving treatment may differ in the degree of pathology compared with those on treatment. However, this patient selection was deliberate. To develop a diagnostic battery, we required patients in whom we were confident of the diagnosis. Thus, some patients with a history of responsiveness to levodopa therapy were included. Patients with early probable iPD (that is, those not yet under treatment) were included to increase the probability that the test battery would be generalizable to the population of concern: patients with very early and subtle iPD.

Patients were recruited from the Movement Disorders Clinics of the Departments of Neurology of the Cleveland Clinic Foundation, the University of Arizona College of Medicine, and the University of Kansas Medical Center. Control subjects were recruited from friends of patients and members of local Parkinson lay organizations. All participants gave written informed consent, and the protocol was approval by the institutional review boards of the three centers. The protocol was consistent with the Declaration of Helsinki. Participants received the complete PD Battery at one of the participating centers.

### Blinded Assessment of the Outcome and Clinical Prediction

Patients were selected by the principal investigators (E.B.M. and W.C.K.) before application of the test battery and therefore, without knowledge of the test battery results. Further, the scoring of the PD Battery is objective and less susceptible to laboratory personnel bias.

### Test Battery Development

In the present study, rapid wrist flexion and extension movements were made to one of two types of targets in response to an auditory "go" signal (Fig. 1). Participants were seated comfortably in front of the instruction panel and manipulandum. The range of motion was divided into seven segments, with the end segments 5° wide and the interior segments 12° wide. The instruction panel consisted of two rows of seven light-emitting diodes (LEDs), each corresponding to one of the seven segments of the range of motion. The top row of LEDs represented targets, and the bottom row represented cursor LEDs indicating present hand position.

Movements to the end segments were limited by mechanical stops. Movement to these targets constituted the bounded tasks. Thus, participants did not have to stop the movement carefully, except to prevent bouncing back out of the target. Movement to the interior targets 2 and 6 required the participants to stop the movement and constituted the unbounded tasks. Flexion and extension and bounded and unbounded trials were presented in a random order. Wrist position was monitored at 125 samples per second. Reaction times and movement velocities (degrees per second) were measured.

Trials began with illuminating an end-target LED (target 1 for extension or target 7 for flexion movements). The participant then moved the hand to align the cursor LED with the target LED. The initial end-target LED was then extinguished, and a target LED at the opposite end illuminated, indicating a new target. For bounded tasks, either target 7 for extension or target 1 for flexion movements was illuminated. For unbounded tasks, the next-to-the-end LED (target 6 for extension or target 2 for flexion movements) was illuminated. After a random time, an auditory "go" signal cued the participant to make a rapid wrist movement to the final target.

All participants used their dominant hand, even though the nondominant hand may have been more affected in the patients with probable iPD. The dominant hand was selected because we anticipated that motor performance for the nondominant hand would be slower, and perhaps more variable, than the dominant hand, particularly for the control subjects. Therefore, mixing the use of dominant and nondominant hands would have increased the variance, thereby decreasing statistical power. Further, selecting the most affected side for the patients with probable iPD would have increased the magnitude of the effect, thereby increasing statistical power. However, choosing the most affected side would have required a clinical assessment, which was not how we envisioned the PD Battery being used.

Participants were allowed to practice until proficient, as evidenced by reaching a plateau in performance. Data from 10 to 20 trials of each of the four tasks (bounded and unbounded, flexion and extension) were collected and analyzed. The number of trials collected for each task differed because the tasks were presented in a random order. Repeatedly, during data acquisition, participants were encouraged to react and move as quickly as possible. Participants were also allowed to rest if they became fatigued.

### Measurement of Olfactory Function

Olfactory function was measured with the University of Pennsylvania Smell Identification Test (UPSIT, Sensonics, Inc, Haddon Heights, NJ). The UPSIT consists of 40 standardized, encapsulated odors, one per page. The participant is asked to identify each odor from among

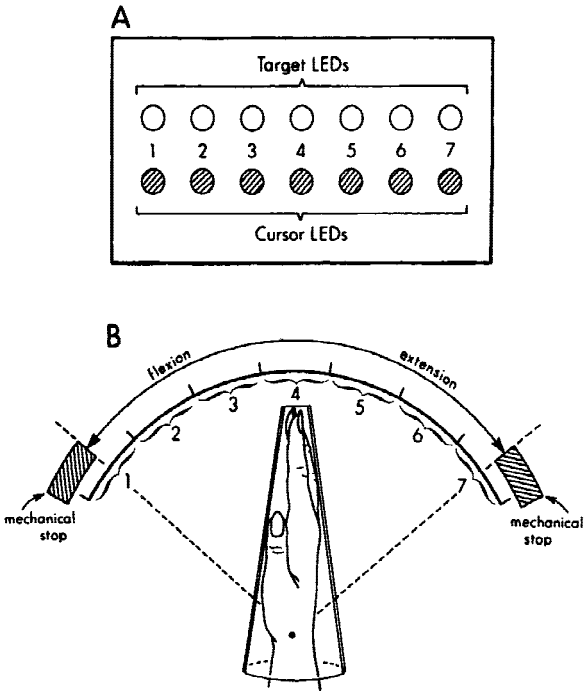


**FIG. 1.** (**A**) Set-up for testing the wrist movement task. Participants executed the task by matching the illuminated cursor light-emitting diodes (LEDs) to the illuminated target LED. (**B**) The manipulandum, showing the range of motion as well as the bounded targets (1 and 7) and the unbounded targets (2 and 6). See text for full explanation.

four alternatives. The UPSIT is highly reliable, with a short-term test–retest correlation of 0.9512. Raw scores consist of the number of correct identifications and were corrected to age- and gender-specific percentile ranks according to previous studies[22] using the second edition of the UPSIT manual.

### Measurement of Mood

Mood was measured with the Beck Depression Inventory (BDI).[22] A number of studies have documented the reliability and validity of the BDI.[23–25]

### Prospective Application of the PD Test Battery

The combined model producing the PD score may separate the patients with probable iPD from normal control subjects because the model was derived from the performance of those patients and control subjects. Development of regression models does not prove the validity of the models beyond the sample.[26] Validity testing requires prospective application of the regression models to an independent sample. Thus, an independent second sample was studied that met the guidelines for the development of clinical tests described above.

The PD score used as a cut-off for determining specificity and sensitivity for the prospective study was determined from the sample on which the test was developed. The PD score that resulted in the greatest specificity and sensitivity in the sample was used in the prospective study. However, the choice of the cut-off score is a subjective decision because of the difficulty of weighing the clinical significance of false-positive and false-negative results.

### Contributions of Tests to the Logistic Regression Model

The relative contribution of each of the three tests to the PD score was tested by forward stepwise logistic regression.

### Statistical Methods

The aim of the study was to develop an equation that relates performance on each of the subtests to produce a score reflective of the probability of being in the probable iPD or control group. Because the outcome of such an analysis is a dichotomous variable, logistic regression analysis was applied. One difficulty is that the data obtained for each subtest is very different (movement velocities, age- and gender-percentile scores on the olfactory test, and a weighted summed response to questions of mood). Further, the large number of variables had to be reduced to allow a reasonable sample size. Consequently, logistic regression analysis was applied to each subtest separately, resulting in a probability score for each subtest (Pwrist for the wrist subtest, Polf for the olfactory subtest, and Pbdi for the mood subtest). These separate probability scores where then used in another logistic regression analysis, which combined the subtests. Pwrist, Polf, and Pbdi constituted the continuous independent variables. In a sense, Pwrist, Polf, and Pbdi are analogous to transformations that result in a type of standard score (for example, a z score) but have the advantage of reducing the number of independent variables. The category of probable iPD or control was the dichotomous dependent variable. This single model resulted in a PD score that indicated the relative probability of being normal as opposed to having probable iPD. The three subtests were considered to be independent measures and, therefore, it is possible that some combination of these subtests would contribute to increases in both specificity and sensitivity.

With respect to the wrist movement subtest, we anticipated that the movement velocities would be skewed toward the most rapid velocity as a result of a physiological limit to the maximum movement velocity. Consequently, the median movement velocity for each task for each participant was used in the analysis. Further, previous studies have shown that reaction times are less consistently abnormal in iPD and, therefore, only movement velocities were studied. Each individual's median movement velocity for each task and the difference in the movement velocities between the flexion bounded and unbounded and extension bounded and unbounded tasks constituted the continuous independent variable for the logistic regression used to determine Pwrist.

## RESULTS

### Test Battery Development

We tested 18 patients with probable iPD who were either newly diagnosed and not on medication ($n = 8$) or were tested after an overnight withdrawal (at least 8 hrs) of their medication ($n = 10$). The mean age of the patients with probable iPD was 64 years (range, 45–79 yrs); nine were women. We also tested 19 normal control subjects. The mean age of the control subjects was 64 years (range, 43–77 yrs); 10 were women. These 37 participants constituted the development group. Those patients under treatment were selected on the basis of mild disease and had a Hoehn and Yahr Scale score of 2 or 2.5. No patient had unilateral disease.

The logistic regression analysis resulted in a PD score for each participant. A PD score of 0.6 was chosen as the cut-off (PD score $\leq 0.6$ for an abnormal result and $>0.6$ for a normal result). This value gave optimal sensitivity and specificity. This model correctly classified 17 of the

18 patients with probable iPD and 18 of the 19 control subjects. Specificities and sensitivities for the combined test battery and the subtests are shown in Table 1.

### Test Contributions to the PD Score

A forward stepwise logistic regression was performed. Pwrist, Polf, and Pbdi constituted the continuous independent variables. The category of probable iPD or control was the dichotomous dependent variable. The olfaction score (Polf) accounted for 48% of the variance, the wrist task (Pwrist) for 26%, and mood (Pbdi) for 9%.

### Prospective Application of the PD Battery

An additional independent sample of 122 control subjects (mean age, 57 yrs; range, 18–93 yrs; 61 women) and 103 patients with probable iPD (mean age, 69 yrs; range, 38–89 yrs; 40 women) were tested. These 225 people constituted the validation group. The PD score was calculated for each participant as described above (Fig. 2). A cut-off PD score of 0.6 was used as determined from the analysis described above. Of 122 control subjects, 108 had PD scores greater than 0.6, and 71 of 103 patients with probable iPD had PD scores less than or equal to 0.6. Specificities, sensitivities, areas under the ROC curves, odds ratios, and chi-square statistics for the combined test battery and the subtests are shown in Table 1.

## DISCUSSION

The regression model combining motor performance, olfaction, and depression scores was robust at distinguishing normal control subjects from patients with mild probable iPD. The specificity, the area under the ROC curve, and the odds ratio were best with the combination of subtests compared with any one test alone, such as olfaction. The sensitivity of the PD Battery combining the subtests in the prospective study was 10 percentage points higher than the olfaction subtest. This result is evidence that these subtests evaluate different domains affected by PD.

We recognize that the symptoms and signs of the patients with probable iPD were severe enough to confirm the diagnosis, even without the PD score. The test battery developed on these patients still could be insufficiently sensitive to recognize early patients with probable iPD whose symptoms and signs were so mild as to make diagnosis difficult for the non-expert. However, we selected newly diagnosed patients with mild disease. Furthermore, if the battery was not successful in distinguishing control subjects from patients with known probable iPD, then there would be little reason to proceed further. In another study, we prospectively studied people with possible iPD but whose symptoms were insufficient for diagnosis to determine the predictive value of the PD Battery.[27]

The sensitivity was low at 69%, which means a large proportion (31%) of patients with early mild PD were missed. However, estimating the significance of the 69% sensitivity is problematic. By comparison, Schoenberg et al. found that 41% of patients with iPD identified on a door-to-door survey had not been previously diagnosed.[28] In a similar survey, Morgante et al. found 35% undiagnosed.[29] These results mean that, by whatever means patients with PD would be diagnosed in these communities (expert vs non-expert physicians), the sensitivity would be 59% and 65%, respectively. However, even this comparison is problematic because these door-to-door surveys overestimate the frequency of missed diagnoses because they did not differentiate between those who were undiagnosed simply because they had not sought medical attention from cases of true missed or

**TABLE 1.** *Characteristics of the test battery for diagnosing probable idiopathic Parkinson's disease in the development and validation groups*

| | Development group | | | |
|---|---|---|---|---|
| Test characteristic | Combined test | Wrist subtest | Olfactory subtest | Mood subtest |
| Specificity (%) | 95 | 89 | 89.5 | 79 |
| Sensitivity (%) | 94 | 63 | 83.5 | 44 |
| | Validation group | | | |
| | Combined test | Wrist subtest | Olfactory subtest | Mood subtest |
| Specificity (%) | 88 | 78 | 78 | 75.4 |
| Sensitivity (%) | 69 | 50 | 74 | 55.3 |
| Area under the ROC curve | 0.87 | 0.67 | 0.83 | 0.73 |
| Odds ratio (95% confidence interval) | 17.1 (8.5–34.3) | 3.5 (2.0–6.2) | 10.1 (5.5–18.0) | 3.8 (2.2–6.7) |
| Chi-square | 76 (df = 1, p <0.001) | 17.1 (df = 1, p <0.001) | 57.9 (df = 1, p <0.001) | 20.9 (df = 1, p <0.001) |

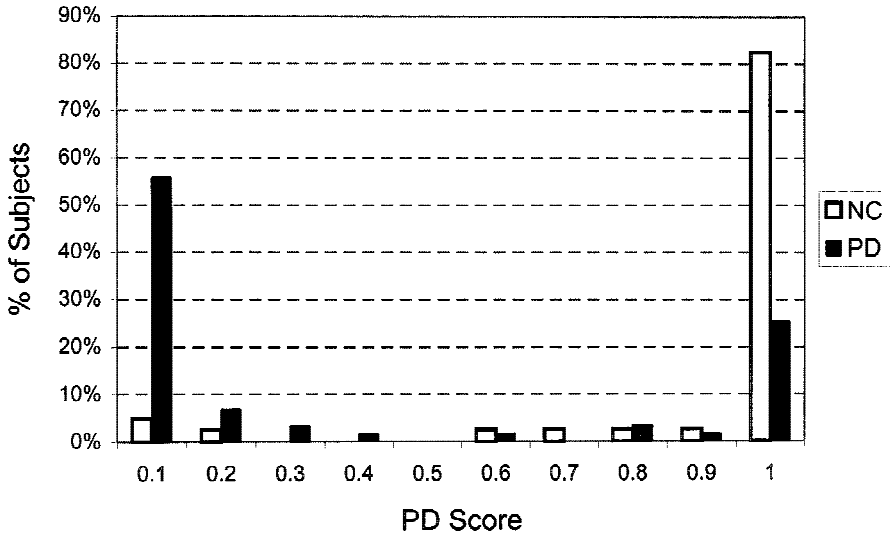


**FIG. 2.** Distribution of test battery scores for patients with probable idiopathic Parkinson's disease and normal control subjects.

false diagnoses who were actively under the care of a physician.

Sensitivity could be increased by changing the PD score cut-off value; however, this change would reduce specificity. What is an acceptable sensitivity and specificity cannot be determined at this time. Acceptable sensitivity and specificity ultimately will be determined by the cost of treating a falsely diagnosed normal person versus the cost of not treating a correctly diagnosed patient with PD.

Even if the PD Battery was not helpful in identifying extremely mild cases of iPD, it still would be of use by identifying the large number of patients with clinically significant iPD who are often not diagnosed or in whom the diagnosis is delayed. The findings of a survey we conducted of 93 consecutive patients with iPD show that most patients (59%) sought medical consultation within 1 year of first noticing a symptom (unpublished observations). However, 41% waited 1 year or more before seeking help, and 21% waited 2 years or more. Either greater effort is needed to educate people about iPD or some screening test of the general population is needed to prevent patients from waiting too long to seek treatment.

The design of this study does not allow us to compare the diagnostic efficacy of the PD Battery with that of expert physicians, such as a movement disorders specialist. Such a study would be difficult to design because there is no antemortem "gold standard" to establish the diagnosis other than the opinion of a movement disorders specialist. What independent antemortem test could be used to prove the diagnosis of a movement disorders specialist as incorrect? Consequently, confirmation of the diagnosis by a movement disorders specialist is the de facto "gold standard." However, the PD Battery could be of use as an initial screening test, especially in those areas in which movement disorders specialists are not readily available. Another use would be if the PD Battery could anticipate the diagnosis of iPD by a movement disorders specialist at a time when symptoms are insufficient to warrant the diagnosis.[27]

Much work remains. As yet it is unclear how comorbidities, such as arthritis, cerebral vascular disease, dementia, and so on, would affect test performance. These and other issues need to be examined before any test battery can be applied to the general population.

**Acknowledgments:** This project was supported in part by grants from the National Institutes of Health (NS R01 33924, E.B.M. and W.C.K.), the Arizona Chapter of the American Parkinson Disease Association (E.B.M.), the Jane K. Pelton Fund for Movement Disorders (E.B.M.), the Krupa Fund, and the Arizona Disease Control Research Commission (E.B.M.).

The authors thank Dr. Andy Harsen, Dr. Cynthia Holmes, Dr. Kenneth Baker, Cindy Kirkpatrick, Berta Leis, and Julie Haynes for their assistance.

## REFERENCES

1. Tetrud JW, Langston JW. The effects of deprenyl (selegiline) on the natural history of Parkinson's disease. *Science* 1989;245:519–522.
2. The Parkinson Study Group. Effects of tocopherol and deprenyl in Parkinson's disease. *N Engl J Med* 1993;328:176–183.
3. Schulzer M, Mak E, Calne DB. The antiparkinson efficacy of deprenyl derives from transient improvement that is likely to be symptomatic. *Ann Neurol* 1992;32:795–798.
4. Olanow CW, Hauser RA, Gauger L, et al. The effect of deprenyl and levodopa on the progression of Parkinson's disease. *Ann Neurol* 1995;38:771–777.
5. Berger BD, Wise CD, Stein L. Nerve growth factor: enhanced

recovery of feeding after hypothalamic damage. *Science* 1973;185: 506–508.
6. Date L, Notter MFD, Felten SY, Felten DL. Stereotaxic injections of Gd1a ganglioside induces limited recovery of striatal dopaminergic system in MPTP-treated aging mice. *J Neurosci Res* 1991; 28:525–530.
7. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules: applications and methodological standards. *N Engl J Med* 1985;313:793–799.
8. Montgomery EB Jr, Gorman D, Nuessen J. Motor initiation versus execution in normal and Parkinson's disease subjects. *Neurology* 1991;41:1469–1475.
9. Montgomery EB, Nuessen J, Gorman D. Reaction time and movement velocity abnormalities in Parkinson's disease under different task conditions. *Neurology* 1991;41:1476–1481.
10. Patrick HT, Levy DM. Parkinson's disease: a clinical study of 146 cases. *Arch Neurol Psychiatry* 1922;7:711–720.
11. Ward CD, Hess WA, Calne DB. Olfactory impairment in Parkinson's disease. *Neurology* 1983;33:943–946.
12. Doty RL, Newhouse MG, Azalina JD. Internal consistency and short-term test–retest reliability of the University of Pennsylvania Smell Identification Test. *Chem Senses* 1985;33:943–946.
13. Quinn NP, Rossor MN, Marsden CD. Olfactory threshold in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1987;50:88–89.
14. Doty RL, Perl DP, Steele JC, et al. Olfactory dysfunction in three neurodegenerative diseases. *Geriatrics* 1991;46:47–51.
15. Celesia GG, Wanamaker WM. Psychiatric disturbances in Parkinson's disease. *Diseases of the Nervous System* 1972;33:577–583.
16. Gotham AM, Brown RG, Marsden CD. Depression in Parkinson's disease: a quantitative and qualitative analysis. *J Neurol Neurosurg Psychiatry* 1986;49:381–389.
17. Brown GL, Wilson WP. Parkinsomism and depression. *South Med J* 1972;33:577–583.
18. Mayeux R, Stem Y, Rosen J, Leventhal I. Depression, intellectual impairment and Parkinson's disease. *Neurology* 1981;31:645–650.
19. Mayeux R, Williams JBW, Stem Y, Cote L. Depression and Parkinson's disease. *Adv Neurol* 1984;40:235–240.
20. Doty RL, Golbe LI, McKeown DA, et al. Olfactory testing differentiates between progressive supranuclear palsy and idiopathic Parkinson's disease. *Neurology* 1993;43:962–965.
21. Doty RL. *The Smell Identification Test Administration Manual.* Haddon Heights, NJ: Sensonics, Inc, 1983.
22. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4: 561–571.
23. Gallagher D, Brenkenridge J, Steinmetz J, Thompson LW. The Beck Depression Inventory and research diagnostic criteria: congruence in an older population. *J Consult Clin Psychol* 1983;51: 945–951.
24. Jacobs KW, Boze MM. Correlations among scales of the Beck Depression Inventory and the profile of mood states. *Psychol Rep* 1993;73:431–434.
25. Leahy JM. Validity and reliability of the Beck Depression Inventory-short form in a group of bereaved females. *J Clin Psychol* 1992;48:64–68.
26. Stevens J. *Applied Multivariate Statistics for the Social Sciences,* 2nd ed. Montvale, NJ: Lawrence Erlbaum Associates, 1992.
27. Montgomery EB Jr, Lyons K, Koller WC. Early detection of or risk for Parkinson's disease. II: a prospective study. *Mov Disord*. In press.
28. Schoenberg BS, Anderson DW, Haerer AF. Prevalence of Parkinson's disease in the biracial population of Copiah County, Mississippi. *Neurology* 1985;35:841–845.
29. Morgante L, Rocca WA, Di Rosa AE, et al. Prevalence of Parkinson's disease and other types of parkinsomsm: a door-to-door survey in three Sicilian municipalities. *Neurology* 1992;42:1901–1907.