INVITED COMMENTARY

Evaluating Dementia Using PET: How Do We Put into Clinical Perspective What We Know to Date?

Decreasing mortality, with consequent progressive aging of the mature adult population, has led to a rising prevalence of senile dementia. The condition is tremendously costly to patients, their families, and society in general. Alzheimer's disease (AD) in the United States affects over 4 million people, who incur associated yearly expenses of nearly $70 billion; when indirect costs such as the lost productivity of caregivers are considered, total annual expenditures approximate $100 billion. As the ''baby boomers'' approach senior citizen status in the 21st century, it has been estimated that over 14 million Americans will suffer from AD by 2050 (*1–4*).

Regional cerebral metabolic patterns of patients with dementia reflect pathophysiologic alterations, even before they lead to symptomatic expression (*5*). The use of PET in evaluating dementia has been studied since the early 1980s (*6–9*) and has been extensively reviewed in recent years (*10–14*). The best studied application of PET in dementia uses FDG to evaluate AD. Assessment of the diagnostic accuracy of PET, even for this application, however, has been hindered by the paucity of data involving scans of patients undergoing subsequent neuropathologic examination.

The article by Hoffman et al. (*15*) in this issue of *The Journal of Nuclear Medicine* addresses an issue of pressing clinical relevance: the establishment of reliable estimates of diagnostic accuracy of FDG PET (the present imaging modality of choice in the assessment of primary neurodegenerative dementias) for AD. In doing so, their article provides a substantial addition to the currently limited base of data in the peer-reviewed scientific literature on patients studied by both brain PET and pathologic analysis. The authors sampled patients at a university-based memory disorder clinic, staffed by board-certified neurologists who specialized in dementia evaluation; patients were included if their impairment of memory was objectively documentable but the diagnosis was ''challenging or difficult'' according to the evaluating clinicians. A nuclear medicine physician who was unaware of clinical information and who interpreted FDG PET was able to predict the ultimate diagnosis of AD in 13 of the 14 patients (92.9%) for whom that was the only evident neuropathology and for 14 of the 16 patients (87.5%) harboring AD-related changes in their brains. In comparison, a diagnosis of AD had been clinically suspected (considered probable) in 64.3% and 62.5%, respectively, and considered possible in another 14.3% and 12.5%, respectively. Thus, under the best of circumstances, AD would be clinically missed in 21% of these patients, whereas under those same circumstances, only 7% of AD cases would escape detection on FDG PET, even in the absence of knowledge of any other clinical information. Because of the lower diagnostic specificity found for FDG PET (75% vs. 100%), however, corresponding overall accuracies for clinical evaluation and PET were identical (86.4%). The difference in specificity was entirely accounted for by 2 cases—1 of Creutzfeldt-Jacob disease and 1 of Lewy Body disease—both with clinically evident diagnoses and associated with bilateral parietotemporal hypometabolism. Thus, if PET had been interpreted in light of the available clinical information (as would likely occur in its routine clinical application), rather than blindly (as done here to address a specific research question), a diagnostic specificity of 100% would pertain. FDG PET did correctly classify as non-AD the cases (1 each) of progressive supranuclear palsy, preamyloid, mesio-limbo cortical degeneration, and nonspecific neuronal degenerative change, as did clinical evaluation.

The most serious problem of the current investigation (*15*) is the inadequacy of its statistical strength to meaningfully assess specificity, because of the low number of patients included in the study who were found on histopathologic exam to be free of AD. This problem arises from the general difficulty, shared by all previous such studies, in obtaining sufficiently sizable samples of patients who have had brain PET and are followed to pathologic confirmation. Because of this, values are reported that would be substantially altered by shifts in the ''truth tables'' of just 1 or 2 patients. This statistical weakness can be appreciated quantitatively by considering that the 95% confidence interval (not reported by the Hoffman et al.) pertaining to the 75% specificity value ranges from 45% to 100%. In other words, the actual specificity for correctly identifying non-AD patients could range anywhere from essentially perfect to a coin flip, given the number and distribution of cases included in the study. Another shortcoming with this study, as presented, is that it is difficult to determine the appropriate setting to which it applies, because the degree of cognitive impairment of the patients examined is nowhere quantified (e.g., by Mini-

Received May 31, 2000; accepted Jun. 1, 2000.
For correspondence or reprints contact: Daniel H. Silverman, MD, PhD, Department of Molecular and Medical Pharmacology, UCLA School of Medicine, CHS AR-144, Los Angeles, CA 90095-6942.

Mental State Examination [MMSE] (*16*)) or otherwise described.

Despite its limitations, the study does provide data derived from compilation of the largest PET-studied autopsy–biopsy series yet reported, which agree with data of other analyses. In a recent review (*10*) that provided a pooled analysis of another 20 patients with dementia for which both PET and autopsy data were available (*17–19*), the sensitivity, specificity, and overall accuracy were 92%, 71%, and 85%, respectively. In a preliminary analysis of a larger (n = 70) multicenter, pooled population that included the current sample, along with those from several other institutions in the United States and Europe, corresponding values were 96%, 67%, and 87% (*20*). Those analyses agree with the values reported for the 22 cases in this study. Thus, diagnostic sensitivity has been found consistently to be quite high (range, 94% ± 2%), with specificity falling in the range of 71% ± 4%. With respect to each of those analyses, it should be kept in mind that the patient samples were not generally representative of people who undergo dementia evaluations but of particularly diagnostically challenging cases—the small subset of patients for whom both brain PET and biopsy–necropsy had been obtained to help establish the diagnosis. This could in turn affect the assessment of diagnostic specificity, causing it to be underestimated relative to a more routine clinical application of FDG PET in which a smaller proportion of non-AD patients would have, for example, Creutzfeldt-Jacob disease.

In addition to the diagnostic value FDG PET may have in evaluation of dementia, it may also serve explicitly as a prognostic tool to determine likelihood of deterioration of mental status in the period after the time of scanning. For example, relative hypometabolism of associative cortex can be used to predict whether cognitive decline will occur (at a pace faster than would be expected for healthy aging) over the several years after a PET evaluation (*21–23*). Moreover, the magnitude of decline over a 2-y period for some standardized measures of memory was recently found to correlate significantly with the degree of initial hypometabolism of inferior parietal, superior temporal, and posterior cingulate cortical regions (*24*), with Pearson correlation coefficients ranging as high as 0.71. In our experience, visually interpreted PET has had a prognostic sensitivity of 90%–93% and a prognostic specificity of 74%–77%, with an overall accuracy of 83%–85% for predicting pathologic clinical progression in the several years (up to 9 y; mean, 3 y) after the scan. This high sensitivity of FDG PET in patients with mild impairment (mean MMSE, 25/30) suggests that by the time a patient presents with symptoms of a progressive neurologic process, sufficient alteration of cortical activity generally has occurred to diminish metabolism of certain areas of the brain to an extent readily detectable on FDG images. At the same time, the lower specificity points to the existence of processes other than those poised to cause imminent cognitive deterioration capable of producing hypometabolic foci discernible with PET. This sensitivity–specificity situation, for clinical progression consequent to metabolically evident disease, bears resemblance to that pertaining to the well-established use of PET for evaluation of a solitary pulmonary nodule, where it is unusual for a focus that is not hypermetabolic to go on to demonstrate progressive growth at a rate indicating malignancy, although several benign diagnostic entities may cause focal hypermetabolism.

Ultimately, how much need there is for functional neuroimaging in the evaluation of dementia depends largely on the adequacy of a traditional diagnostic work-up without imaging. What does the primary scientific literature reveal regarding accuracy of clinical diagnosis? Considering the critical importance of this question, surprisingly few studies have addressed it systematically—with a representative patient mix and in a way capable of yielding measures of true sensitivity and specificity. In 1 study that was so designed, of 421 cognitively impaired patients being clinically followed, diagnostic comparisons were made for the first 58 who died (*25*). The patients' mean age was 79 y, and all but 3 subjects were at least 65 y. Histologic review revealed that nearly half (28/58) had neuropathologic hallmarks of AD, a diagnosis that was missed in 8 subjects. Conversely, the diagnosis of AD was made clinically in 8 of the 30 subjects who had no pathologic evidence for AD. The sensitivity, specificity, and overall accuracy were thus 71%, 73%, and 72%, respectively. Compared with PET studies, this represents a substantially lower sensitivity, at a comparable level of specificity; these numbers suggest that for every 100 patients with AD who are examined, PET would find about 20 whose diagnosis would have been missed by clinical evaluation. Few clinical diagnosis studies have provided an analysis that is stratified for severity, as would be needed to address the issue of diagnostic accuracy in the earlier stages of disease. One study that did specifically address the question of clinical detection of very-mild disease followed patients who initially appeared normal or minimally affected for an average of 4 y (*26*). Even by the end of this longitudinal follow-up period, a neurologist examiner detected AD in only 70% of the patients who were histologically positive. On the other end of the severity spectrum, 65 elderly patients with moderate to severe dementia were evaluated clinically and followed longitudinally until death (*27*). The mean duration of their symptoms was 8.9 y. Autopsies revealed that AD existed in 48 patients (10 of whom also had evidence of multiple infarcts). By the time those patients had reached this late stage of disease, sensitivity of AD diagnosis made clinically (45/48 = 94%) was similar to that made with PET at much earlier stages of disease, but 10 of the remaining 17 patients were also clinically (mis)diagnosed as having AD, indicating a specificity of only 41%. Another study was performed with 25 patients who had advanced-stage dementia; functional status had deteriorated to the point where they were living in a long-term care

facility. All of the patients met the *Diagnostic and Statistical Manual for Mental Disorders, third edition* (*28*), criteria for primary neurodegenerative disorder and National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) criteria for AD, yet only 68% were found on pathologic review to actually have AD. Most of the misdiagnosed cases were in younger patients. Finally, in a study of 54 patients with dementia who were not stratified by severity but included a good diagnostic mix (in descending frequency: AD, multiple infarct dementia, Parkinson's dementia, Creutzfeldt-Jakob disease, subcortical gliosis, progressive supranuclear palsy, and Huntington's disease), 2 independent neurologists agreed on the correct diagnosis in just 63% of cases; their average individual accuracy rate was 71.5% (*29*).

As well-articulated in the article by Hoffman et al. (*15*), the gold-standard pathologic diagnosis has not always turned out to be as unambiguous a determination as one might suppose. This is especially the case when pathologic diagnosis is based on biopsy, which allows only a small area of (what one hopes to be representative) tissue to be sampled; even when full brain autopsy is performed, the application of different histopathologic criteria can result in divergent diagnoses in a substantial proportion of cases. Efforts over the last decade to bring standardization to this field, however, have resulted in schemes of consensus criteria with much higher agreement among them, as discussed by Mirra et al., who applied 1 such scheme (*30*) in their study.

In conclusion, to place the above considerations in their wider clinical context, it is important to note that dementing illness in patients with mild cognitive dysfunction frequently is unrecognized and, consequently, goes untreated because of the difficulty in making an accurate clinical diagnosis in the early stages of dementia (*31*). This may have been less relevant at a time when little could be done to improve the cognitive symptoms of most patients with dementia. However, advances occurring over the last half-dozen years have changed that situation. For example, tacrine, donepezil, and rivastiamine have been cleared by the U.S. Food and Drug Administration for improving cognition or slowing intellectual decline associated with mild to moderate AD. There is evidence accumulating that such cholinesterase inhibitors can also improve global function and decrease the need for placement of patients in long-term care facilities (*32*,*33*) and that to derive maximal benefit, the use of these drugs should be initiated early in the course of disease. Moreover, both vitamin E and selegiline were recently demonstrated to delay progression of symptoms, in a double-blind, placebo-controlled, randomized, multicenter trial of 341 patients with moderate AD (*34*). In addition, patients predisposed to or with early signs of vascular dementia may modify their risk of progression with antihypertensive and antithrombotic medications or vascular surgery. In addition to the significant strides that have been made in the past decade in developing therapies for patients with AD, the present decade has opened with an improved understanding of etiologic factors, now being exploited in the development of more specifically targeted therapies aimed at arresting progression of AD (*35*). These 3 areas of advancement—better understanding of the basic neurodegenerative mechanisms, development of more effective therapies administered before significant irreversible neurodegeneration has occurred, and enhanced ability to establish accurate diagnoses early on—are concomitantly driving progress on each other's fronts. The immediate diagnostic challenge—to accurately identify minimally affected patients early enough to allow them to reap the greatest potential therapeutic benefits—will be met best by using the tools most well equipped to facilitate making the diagnosis with a high degree of sensitivity and overall accuracy at the earliest stage of disease. The current study bolsters the use of FDG PET, to improve accuracy of evaluation of patients with dementia, in meeting that challenge.

Daniel H.S. Silverman
Michael E. Phelps
University of California Los Angeles
School of Medicine
Los Angeles, California

REFERENCES

1. Evans DA. Estimated prevalence of Alzheimer's disease in the US. *Milbank Q*. 1990; 68:267–289.
2. Carr DB, Goate A, Phil D, Morris JC. Current concepts in the pathogenesis of Alzheimer's disease. *Am J Med*. 1997;103:3S–10S.
3. Ernst RL, Hay JW. The U.S. economic and social costs of Alzheimer's disease revisited. *Am J Pub Health*. 1994;84:1261–1264.
4. National Institute on Aging. *Progress report on Alzheimer's disease* [NIH Publication No. 96–4137]. Bethesda, MD: National Institute on Aging; 1996.
5. Small GW, La Rue A, Komo S, Kaplan A, Mandelkern MA. Predictors of cognitive change in middle-aged and older adults with memory loss. *Am J Psychiatr.* 1995;152:1757–1764.
6. Farkas T, Ferris SH, Wolf AP, et al. ^{18}F-2-deoxy-2-fluoro-D-glucose as a tracer in the positron emission tomographic study of senile dementia. *Am J Psychiatr.* 1982;139:352–353.
7. Benson DF, Kuhl DE, Phelps ME, Cummings JL, Tsai SY. Positron emission computed tomography in the diagnosis of dementia. *Trans Am Neurol Assoc*. 1981;106:68–71.
8. Frackowiak RS, Pozzilli C, Legg NJ, et al. Regional cerebral oxygen supply and utilization in dementia: a clinical and physiological study with oxygen-15 and positron tomography. *Brain*. 1981;104:753–778.
9. Foster NL, Chase TN, Fedio P, et al. Alzheimer's disease: focal cortical changes shown by positron emission tomography. *Neurology*. 1983;33:961–965.
10. Silverman DHS, Small GW, Phelps ME. Clinical value of neuroimaging in the diagnosis of dementia: sensitivity and specificity of regional cerebral metabolic and other parameters for early identification of Alzheimer's disease. *Clin Positron Imaging.* 1999;2:119–130.
11. Friedland RP, Jagust WJ. Positron and single photon emission tomography in the differential diagnosis of dementia. In: Duara R, ed. *Positron Emission Tomogaphy in Dementia*. New York, NY: Wiley-Liss, Inc;1990:161–177.
12. Haxby JV. Resting state regional cerebral metabolism in dementia of the Alzheimer type. In: Duara R, ed. *Positron Emission Tomogaphy in Dementia.* New York, NY: Wiley-Liss, Inc;1990:93–116.
13. Mazziotta JC, Frackowiak RSJ, Phelps ME. The use of positron emission tomography in the clinical assessment of dementia. *Semin Nucl Med*. 1992;22: 233–246.
14. Herholz K. FDG PET and differential diagnosis of dementia. *Alzheim Dis and Assoc Disord*. 1995;9: 6–16.
15. Hoffman JM, Welsh-Bohmer KA, Hanson M, et al. FDG PET imaging in patients with pathologically

verified dementia. *J Nucl Med.* 2000;41:1920–1928.
16. 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.* 1975;12:189–198.
17. Salmon E, Sadzot B, Maquet P, et al. Differential diagnosis of Alzheimer's disease with PET. *J Nucl Med.* 1994;35:391–398.
18. Tedeschi E, Hasselbach SG, Waldemar G, et al. Heterogeneous cerebral glucose metabolism in normal pressure hydrocephalus. *J Neurol Neurosurg Psychiatr.* 1995;59:608–615.
19. Mielke R, Schröder R, Fink GR, et al. Regional cerebral glucose metabolism and postmortem pathology in Alzheimer's disease. *Acta Neuropathol.* 1996;91:174–179.
20. Silverman DHS, Small GW, Kung de Aburto MA, et al. Diagnostic accuracy of FDG-PET in evaluation of dementia: international multi-center pooled brain scan and autopsy data [abstract]. *J Nucl Med.* 2000;41(suppl):63P.
21. Silverman DHS, Chang CY, Cummings JL, et al. Prognostic value of regional brain metabolism in evaluation of dementia: comparison with long-term clinical outcome [abstract]. *J Nucl Med.* 1999; 40(suppl):71P.
22. Silverman DHS, Lu CL, Czernin J, Small GW, Phelps ME. Prognostic value of brain PET in patients with early dementia symptoms, treated or untreated with anticholinesterase therapy [abstract]. *J Nucl Med.* 2000;41(suppl):64P.
23. McKhann G, Drachmann 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.* 1984;34:939–944.
24. Small GW, Ercoli LM Silverman DHS, et al. Cerebral metabolic and cognitive decline in persons at genetic risk for Alzheimer's disease. *Proc Nat Acad Sci USA.* 2000;97:6037–6042.
25. Molsa P, Parjavi L, Rinne J, Rinne U, Sako E. Validity of clinical diagnosis in dementia: a prospective clinicopathological study. *J Neurol Neurosurg Psychiatr.* 1985;48:1085–1090.
26. Morris JC, McKell DW, Storandt M, et al. Very mild Alzheimer's disease: informant-based clinical, psychometric, and pathologic distinction from normal aging. *Neurology*. 1991;41:469–478.
27. Wade JP, Mirsen TR, Hachinski VC, et al. The clinical diagnosis of Alzheimer's disease. *Arch Neurol.* 1987;44:24–29.
28. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 3rd ed. Washington, DC: American Psychiatric Association; 1987.
29. Boller F, Lopez OL, Moossy J. Diagnosis of dementia: clinicopathologic correlations. *Neurol.* 1989;39:76–79.
30. Mirra S, Heyman A, McKeel D, et al. The consortium to establish a registry for Alzheimer's disease (CERAD), Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology.* 1991;41:479–486.
31. Small GW, Rabins PV, Barry PP, et al. Diagnosis and treatment of Alzheimer disease and related disorders: consensus statement of the American Association for Geriatric Psychiatry, the Alzheimer's Association, and the American Geriatrics Society. *JAMA*. 1997;278:1363–1371.
32. Knopman D, Schneider LS, Davis K, et al. Long-term tacrine (Cognex) treatment effects on nursing home placement and mortality. The Tacrine Study Group. *Neurology.* 1996;47:166–177.
33. Rogers SL, Farlow MR, Doody RS, Mohs R, Friedhoff LT. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group. *Neurology.* 1998;50:136–145.
34. Sano M, Ernesto C, Thomas RG, et al. Controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. *N Engl J Med*. 1997;336:1216–1222.
35. Selkoe DJ. The origins of Alzheimer disease. *JAMA.* 2000;283:1615–1617.