Published in final edited form as:

Am J Geriatr Psychiatry. 2010 April ; 18(4): 363–370.

Reliability and Validity of an Algorithm for the Diagnosis of Normal Cognition, MCI and Dementia: Implications for Multi-Center Research Studies

Ranjan Duara, MD^{1,2,4,5}, David A. Loewenstein, PhD^{1,2}, Maria Greig, MD¹, Amarilis Acevedo, PhD², Elizabeth Potter, PhD¹, Jason Appel, MD³, Ashok Raj, MD^{4,5}, John Schinka, PhD^{4,5,6}, Elizabeth Schofield, MPH^{4,5}, Warren Barker, MS¹, Yougui Wu, PhD^{4,5}, and Huntington Potter, PhD^{4,5}

¹ Mount Sinai Medical Center, Wien Center for Alzheimer's Disease and Memory Disorders, Miami Beach, FL, USA

² University of Miami, Miller School of Medicine, Psychiatry & Behavioral Sciences, Miami, FL, USA

³ Mount Sinai Medical Center, Radiology, Miami Beach, FL, USA

⁴ Johnnie B. Byrd, Sr. Alzheimer's Center & Research Institute, Tampa, FL, USA

⁵ University of South Florida, Tampa, FL

⁶ James A. Haley Veterans' Hospital, Tampa, Florida

Abstract

Background—The traditional consensus diagnosis (ConsDx) of normal cognition, mild cognitive impairment (MCI), and dementia relies on the reconciliation of an informant-based report of cognitive and functional impairment by a physician (PhyDx), and a neuropsychological diagnosis (NPDx). As this procedure may be labor-intensive and influenced by the philosophy and biases of a clinician, the diagnostic algorithm (AlgDx) was developed to identify individuals as cognitively normal, with MCI or dementia.

Methods—The AlgDx combines the PhyDx with the NPDx, using a computational algorithm that provides cognitive diagnoses, as defined by the National Alzheimer Coordinating Center/Uniform Data Set (NACC/UDS) nomenclature. Reliability of the AlgDx was assessed in 532 community-dwelling elderly subjects by its concordance with the ConsDx, and association with two biomarkers, medial temporal atrophy (MTA) scores of brain MRI scans, and ApoE-ε4 genotype.

Results—A high degree of concordance was observed between ConsDx and AlgDx with a weighted Cohen's kappa of 0.84. Concordance of the AlgDx to the same ConsDx categories ranged from 85% to 92%. Excellent discriminative validity was observed as AlgDx, MTA scores, and APOE-ε4 allele frequencies distinguished amnesic MCI and dementia subjects from normal subjects.

Conclusion—The AlgDx of normal cognition, MCI, and dementia is a valid alternative that reduces time, effort and biases associated with the ConsDx. The inherent reliability of a fixed algorithm, together with its efficiency and avoidance of individual bias, suggests the AlgDx may be used in longitudinal, multi-site clinical trials and population studies of MCI and dementia.

Address Correspondence To: Ranjan Duara, M.D., Wien Center for Alzheimer's Disease and Memory Disorders, Mt. Sinai Medical Center, 4300 Alton Rd, Miami Beach, FL 33140, Telephone: 305-674-2543; Fax: 305-532-5241, ranjan-duara@msmc.com.

Disclosure: The authors report no conflicts of interest.

Statistical analysis was performed by David A. Loewenstein, Ph.D., University of Miami

Keywords

Diagnosis; Algorithmic Diagnosis; Mild Cognitive Impairment; Dementia; Longitudinal Studies

INTRODUCTION

Progressive dementing disorders, most typically represented by Alzheimer's disease, have a prodromal phase during which subtle deterioration in cognitive and functional abilities progress from a cognitively normal stage to mild cognitive impairment (MCI), eventually leading to dementia (1–5). Methods have been proposed for making an accurate diagnosis of Alzheimer's disease (AD) well before criteria for dementia are fulfilled (2–4), typically during the stage characterized as amnesic MCI (aMCI). The motivation for making an earlier diagnosis of AD is to enable interventions such as pharmacological agents to be used when they are most effective. Evaluating efficacy of these interventions requires reliable and reproducible methods in diagnosing the early stages of AD.

Mild Cognitive Impairment and Cognitive Impairment No Dementia (CIND) are defined as pre-dementia states with identifiable cognitive impairment, but without functional impairment (6,7). Yet it is clear most individuals diagnosed with MCI or CIND do have subtle functional impairments; this is supported by the observation that Clinical Dementia Rating (CDR) scale scores progressively increase during the transition from *no cognitive impairment* (NCI) to MCI and dementia (8–13). Further, it is evident that MCI and CIND do not represent single points in the continuum between NCI and dementia, but rather a range of cognitive and functional impairments. Thresholds for diagnosing MCI and CIND can be arbitrary (6,7,14), leading to considerable diagnostic variability among clinicians and research teams. Methods that ensure high inter-individual and inter-site reliability in diagnosing MCI and dementia should provide greater power, and require fewer subjects to obtain reliable results for clinical, epidemiological, and especially longitudinal research studies.

The National Alzheimer's Coordinating Center (NACC) Uniform Data Set (UDS; <http://www.alz.washington.edu/>) provides the nomenclature for diagnostic entities interposed between NCI and dementia. NACC criteria (15) rely on the traditional consensus diagnosis (ConsDx) that is achieved when two or more clinicians discuss the patient's clinical history and neuropsychological test performance. Homogeneity of methods in reconciling the physician's (PhyDx) and the neuropsychologist's diagnosis (NPDx), which tend to be discrepant during pre-dementia states, is particularly important for achieving reproducible diagnoses in cross-sectional and longitudinal studies. In addition to its potential for introducing variability (16,17), the ConsDx requires dedicated time and effort by highly skilled professionals who have already invested their time and effort in evaluating and diagnosing the subject. Thus, methods that achieve high reproducibility in reconciling potentially divergent opinions regarding the diagnosis of dementia and pre-dementia states could reduce costs and improve the reliability and efficiency of multi-center and longitudinal studies of AD and MCI states.

This study evaluates performance of a computational algorithm (AlgDx) that was developed to diagnose dementia and pre-dementia states as described by NACC. Inter-rater reliabilities of PhyDx and NPDx were assessed by comparing diagnostic results from two independent teams of raters. Reliability of the AlgDx was assessed by comparison to the ConsDx. Validity of the AlgDx was assessed by its relationship to medial temporal atrophy (MTA) on brain MRI scans, and Apolipoprotein E (ApoE)- ϵ 4 allele frequency; MTA is a biomarker correlated to the Braak stages of AD pathology (18–20), while ApoE- ϵ 4 allele expression has been correlated to beta amyloid deposition in the AD brain (21).

METHODS

Subject Recruitment

Subjects (n=532) were recruited into the Florida Alzheimer's Disease Research Center Clinical Core (FADRC-CC) by advertisement from the community, and from clinics in two major cities, Miami and Tampa, Florida, and adjacent cities. Subjects recruited in the Miami area were likely to be clinic-based, i.e., the Wien Center for Alzheimer's Disease and Memory Disorders, at Mount Sinai Medical Center, Miami Beach, Florida, or had responded to advertisements for a well-established memory screening study. Subjects recruited from the Tampa area were more likely to be community-based, and to have responded to advertisements asking for healthy volunteers to participate in a normal aging study.

Subjects were 52 to 92 years of age and had a reliable informant who could provide a detailed history. The following evaluations were done: (1) full clinical history, obtained from a reliable informant; (2) neurological evaluation; (3) Mini-Mental State Evaluation (MMSE) (22); (4) CDR scale (8); (5) neuropsychological test battery (15), as outlined in NACC protocol. Additional tests included the Three Trial Fuld Object Memory Evaluation (OME) (23), Hopkins Verbal Learning Test (HVLT) (24), and Stroop Color-Word Test (25). In addition, brain MRI scans and standard blood tests for dementia work-up were obtained. Details of the evaluation methods have been described previously (26).

Diagnostic Procedures

1. Physician's Cognitive Diagnosis (PhyDx): *The physician assigned a cognitive diagnosis of NCI, MCI or Dementia, based on the subject's entire clinical history, including his/her functional status (which was derived from the history itself and from the CDR rating and the Functional Activity Questionnaire) as well as the MMSE score and sub-scores.* Factors that could modify the physician's impression about the subject's cognitive and functional abilities were taken into consideration. These factors included the subject's educational and cultural background, *sensory (especially visual and hearing) and motor deficits*, language and speech disorders, medical and psychiatric conditions and the perceived reliability of the informant.
2. Neuropsychological Diagnosis (NPDx): Each subject in the study was administered a neuropsychological test battery in his/her native language (English or Spanish) as described above. In our cohort, we have a large battery of NACC and non-NACC neuropsychological tests. We have previously found that the probability of spurious findings at a specific cut-off for impairment is greatly enhanced by multiple measures. Therefore, we wanted to limit the size of the diagnostic battery and to employ normative data applicable to both Spanish-speaking and English-speaking subjects. To assess memory, we employed the 3-trial Fuld Object Memory Evaluation (23) and Delayed Visual Reproduction of the Wechsler Memory Scale-R (27). Tests of non memory function included category fluency (language) (28), letter fluency (language) (29), Block Design-WAIS-III (visuospatial) (30), Trails B (Executive) (31) and Similarities-WAIS-R (Executive) (30). Neuropsychological classification was achieved employing methods developed by Loewenstein et al. (32). The threshold for MCI syndromes was a test score of 1.5 SD or greater below expected normative values, accounting for age, educational level, and language of administration based on a large co-normed normative database used in previous studies (32–34). The threshold for dementia was a score that was 2.0 SD or greater below expected normative values in one memory and one non-memory test. This corresponds to a confirmation of a dementia syndrome at or below the 5th percentile specified by NINCDS-ADRDA criteria, whereas 1.5 SD is typically the cut-off used for the cut-off for MCI (34–35). Nomenclature used for NPDx was as follows: No Cognitive

Impairment (NCI), Non-Amnesic MCI (na-MCI, single or multi-domain), Amnesic MCI (a-MCI, single or multi-domain) and Dementia.

3. Cognitive Diagnoses: NACC/UDS (D1 diagnosis) nomenclature was used as described for the NPDx. The final cognitive diagnosis was made by two separate methods: (a) traditional Consensus Diagnosis (ConsDx), and (b) Algorithmic Diagnosis (AlgDx).
 - a. ConsDx was achieved by discussion between two members of the consensus team, namely, the physician and neuropsychologist. These professionals reviewed and discussed the subject's clinical history, CDR scores, and results of neuropsychological evaluation, taking into consideration any modifying circumstances, such as those described for the PhyDx. A ConsDx of dementia was based on DSM-IV (36) criteria for dementia. If the subject was considered not to have NCI or dementia, a diagnosis of MCI was assigned. The subtype of MCI (e.g., MCI amnesic-single domain) was determined by the neuropsychological diagnosis, following NACC/UDS protocol (D1 Diagnosis) nomenclature. When the consensus team was unable to agree whether the subject had a subtype of MCI or was cognitively normal, the subject was classified as normal, because most nomenclature for a cognitive diagnosis of MCI depends on a clinical diagnosis of cognitive impairment confirmed by objective deficits on neuropsychological tests. *When the consensus team was unable to agree whether the subject had a diagnosis of a subtype of MCI or dementia, the PhyDx was used for the consensus diagnosis. The rationale for this rule is that in the presence of neuropsychological deficits, additional evidence of significant functional deficits provided by the PhyDx is sufficient to meet criteria for a diagnosis of dementia and the lack of such functional deficits would place the cognitive diagnosis in the MCI domain.*
 - b. AlgDx was determined using a computational algorithm that assigned each NACC diagnosis by combining the PhyDx (NCI, MCI or Dementia) with the NPDx (NCI, naMCI, aMCI or Dementia; see Table 1). The rules used to assign these diagnoses were as follows: (A) when both the PhyDx and NPDx was NCI, the AlgDx was *NCI*; (B) when the PhyDx was NCI, but the NPDx was naMCI or aMCI, the AlgDx was NCI; (C) when the PhyDx was MCI and the NPDx was NCI, the AlgDx was NCI; (D) when the PhyDx was MCI, and the NPDx was naMCI or aMCI, the AlgDx was *naMCI* or *aMCI*; (E) when the PhyDx was dementia and the NPDx was aMCI or Dementia, the AlgDx was *Dementia*. Unusual combinations, such as a PhyDx of NCI and a NPDx of dementia, which have occurred in less than 1% of all subjects, were excluded from AlgDx and flagged for individual adjudication by the consensus team.
4. MRI Scans: MRI scans were acquired using a proprietary 3-D (volumetric) protocol on a Siemens Symphony, 1.5 Tesla machine (Iselin, NJ). Medial temporal atrophy was assessed by visual ratings of the hippocampus (HPC), entorhinal cortex (ERC), and perirhinal cortex (PRC) on a coronal brain MRI slice intersecting the mamillary bodies, using a visual rating system (VRS) that achieves high intra-rater and inter-rater reliability by providing drop-down reference images to guide the rater in using a five point scale (26). The total medial temporal atrophy (MTA) score was calculated by summing the VRS scores of HPC, ERC and PRC, for the right and left side of the brain.

5. ApoE Genotype Analysis: ApoE genotype was determined using standard methods (37). ApoE ϵ 4 frequencies were subsequently calculated for each diagnostic group.

Statistical Analysis

Inter-rater reliability was assessed using Cohen's weighted kappa (38) for: (1) physician's cognitive diagnosis (PhysDx); (2) neuropsychological diagnosis (NPDx); (3) traditional consensus diagnosis (ConsDx). Inter-rater reliability of the AlgDx was not assessed, as this is purely a function of the combined reliability of the CDR rating and NPDx. A comparison between the ConsDx and AlgDx was examined using cross tabulations and Cohen's weighted kappa to assess concordance. Association of AlgDx to MTA scores and ϵ 4 frequency, was evaluated with ANOVA and Sidak post-hoc procedure for the test of mean differences.

RESULTS

Subject Demographics

The mean age subjects in this study was 74.1 ± 6.5 years; mean education level was 13.9 ± 3.5 yrs; 57% of subjects were female; mean MMSE scores were 28.7 ± 1.6 for NCI, 26.6 ± 2.3 for na-MCI, 26.7 ± 2.5 for a-MCI and 23.0 ± 3.4 for Dementia. The majority of subjects (81%) spoke English as their primary language, while 19% spoke Spanish as their primary language.

Inter-rater Reliability of The Consensus Diagnosis: Inter-rater reliability of PhysDx and NPDx was calculated for two separate physician/neuropsychologist teams, who assessed the same 30 subjects (10 with NCI, 10 with MCI and 10 with dementia). Inter-rater reliability for PhysDx was 0.69 ($SE = .11$; *Cohen's weighted kappa*), with 70% agreement for NCI, 70% agreement for MCI, and 80% agreement for Dementia. Inter-rater reliability for NPDx was 0.88 ($SE = .07$), with 89% agreement for NCI, 91% agreement for MCI, and 88% agreement for dementia. Inter-rater reliability for ConsDx was 0.78 ($SE = .07$), with 90% agreement for NCI, 70% agreement for MCI and 80% agreement for Dementia.

Correspondence Between ConsDx and AlgDx—Of 532 subjects diagnosed by AlgDx, 305 subjects (57.3%) were classified as NCI, 52 (9.8%) as na-MCI, 87 (16.4%) as aMCI, and 88 (16.5%) as Dementia. Using ConsDx, 274 (51.5%) were classified as NCI, 160 (30.1%) as MCI, and 98 (18.4%) as dementia (for subjects diagnosed by ConsDx the MCI group was not further subdivided into naMCI and aMCI to avoid circularity. That is, once subjects were diagnosed as MCI by ConsDx, we used the same neuropsychological diagnosis to sub-divided them into naMCI and aMCI as with AlgDx). AlgDx diagnoses were confirmed by ConsDx for 88.2% of NCI, 92.3% of na-MCI, 85.1% of aMCI and 90.9% of Dementia cases. Cohen's weighted kappa for agreement was 0.84 ($SE = .02$). This indicates very high levels of agreement between the two diagnostic systems of AlgDx and ConsDx for NCI, MCI, and Dementia.

AlgDx Classification, MTA and ϵ 4 frequencies—MTA scores were available for 427 AlgDx subjects, and ApoE ϵ 4 genotype was available for 314 subjects. As indicated in Table 3, using AlgDx, post-hoc Sidak tests of means indicated subjects diagnosed with aMCI had higher MTA scores, in comparison to naMCI and NCI subjects, while subjects with Dementia had higher MTA scores, in comparison to subjects in the other three diagnostic groups [$F(3, 423) = 41.02$; $p < .001$]. Regarding ϵ 4 frequencies, there were statistically significant differences between groups [$X^2 (df = 3) = 31.33$; $p < .001$]. Subjects diagnosed as Dementia had the highest ϵ 4 frequency (34.9%), followed by aMCI (28.3%), naMCI (16.7%) and NCI (12.9%) subjects. A similar pattern of results was observed for subjects diagnosed by ConsDx. Using AlgDx, post-hoc Sidak tests of means indicated subjects diagnosed with aMCI had higher MTA scores, in comparison to naMCI and NCI subjects, while subjects with Dementia had higher MTA scores, in comparison to subjects in the other three diagnostic groups [$F(3, 451) = 35.53$;

$p < .001$]. Regarding $\epsilon 4$ frequencies, there were statistically significant differences between groups [X^2 (df= 3) = 32.58; $p < .001$]. Subjects diagnosed as Dementia had the highest $\epsilon 4$ frequency (33.3 %), followed by aMCI (27.1%), naMCI (17.5%) and NCI (12.1%) subjects.

DISCUSSION

Obvious functional deficits required for a diagnosis of dementia occur relatively late in the pathological course of Alzheimer's disease (AD). However, AD pathology is present in the brain years, or even decades before cognitive symptoms, functional impairments, and a clinical diagnosis of MCI, or dementia is established (18,19). Approximately one third of individuals with evidence of AD on autopsy (Braak Stage V or VI), were not found to be demented during life (39). Even though current criteria require the presence of dementia for a clinical diagnosis of AD, patients diagnosed with amnesic MCI, often considered a prodromal form of AD (35,40), were found at autopsy to have brain pathology indistinguishable from that observed in clinically diagnosed AD patients (41,42). As earlier diagnosis and treatment of AD is a desirable goal, it becomes imperative to develop practical, reliable, and cost efficient methods to characterize cognitive and functional changes occurring during the early stages of AD, preceding onset of dementia.

Many clinical trials and epidemiological studies use a consensus diagnosis of MCI and dementia as standard procedure in dementia research (15). However, there can be considerable variability among diagnosticians, and even single raters may vary considerably in how they determine diagnoses over a period of time. Variability in multicenter studies can occur if the raters are guided by different philosophies, criteria, or biases in combining individual elements of the ConsDx. Although it plays a critical role in multi-site and longitudinal studies, reliability of the consensus diagnosis has been assessed only infrequently (13,16). Variable factors that may influence a consensus diagnosis of MCI or dementia (9,12,13) include: (1) history obtained from subjects and/or collateral sources; (2) subject's performance on neuropsychological tests; (3) behavioral observations of the subject during the evaluation; (4) subject's pre-morbid cognitive or occupational status; (5) presence of sensory deficits affecting hearing and/or vision; (6) co-existing psychiatric conditions which may affect attentiveness, judgment and comprehension; (7) the subject's medications and current medical conditions that may impact cognition and/or attentiveness. We demonstrated high reliability for the NPDx (which is strongly based on tests scores), but only acceptable reliability for the PhyDx (which is not dependent on scores, but on overall clinical impression). As expected, the reliability of ConsDx was intermediate between that of PhyDx and NPDx.

To address the issue of variability in arriving at a consensus diagnosis, we developed a simple algorithm combining the PhyDx and NPDx in a transparent and predictable way, to yield the AlgDx. The AlgDx obviates the requirement for the physician and neuropsychologist to meet and discuss each case, so as to arrive at a ConsDx. Although all variability in achieving the NPDx and PhyDx is not avoided, the AlgDx provides clear decision rules for the disposition of each case, and a high level of concordance with the ConsDx is obtained. In fact, a major source of variance between different sites arises from the use of different neuropsychological tests and different normative data for the same tests. However, we have not addressed this issue in this study because our purpose is not to recommend a specific set of tests and normative data. Rather, our purpose is to demonstrate that given uniform use of specific tests and normative data, the diagnostic algorithm we have developed avoids much of the effort, time and unnecessary variability associated with combining the results of neuropsychological and physician evaluations to obtain a ConsDx.

The current algorithm for AlgDx suffers from a limitation in classifying subjects diagnosed as MCI, either by the physician or the neuropsychologist, but not by both. In this study we

simplified the AlgDx algorithm, so as to test its reliability and validity for only three groups of subjects, namely, NCI, MCI and dementia. We chose to classify subjects as NCI when they did not meet formal MCI criteria because either the physician or neuropsychologist had classified them as NCI. However, in future studies, if and when a “Pre-MCI” label gains acceptance, it may be desirable to include such a label in the diagnostic algorithm. In fact, NACC has recommended the use of a Pre-MCI term, namely, “Impaired-not MCI”, for subjects who meet neuropsychological criteria for aMCI or naMCI but who are classified as NCI by the physician. We propose that two Pre-MCI diagnostic entities should be recognized, namely, “PreMCI-NP” (in which the PhysDx is NCI but the NPDx is naMCI or aMCI) and “PreMCI-MD” (in which the PhysDx is MCI, but the NPDx is NCI).

Validity of the AlgDx was assessed by its correspondence to the ConsDx (Table 2), and the extent to which the AlgDx was associated with biological and genetic markers of AD, i.e., MTA scores and ApoE ϵ 4 allele frequencies (Table 3). In these respects, it is evident AlgDx corresponded well with ConsDx, when the diagnostic groups NCI, aMCI, naMCI and Dementia were analyzed. The finding that mean MTA scores for subjects assigned to groups by the AlgDx very closely resembled analogous groups assigned by the ConsDx, further supports the idea that both classification systems are equally sensitive to the biological underpinnings of neurodegenerative disease.

The rationale for using MTA scores and ϵ 4 frequencies as independent surrogate biomarkers for validating the AlgDx, is the high prevalence of AD pathology among subjects presenting with dementia or memory impairment. The vast majority of amnesic MCI and dementia cases are likely to harbor Alzheimer’s disease pathology, regardless of the clinical diagnosis they carry (39). It is also very likely that AD pathology in the medial temporal lobe drives the correlations between MTA and cognitive scores, even though non-AD pathologies may also be responsible for some of the cognitive deficits. We have previously shown that MTA scores evidence quantitative increases from NCI and naMCI, to amnesic MCI and Probable AD (26). Moreover, the severity of medial temporal atrophy evaluated by either visual rating or volumetric methods is closely related to the severity of underlying AD pathology in the medial temporal lobes, and to the Braak stage (18,20). The assessment of MTA on MRI scans should be effective in identifying AD pathology, independent of the cognitive status, i.e., among subjects with dementia, aMCI and even among cognitively normal elderly individuals.

One alternative to the AlgDx is to use the Clinical Dementia Rating scale as an algorithmic diagnostic system. However, the CDR is not a good substitute for the PhysDx, as it was not designed for diagnosing or staging pre-dementia states and reliability of the CDR in rating mild dementia has been reported to be suboptimal (12). We would anticipate that reliability of the CDR would be even worse for rating MCI and pre-MCI states. Moreover, the CDR does not incorporate formal and validated neuropsychological tests that are required for categorizing various pre-dementia states.

While concordance between the AlgDx and ConsDx was high, it was not perfect. The biggest discrepancy between the two systems was a tendency for NCI subjects diagnosed by AlgDx to be classified as amnesic MCI by ConsDx. This occurred in 10.5% of cases, reflecting the inflexibility of AlgDx criterion for MCI, which requires PhysDx and NPDx to be congruent in the MCI diagnosis. It is likely that in some cases a ConsDx diagnosis of MCI is based solely on the PhysDx or the NPDx, when the findings by the respective clinician data is particularly strong. Whether the ConsDx, as compared to the AlgDx, provides a true reflection of earlier cognitive/functional impairment or reflects false positive classification cannot be answered by the present study, although this could be further investigated by longitudinal studies of sufficiently long duration.

In summary, the AlgDx provides a simple, reliable and valid alternative to the classical consensus diagnosis. As such, AlgDx may have particular utility for assessing the effects of interventions in clinical trials. It is apparent that convenience, user-friendliness, increased reliability, brevity, and equivalent validity of the AlgDx makes it an appropriate and cost-effective method for diagnosing MCI states and early dementia, especially in multi-center longitudinal studies and clinical trials.

Acknowledgments

Research was supported by the National Institute of Aging, NIH 5R01AG020094-05 and NIH 1P50AG025711-03, & Byrd Alzheimer Center and Research Institute grants

NIA 1P50 AG025711-01 (Florida ADRC), NIA RO1 AG 30561-03

State of Florida Department of Elder Affairs

The Johnny Byrd, Sr. Alzheimer's Center and Research Institute

University of South Florida, School of Medicine, Tampa, Florida

References

1. Bowen J, Teri L, Kukull W, et al. Progression to dementia in patients with isolated memory loss. *Lancet* 1997;349:763–765. [PubMed: 9074575]
2. Grober E, Lipton RB, Hall C, et al. Memory impairment on free and cued selective reminding predicts dementia. *Neurology* 2000;54:827–32. [PubMed: 10690971]
3. Howieson DB, Dame A, Camicioli R, et al. Cognitive markers preceding Alzheimer's dementia in the healthy oldest old. *J Am Geriatr Soc* 1997;45:584–9. [PubMed: 9158579]
4. Fabrigoule C, Rouch I, Taberly A, et al. Cognitive process in preclinical phase of dementia. *Brain* 1998;121:135–41. [PubMed: 9549494]
5. Tabert MH, Albert SM, Borukhova-Milov L, et al. Functional Deficits in Patients with Mild Cognitive Impairment: Prediction of Alzheimer's Disease. *Neurology* 2002;58:758–764. [PubMed: 11889240]
6. Graham JE, Rockwood K, Beattie BL, et al. Standardization of the diagnosis of dementia in the Canadian Study of Health and Aging. *Neuroepidemiology* 1996;15:246–56. [PubMed: 8878077]
7. Luis CA, Barker WW, Loewenstein DA, et al. Conversion to Dementia among Two Groups with Cognitive Impairment. A Preliminary Report. *Dement Geriatr Cogn Disord* 2004;18:307–13. [PubMed: 15305108]
8. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* 1993;43:2412–2414. [PubMed: 8232972]
9. McCulla MM, Coats M, Van Fleet N, Duchek J, Grant E, Morris JC. Reliability of clinical nurse specialists in the staging of dementia. *Arch Neurol* 1989;46:1210–1211. [PubMed: 2818255]
10. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *Int Psychogeriatr* 1997;(Supp 1):173–176. [PubMed: 9447441]
11. Fillenbaum GG, Peterson B, Morris JC. Estimating the validity of the clinical Dementia Rating Scale: the CERAD experience. Consortium to Establish a Registry for Alzheimer's Disease. *Aging (Milano)* 1996;8:379–385. [PubMed: 9061124]
12. Rockwood K, Strang D, MacKnight C, et al. Interrater reliability of the Clinical Dementia Rating in a multicenter trial. *J Am Geriatr Soc* 2000;48:558–559. [PubMed: 10811551]
13. Schafer KA, Tractenberg RE, Sano M, et al. Alzheimer's Disease Cooperative Study. Reliability of monitoring the clinical dementia rating in multicenter clinical trials. *Alzheimer's Disease and Related Disorders* 2004;18:219–222.
14. Bennett DA, Schneider JA, Bienias JL, et al. Mild Cognitive impairment is related to Alzheimer disease pathology and cerebral infarctions. *Neurology* 2005;64:834–841. [PubMed: 15753419]

15. Beekly DL, Ramos EM, Lee WW, et al. NIA Alzheimer's Disease Centers. The National Alzheimer's Coordinating Center (NACC) database: the Uniform Data Set. *Alzheimer Dis Assoc Disord* 2007;21(3):249–58. [PubMed: 17804958]
16. Hogervorst E, Bandelow S, Combrinck M, et al. The validity and reliability of 6 sets of clinical criteria to classify Alzheimer's disease and vascular dementia in cases confirmed post-mortem: added value of a decision tree approach. *Dement Geriatr Cogn Disord* 2003;16:170–80. [PubMed: 12826744]
17. Jobst KA, Barnetson LP, Shepstone BJ. Accurate prediction of histologically confirmed Alzheimer's disease and the differential diagnosis of dementia: the use of NINCDS-ADRDA and DSM-III-R criteria, SPECT, X-ray CT, and APO E4 medial temporal lobe dementias. The Oxford Project to Investigate Memory and Aging. *Int Psychogeriatr* 1998;10:271–302. [PubMed: 9785148]
18. Bobinski M, Wegiel J, Wisniewski HM, et al. Neurofibrillary pathology - correlation with hippocampal formation atrophy in Alzheimer disease. *Neurobiol Aging* 1996;17:909–919. [PubMed: 9363803]
19. Jack CR, Dickson DW, Parisi JE, et al. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. *Neurology* 2002;58:750–7. [PubMed: 11889239]
20. Jarvenpaa T, Laakso MP, Rossi R, et al. Hippocampal MRI volumetry in cognitively discordant monozygotic twin pairs. *J Neurol Neurosurg Psychiatr* 2004;75:116–120. [PubMed: 14707319]
21. Ye S, Huang Y, Mullendorff K, Dong L, Giedt G, Meng EC, Cohen FE, Kuntz ID, Weisgraber KH, Mahley RW. Apolipoprotein (apo) E4 enhances amyloid beta peptide production in cultured neuronal cells: apoE structure as a potential therapeutic target. *Proc Natl Acad Sci U S A* 2005;102:18700–5. [PubMed: 16344478]
22. Folstein M, Folstein S, McHugh P. Mini-mental state. A practical method for grading the cognitive state of patients for the physician. *J Psychiatr Res* 1975;12:189–198. [PubMed: 1202204]
23. Fuld, PA. Fuld Object-Memory Evaluation. Illinois: Stoelting Co; 1981.
24. Brandt J. The Hopkins Verbal Learning Test: Development of a new verbal memory test with six equivalent forms. *Clin Neuropsychol* 1991;5:125–142.
25. Golden C. Stroop Color and Word Test. *J of Personality Assessment* 1975;39:502–506.
26. Duara R, Loewenstein DA, Potter E, et al. Medial Temporal Lobe Atrophy on MRI Scans and the Diagnosis of Alzheimer's Disease. *Neurology* 2008;71:1986–1992. [PubMed: 19064880]
27. Wechsler, D. The Wechsler Memory Scale-Revised. The Psychological Corporation; San Antonio, TX: 1987.
28. Spreen, O.; Strauss, E. A compendium of neuropsychological tests: Administration, norms, and commentary. 2. New York: Oxford University Press; 1998.
29. Acevedo A, Loewenstein DA, Barker WW, Harwood DG, Luis C, Bravo M. Category fluency test: Normative data for English-and Spanish-speaking elderly. *J Int Neuropsychol Soc* 2000;6:760–69. [PubMed: 11105466]
30. Wechsler, D. The Wechsler Adult Intelligence-Revised. New York: Psychological Corporation; 1981.
31. Army Individual Test Battery. Manual of directions and scoring. Washington: War Department, Adjutant General's Office; 1944.
32. Loewenstein DA, Acevedo A, Agron J, et al. The use of amnesic and nonamnesic composite measures at different thresholds in the neuropsychological diagnosis of MCI. *J Clin Exp Neuropsychol* 2007;29:300–307. [PubMed: 17454350]
33. Acevedo A, Loewenstein DA, Agron J, Duara R. Influence of socio-demographic variables on neuropsychological test performance in Spanish-speaking older adults. *Journal of Clinical and Experimental Neuropsychology* 2007;29:530–544. [PubMed: 17564918]
34. Loewenstein DA, Acevedo A, Ownby R, Agron J, Barker WW, Issacson RS, et al. Using different memory cut-offs to assess MCI. *American Journal of Geriatric Psychiatry* 2006;14:911–919. [PubMed: 17068313]
35. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med* 2004;256:183–94. [PubMed: 15324362]
36. Diagnostic and Statistical Manual of Mental Disorders DSM-IV. American Psychiatric Association; 2000.

37. Wenham PR, Price WH, Blundell G. Apolipoprotein E genotyping by one stage PCR. *Lancet* 1991;337:1158–1159. [PubMed: 1674030]
38. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psych Bull* 1968;70:213–220.
39. Snowdon DA, Greiner LH, Mortimer JA, Riley KP, Greiner PA, Markesbery WR. Brain infarction and the clinical expression of Alzheimer disease. The Nun Study. *JAMA* 1997;277:813–7. [PubMed: 9052711]
40. Dickerson BC. Pre-Dementia Diagnosis of Alzheimer's Disease: Translating Clinicobiologic Research into Practice. *Curr Psychiatr Rev* 2008;4:73–79.
41. Markesbery WR, Schmitt FA, Kryscio RJ, Davis DG, Smith CD, Wekstein DR. Neuropathologic Substrate of Mild Cognitive Impairment. *Arch Neurol* 2006;63:38–46. [PubMed: 16401735]
42. Barker WW, Luis CA, Kashuba A, Luis M, Harwood DG, Loewenstein D, Waters C, Jimison P, Shepherd E, Sevush S, Graff-Radford N, Newland D, Todd M, Miller B, Gold M, Heilman K, Doty L, Goodman I, Robinson B, Pearl G, Dickson D, Duara R. Relative frequencies of Alzheimer disease, Lewy body, vascular and frontotemporal dementia, and hippocampal sclerosis in the state of Florida brain bank. *Alzheimer Dis Assoc Disord* 2002;16:203–12. [PubMed: 12468894]

Table 1

Algorithm for Combining PhysDx with NPDx to Derive AlgDx

PhysDx	Neuropsychological Diagnosis			
	Normal	naMCI	aMCI	Dementia
Normal	Normal	Normal	Normal	MCI
MCI	Normal	MCI	MCI	MCI
Dementia	MCI	MCI	Dementia	Dementia

PhysDx = Physician's Diagnosis; NPDx = Neuropsychological Diagnosis; AlgDx = Algorithmic Diagnosis; naMCI = non-amnestic MCI; aMCI = amnestic MCI

Table 2

Agreement Between the Consensus Diagnosis and Algorithmic Cognitive Diagnosis (N= 532)

Algorithmic Diagnosis	Consensus Diagnosis		
	NCI (n = 274)	MCI (n = 160)	Dementia (N = 98)
NCI (n = 305)	269 (88.2%)	32	4
AMCI (n = 87)	1	74 (85.1%)	12
naMCI (n = 52)	2	48 (92.3%)	2
Dementia (n=88)	2	6	80 (90.9%)

aMCI = amnesic MCI; naMCI = non-amnesic MCI; parentheses represent percentage of true concordance between AlgDx and ConsDx

Table 3

MTA Scores and ApoE ε4 frequencies In Groups Diagnosed by AlgDx Method

Diagnostic Classification	AlgDx		Consensus Dx	
	Mean MTA Score	APOE (ε4%)	Mean MTA Score	APOE (ε4%)
NCI	.632 ^a (.69)	12.9%	.625 ^a (.68)	12.4%
naMCI	.628 ^{ab} (.62)	16.7%	.550 ^a (.60)	17.5%
aMCI	.944 ^b (.76)	28.3%	.939 ^b (.79)	27.1%
Dementia	1.78 ^c (1.10)	34.9%	1.73 ^c (1.1)	31.3%

Note: Means with different alphabetic superscripts compared vertically within each diagnostic subgroup are statistically significant by the Sidak test of means procedure; MTA = Medial Temporal Atrophy; AlgDx = Algorithmic Diagnosis; naMCI = non-amnestic MCI; aMCI = amnestic MCI