Table 1. Summary of studies: Methods, sample, analyses and recommendations of investigators..

Reference/ Instrument	Method of data collection/ Clinical criteria	Sampling method/Sample characteristics/ Response rate	Method of analysis: Results by group / Validity (Criterion)	Recommendations of investigators
Anthony, et al., 1982 MMSE	Non-psychiatrist investigators administered the MMSE, obtained social and demographic information within 24 hours of admission; psychiatrist (blinded to MMSE score) tested for dementia and delirium using Folstein and McHugh (1976) criteria (based on DSM-III). MMSE was independent of clinical evaluation. Those with discrepant MMSE and clinical evaluations, were re-examined on day 2 by psychiatrists blinded to group. Cut-score of 23/24 used because recommended in most publications (0-23 impaired).	Consecutively admitted patients to the general medical ward of Johns Hopkins Hospital, Baltimore, Md. N=97; 23% White, 77% Black; 63% female. age range / %: 20-29 / 16; 30-39 / 13; 40-49 / 18 50-59 / 13; 60-69 / 12; 70-79 / 23; 80+ / 6 Education: ≤8 years: 47% Response rate: Initial: 98%	Comparison of sensitivity/specificity at the standard cut-score. At 23 cut-score: 87% sensitivity; 82% specificity Subgroup specificities: Blacks vs Whites: 78.2% vs 94.7% women vs men: 76.6% vs 92.6% ≤8years education vs 9 or more years: 63.3% vs 100% ≥60 years old vs <60 years old: 65.2% vs 92.0%	Sensitivity levels support the MMSE as a *screen* for dementia and delirium. It should not be used as the sole criterion for diagnosis due to the high false-positive ratio, especially for the elderly and poorly educated. Modifications in scoring to improve sensitivity and specificity were tested (lower cut-score, eliminate difficult items): these not satisfactory due to the decreases in sensitivity that accompany increases in specificity.
Bohnstedt, et al., 1994 MMSE	Subjects diagnosed by interdisciplinary teams. Staff completed neuropsychologic tests on a variety of patients and collected clinical diagnostic, neuroimaging, demographic, and neuropathologic data. DSM-III-R criteria for clinical dementia; NINCDS-ADRDA criteria for AD. Standard cut-score of ≤23 used. By reducing the cut-score to ≤19 for Blacks (B) and Latinos (L), sensitivity and specificity values are closer to those for Whites. Diagnoses were NOT made independent of MMSE.	8 Alzheimer's Disease Diagnostic and Treatment Centers throughout Ca. N=1888; 83% White (W), 10% Black, 7% Latino; 67% female; median sample patient age at enrollment into the centers=76 years.	Comparison of sensitivity/specificity/PPV/NPV at standard and adjusted cut-scores. % demented according to MMSE (≤23) / clinically: White: 72 / 93; Black: 90 / 96; Latino: 88 / 96; MMSE: p<.001 between groups; Clinically: p>.10 between groups MMSE vs. Clinical diagnosis., at cut points 23 / 19 for subgroups: Sensitivity- B: 93.0 / 78.0%; L: 90.0 / 75.4%; W: 76.6% Specificity- B: (n=8) 87.5 / 100%; L: (n=5) 60.0 / 100%; W: 93.6% PPV- B: 99.4 / 100%; L: 98.3 / 100%; W: 99.4% NPV- B: 35.0 / 16.3%; L: 18.8 / 13.5%; W: 23.1%	Consider lowering the cut-score for Blacks and Latinos to 19 (from 23) in order to provide a more accurate estimate of cognitive functioning compared to a diagnosis of dementia.
Fillenbaum, et al., 1990 Katzman OMC, MSQ, MMSE, Storandt, Iowa, Kendrick, SPMSQ	SPMSQ score used as the sampling screen, and to aid in the diagnosis of dementia. Administered to Duke EPESE cohort, and sample classified by score (impaired, marginal, unimpaired). Cut-score determined so that 10% of an older community population would be impaired (Pfeiffer, 1975). Clinical diagnosis: semi-structured interview by a neurologist using DSM-III and NINCDS/ADRDA criteria to determine dementia status, battery administered by 2 middle-aged, White college graduates, mental status information from close family member. 22 of 26 dementia cases confirmed through reevaluation at Duke Memory Disorder Clinic using MRI and lab tests.	Used random sampling with replacement, matched for race, age, SPMSQ score, & sex. Numbers weighted to allow generalization to entire sample. Duke EPESE, 5 counties in North Carolina; N=164; Black: n=83, White: n=81; 57.9% female; demented: n=26; non-demented: n=132; 1 person untestable, 5 with unclear diagnoses; minimum age=65. Response rate: 67% (33% refusal)	Comparison of sensitivity/specificity at the standard cut-scores. Sensitivity / specificity by race (weighted to represent all community Black & White residents): Black / White / Overall (Total) OMC: 100 / 38.2% — 100 / 78.5% — 100/58.1% MSQ: 100 / 70.5% — 100 / 96.0% — 100 / 83.1% MMSE: 100 / 26.4%; — 100 / 69.1% — 100 / 48.8% (uncodable) 5.1% — 3.5% — 4.4% Storandt: 100 / 42.0% — 100 / 69.2% — 100 / 56.6% (uncodable) 4.0% — 3.1% — 3.6% Iowa: 100 / 26.4% — 100 / 69.1% — 100 / 48.8% (uncodable) 5.1% — 3.5% — 4.4% Kendrick: 64.8 /92.4% — 77.6 /97.2% — 66.6 /94.7% SPMSQ: 89.6 /89.9% — 100 / 90.0% — 91.1 /90.0% Follow up study of Fillenbaum (1998) used ROC curve: cut-score at the posterior probability of .465. 69% sensitivity; 92% specificity	Most tests showed a high 'false-positive' rate, particularly for Blacks. This limits the usefulness of these tests for large scale screening for prevalence. Briefest measures with the easiest scoring systems were the most accurate- SPMSQ one of the best. SPMSQ and Storandt least biased for race and education. Storandt and Iowa valuable in clinical settings

Reference/ Instrument	Method of data collection/ Clinical criteria	Sampling method/Sample characteristics/ Response rate	Method of analysis: Results by group / Validity (Criterion)	Recommendations of investigators
Fuh, et al., 1995 IQCODE, MMSE-CE	IQCODE administered to an informant by trained research assistants who were medical or nursing students blinded to dementia status. Dementia status determined by a physician according to DSM-III-R criteria based on semistructured interview and testing, neurologic exam, and standardized assessment. Clinical Dementia Rating (CDR) level measured. Family member interviewed for history, daily activities, and social functioning when dementia was suspected.	Community sampling frame: all those aged 50 and over according to 1990 housing registration records in the township of Kin-Hu (Kinmet, west of Taiwan); demented: outpatients at a Veterans General Hospital, or those identified from community sample (n=16). Random sample stratified by age and sex. N=460; overall: 49% female, demented population: 36% female (n=22), control population: 51% female (n=203); n=399 non-demented community elderly (control), n=61 demented patients. age: control-68.1±11.1; demented- 73.3±8.3; education: control- 1.6±3.0; demented-6.6±6.3. (63% had 0 years of formal education) Response rate: Non-demented community residents: 49.1%	IQCODE: area under ROC curve (AUC)= 91.3% (SE=2.4%); at cutscore $\geq$3.4: 88% specificity; 89% sensitivity; PPV=45%; NPV=99%. 17 item IQCODE: AUC=91.1% (SE=3.1%) MMSE: AUC= 84.0% (SE=3.1%); at cut-score $\leq$20: 72% sensitivity; 81% specificity; PPV=30%; NPV=96%. {p=.01 (z=2.57) between measures}	IQCODE has high internal consistency, score has little to no association with education, age or gender, high association with cognitive abilities. It effectively discriminates between those with and without dementia. IQCODE better in terms of AUC than MMSE in detecting dementia among a sample of individuals with low education. IQCODE can be shortened from 26 to 17 items for predominantly illiterate Chinese elders.
Gurland, et al. (1992) Kahn Golfarb MSQ; Short Portable MSQ; CARE-Diagnostic; Blessed MIC; MMSE	Cases initially identified based on physician classification of possibly, probably, or definitely organic brain syndrome. Most subjects were referred to a clinical evaluation team who used a structured medical and neurological examination, assessment of extrapyramidal signs and other involuntary movements, a history, functional evaluation, psychiatric evaluation, and a battery of neuropsychological tests. Diagnoses were assigned at a conference attended by the examining research neurologist, and physician. A case was defined by criterion diagnosis of dementia according to DSM-III-R criteria (Gurland 1995).	Subjects recruited from outpatient primary care clinics (AIM) at Columbia Presbyterian & two local senior centers, and a local nursing home. N=550; Black: n=182, Latino: n=184, White: n=184. age range (%): 65-74, 75-84, 85+; education: (%): 0-6, 7-11, 12+ B 41.8 39.6 15.9 22.5 45.6 26.4 L 50 34.8 13.6 55.4 30.6 11.4 W 17.4 41.3 36.4 12.0 33.9 43.7 Response rate: 90%	Sensitivity set at 90% or better. Positive classification using cut-scores without adjustments (%): B L W T CARE Diagn [cut-score 0-6/7+]: 11.5 15.2 24.5 17 KG- MSQ [cut-score 0-2/3+]: 20.3 28.8 37.0 28.6 SP MSQ [cut-score 0-2/3+]: 28 38 36.4 34.2 Blessed MIC [cut-score 0-7/8+]: 34.1 55.4 41.8 43.8 MMSE [cut-score 0-6/7+]: 57.1 64.7 52.2 58.1 No. of false positives using criterion diagnosis: CARE Diagn [cut-score 3+] 5 20 2 27 KG- MSQ [cut-score 1+] 12 32 10 57 SPMSQ [cut-score 2+] 8 24 5 37 Blessed MIC [cut-score 7+] 7 34 2 43 MMSE [cut-score 7+] 11 36 7 54	Conflicts between scales can be partially resolved through adjustments of the cut-scores for determination of a possible case. Efforts should be made to reduce discrepancies between methods introduced by varying sensitivities (and specificities).

Reference/ Instrument	Method of data collection/ Clinical criteria	Sampling method/Sample characteristics/ Response rate	Method of analysis: Results by group / Validity (Criterion)	Recommendations of investigators
Hall, et al., 1996 CSI'D'	Pilot study conducted to determine cut-score. Cut-score was selected at 100% sensitivity, and 89% specificity for dementia (Hall 1993). Sampling: Indiana- simple random sample of 60% of homes in 29 contiguous tracts in Indianapolis, IN with an avg of 80%Black residents; Ibadan- a census to enumerate all households with residents aged 65 and over (total population survey), door to door screen. All of poor performance, 50% of intermediate performance, & 5% of good performance (75% aged ≥75) offered clinical assessment. Clinical assessment of selected subjects by senior neurologists, & psychiatrists for DSM-III-R & ICD10 criteria, CDR for severity. Clinical exam includes CERAD-NB, CAMCOG, CT scans, relative interview, neurological assessment, lab tests. Consensus diagnosis.	Indiana (IN): n=2212; Ibadan (Ib): n=2494; Idakan area of Ibadan, mainly Yoruba. IN: 65% female; Ib: 65% female; age: IN: 74±7; Ib: 72.3±7.5. Literacy: IN: 97.9%; Ib: 15.2%. education: IN: 9.6±3.1; Ib: .8±2.3. Subjects with clinical assess: IN demented: n=65; IN non-demented: n=286; Ib demented: n=28; Ib non-demented: n=395. Response rate**:** IN: 85.6%; Ib: 98.4% Relative interview: IN: 67.8% Ib: 100%	Combined sites: 87.02% sensitivity: 83.12% specificity (Standard errors: 6.76%, .57%) if applied to population age ≥65, with typical dementia prevalence of 8%: PPV=.31; NPV=.99 area under ROC for discriminant function: IN: .98; Ib: .90 area under ROC for cognition only: IN: .94; Ib: .84	The screening instrument shows minimal cultural bias, and performed well in both cultures. The authors do not adjust screen scores for education because it is a significant risk factor for Alzheimer's Disease (AD).
Hall, et al., 1993 CSI'D'	Sampling: English (Eng) sample taken from the database for Manitoba's health insurance system, stratified by age to oversample older age groups. The proportion of each age group was based on prevalence rates of dementia- the sampling fractions were roughly proportional to these rates. Stratified by sex to reflect the distribution in Winnipeg. Random selection within age/sex strata. All Cree aged 65 and older were eligible for participation. Pretest to determine cut-score at 100% sensitivity. One week interviewer training; interviewers for Cree originate from Nelson and Norway Houses, interviewers for English were Winnipeg residents; clinical evaluations for Cree and English speakers made by same neurologists & psychiatrists (at least 2 saw each subject); diagnosis made at consensus conference by 4 physicians blind to screen results; dementia status based on DSM-III-R criteria, mild dementia criteria of Roth, AD diagnosis by NINCDS/ADRDA criteria.	Cree: N=198; community: n=192; nursing home: n=6; clinical assessment subset: N=91; cognitive and informant sections complete: N=171; non-demented: n=165; demented: n=6. Eng: N=252; community: n=227; nursing home: n=25; clinical assessment subset: N=67; cognitive and informant sections complete: N=216; non-demented: n=200; demented: n=16; two geographically remote Cree reserves; Manitoba & Eng-speaking, non-Indian residents of Winnipeg; Cree: older subjects oversampled; Cree: 52% female; Eng: 56% female. Age range: 65-69: n=68; 70-74: n=52; 75-79: n-32; 80+: n=46. Eng: 65-69: n=21; 70-74: n=35; 75-79: n=55; 80+: n=141. Years of education: Cree: 0: 47%; 7-11: 7%; 12: 1%; Eng: 0: 8%; 7-11: 45%; 12: 13%. Response rate: Cree: screening: 92%, clinical: 94%. Eng: screening: 70%, clinical: 86%.	Using cut-score obtained in pretest: discriminant function: 100% sensitivity, 79% specificity, no false negatives. Total cognitive score had the lowest specificity: 60% specificity, (95%CI: 55-65%) combining cognition without calculation+ informant score+education, best specificity: 89% (86-92%). Area under ROC curve (using cognitive score without calculation and informant score)=.976 Note: For more specificity data see cited article table 5, page 9- has cognitive score with & without calculation items & education & informant score- for several combinations of variables.	Construction of comparative cognitive scales proved more difficult than construction of an informant interview. This new instrument exhibits evidence of validity and reliability. It may be useful in other cross-cultural studies.

Reference/ Instrument	Method of data collection/ Clinical criteria	Sampling method/Sample characteristics/ Response rate	Method of analysis: Results by group / Validity (Criterion)	Recommendations of investigators
Hendrie, et al., 1995 CSI 'D'	In Indiana, battery administered by trained technicians, and in Ibadan by faculty members; 40% of clinical evaluations done in home; median time between screen and evaluation: IN: 5 months, Ibadan: 10 months. Sampling, cut-score determination, case definition: see Hall 1996. 60% simple random sample.	Nursing home: n=106 screen & clinical evaluation; IN community: n=2212 screen, n=351 clinical evaluation; Ibadan (Ib): n=2494 screen, n=423 clinical evaluation. Community and nursing home dwellers in 29 census tracts in Indianapolis, IN and community residents (mainly Yoruban) in Ibadan, Nigeria. Nursing home: 66% female; IN community & Ib: 65% female. $\bar{x}$ age: nursing home: 80.7±9.0; IN community: 73.9±7.0; Ib: 72.3±7.5; $\bar{x}$ education: IN: 9.6±3.1; Ib: .8±2.3; Literacy rate: IN: 97.9%; Ib: 15.2%. Response rate: IN: 85.7%; Ibadan: 98.4%.	Dementia: good screen performance: IN: 1.0%; Ibadan: 1.7%; intermediate screen performance: IN: 6.4%; Ibadan: 2.1%; poor screen performance: IN: 33.9%, Ibadan: 11.2% good / intermediate&poor performance (2 groups): 87% sensitivity; 83.1% specificity.	
Heyman, et al., 1991 SPMSQ	Four stage sampling design for initial sample, here: random selection by trichotomized SPMSQ score, race (Black or White), sex, and 5-year age group. Tried selecting the same number of Black men as White men, and as many Black women as White women. Demographics, health characteristics, & SPMSQ collected at baseline; semistructured interview administered by a neurologist; diagnosis based on DSM-III, and NINDS/ADRDA, severity based on CDR; medical, and psychiatric history obtained; brief physical and neurologic exam; SPMSQ screen included in diagnosis.	N=164; Duke EPESE; minimum age: 65 Response rate: 74%	27% of White subjects with impaired SPMSQ were diagnosed as demented by neurologist; 44% of Black subjects with impaired SPMSQ found demented by neurologist; of the 26 found impaired by neurologist: 19 scored as impaired on SPMSQ (73%), 4 marginal (1 point above cutoff; 15%), 3 unimpaired (12%); Of these 26: according to medical records, 13 (50%) had prior diagnosis of dementia, 7 had been referred to a memory disorders clinic for a second evaluation.	

Reference/ Instrument	Method of data collection/ Clinical criteria	Sampling method/Sample characteristics/ Response rate	Method of analysis: Results by group / Validity (Criterion)	Recommendations of investigators
Hill, et al., 1993 CMMS	Investigators from the Shanghai Mental Health Center screened in subject homes; clinical examination for those below the cut-score, and 5% of those above; single-stage cluster sampling, random probability sample: in 1/3 of all neighborhoods all those aged 55 and over were included, in 1/3 all those aged 65 and over were included, and in 1/3 all those aged 75 and over were included; age and education adjusted cut-scores used (per Kittner et al 1986); a cut-score for the predictive equation with approximately equal sensitivity and specificity chosen, operating characteristics as described by Vecchio (1966); presurvey cutoff of ≤20 to determine direct or proxy interview; dementia determined by history, NINCDS-ADRDA criteria, Hachinski score, DSM-III criteria & neuropsychological instruments; clinical diagnosis made independently by 3 psychiatrists blind to CMMS score and reviewed by a US clinician and neuropsychologist.	N=643; Jing-An district of Shanghai; 159 cases of dementia according to DSM-III, approximately 2/3 of these met AD criteria Response rate: 85.6%	CMMS: 69.5% sensitivity, 90.2% specificity, PPV=62.4%, NPV= 92.7%; CMMS+ function measures: 92.4% sensitivity, 93.3% specificity, PPV= 76.4%, NPV= 98.1%; Education adjusted CMMS + function measures: 92.4% sensitivity, 92.0% specificity, PPV= 72.9%, NPV= 98.1%. * predictive value estimates do not have the usual meaning since there was no attempt to have this sample emulate the population prevalence rate for dementia. Note: For other results see cited article table 3, pg. 142. Different combinations of variables (e.g., history, IADL, POD)	Findings suggest the possible usefulness of combining the CMMS, functional scales, and history in determining dementia status. Using DSM-III criteria, subjects with no education or 6 years or less have a significantly greater risk of being diagnosed with dementia than those with more than 6 years of education.
Li, et al., 1989 MMSE, CRBRS	All persons aged 60 and older in all households in 4 geographically representative residents' committees in the West District of Beijing were contacted for the survey. Pretest to determine cut-score for MMSE. All those who scored ≤17 on the MMSE, and a random sample of 5.5% of those who scored above 17 received clinical evaluation from a psychiatrist; final diagnosis was made according to operating diagnostic criteria, a restricted DSM-III definition of dementia. The clinical exam consisted of the Geriatric Mental State Exam, the Dementia Differential Diagnostic Schedule. All those who scored ≤2 on CRBRS received clinical evaluation from a psychiatrist, information received from informant.	Pretest: N=137; age range: 50-89; 44% female; in- or out-patients of the Institute of Mental Health or nursing home residents. Tested on 17 demented, 47 other mental disorders, & 73 normal patients; study N=1090; 53.7% female; age groups: 60-64: 34.4%, 65-69: 27.3%, 70-74: 21.4%, 75-79: 10.0%, 80-84: 5.2%, 85+: 1.7%; 33.7% illiterate. Response rate: 81.9%	Pretest to determine cut-score, at 100% sensitivity, 89% specificity. 98.3% took the MMSE. N=42 scored ≤17, 6 (14% of the 42) clinical diagnosis of moderate or severe dementia, 4 (9.5% of 42) diagnosed with potential dementia, 1 (2.4% of 42) refused exam, 31 (73.8% of 42) non- dementia cases; for 57 subjects scoring ≥17 with clinical evaluation: no cases of dementia were found; CRBRS: 8 (44.4% of 18) clinical diagnosis of moderate or severe dementia, 1 (5.5% of 18) diagnosed with possible dementia	MMSE is simple to use and acceptable to the Chinese population. MMSE score was significantly correlated with age, education level, and level of physical ability, perhaps different cut-scores for different education levels should be developed

Reference/ Instrument	Method of data collection/ Clinical criteria	Sampling method/Sample characteristics/ Response rate	Method of analysis: Results by group / Validity (Criterion)	Recommendations of investigators
Loewenstein, et al., 1995 OME, MMSE	Normal subjects (MMSE ≥23) recruited from the community as part of an ongoing longitudinal study. Entire OME 5 trials takes 20 minutes to administer, first trial alone takes 5 minutes. Complete neurological, medical, psychiatric, and neuropsychological evaluation, MRI, NINCDS/ADRDA criteria for AD. MMSE cut-scores derived on the basis of normal control scores applied to those with a clinical diagnosis of AD but with an MMSE score of 18 or above, equivalent to cut-scores derived in the ECA investigation.	N=191; control (MMSE ≥23, no reported memory deficit): n=53 recruited from the community, English (Eng) speakers: n=30, Spanish (Sp) speakers: n=23; AD: n=138, Eng speakers: n=111, Sp speakers: n=27; recruited from Memory Disorders Clinic. Virtually all Spanish speakers of Cuban descent. $\bar{x}$ age: Sp AD: 72.0±7.7; Eng AD:77.81±6.8; Sp control: 71.91±4.4; Eng control: 75.59±5.6. $\bar{x}$ education: Sp AD: 10.25±4.8, Eng AD: 12.03±3.8, Sp control: 13.43±5.2, Eng control: 14.37±2.8	The terms sensitivity and specificity apply to well-defined AD and normal control groups. OME: (29 cut-score): Sensitivity- Spanish AD: 95.9%, English AD: 95.5% (for those age ≤79 sensitivity: 94.1%); Specificity- Spanish control: 100%, English control: 96.7%; Sensitivity trial 1 (<5)- English: 92.8% (for those age ≤79 sensitivity: 91%), Spanish: 100%; Specificity trial 1- total: 91.6%, Spanish only: 87%.: MMSE: (23 cut-score): Sensitivity- English: 77.8%, Spanish: 73.1%; at 24 cut-score: Sensitivity- English: 81.0%, Spanish: 88.9%	OME sensitivities greater than MMSE, and has several advantages: easy to administer, uses common items, easily translated, limits confounding due to poor auditory or visual perception. Cut-scores of 29 for total OME, and 5 for the first trial are culturally appropriate
Morales, et al., 1995 S-IQCODE, SS-IQCODE, MMSE	Sample stratified by age and sex randomly selected from the census of three districts in Madrid, Spain. Door-to-door screening, including MMSE. Interview with informant who lived with subject. Dementia diagnosed in the study sample by a neurologist using neurological and neuropsychological assessment, CAT scan, lab tests, diagnosis according to DSM-III-R criteria. Takes approximately 10 minutes to administer. Refusals were assessed by mailed questionnaires and/or medical records.	Whole sample: N=352; study sample (those who participated in phase 2 of the study without those excluded): N=68 (no significant differences between whole and study samples in sex, age, # of chronic diseases, # of drugs, GDS Scale. MMSE higher). Demented: n=7; non-demented: n=61. Madrid, Spain. Study sample: 51.5% female; $\bar{x}$ age: total population: 73.1±5.2, demented: 74.8±5.5, non-demented: 72.9±5.2. education: 7.4% illiterate; 92.6% had at least primary schooling. Response rate: 73% (whole sample for initial interview). All screened invited for clinical evaluation: 52.5% response rate	S-IQCODE: 86% sensitivity, 92% specificity, PPV=.54; NPV=.98; accuracy=.91. Item score and total score correlation using ROC curves: r ranges from .4120 (remember present date) -.8558 (has his/her intelligence changed) for items selected for SS-IQCODE (some items with r in range not selected). SS-IQCODE: 86% sensitivity; 91% specificity; PPV=.50; NPV=.98; accuracy=.90. MMSE: 57% sensitivity; 84% specificity; PPV=.29; NPV=.94; accuracy=.81.	S-IQCODE valid for detection of mild dementia. It is not affected by the most frequent confounding variables
Murden, et al., 1991 MMSE	Administered to subject by a geriatrician or physician's assistant trained in geriatrics with the collaboration of a geriatrician. Dementia diagnosed using 4 criteria: history of memory decline & other cognitive function, no impaired consciousness, abnormal MMSE (≤23), dementia presence according to examiner. Those with uncertain diagnoses followed up with MMSE and clinical exams a minimum of 6 months later in order to determine certain diagnoses. MMSE performed by initial clinical examiner.	N=358; Demented: n=110 (100% Black); Non-demented: Black: n=148, White: n=100; Kings County Hospital outpatient clinic and geriatric clinic at Bellevue Hospital (NY). 76% female; age range: 60-99; Black: n=258, White (Non-Latino): n=100; Nondemented patients- ≥9 years ed: White: n=74, Black: n=38; <9 years ed: White: n=26, Black: n=110. Response rate: (Follow up) 84.1%	MMSE sensitivity and specificity lowering cut-score to 17 from 23 for Black subjects: (23 ⇨17): ≥9 years education: sensitivity- 93%⇨52%, specificity- 100⇨100%; <9 years education: sensitivity- 98%⇨81%, specificity-75%⇨100%	Reduce the cut-score to 17 to reduce false labeling of dementia in Blacks with less than a 9th grade education. The authors are working on a revised MMSE that may be more appropriate for people of all education levels.

Reference/ Instrument	Method of data collection/ Clinical criteria	Sampling method/Sample characteristics/ Response rate	Method of analysis: Results by group / Validity (Criterion)	Recommendations of investigators
Pittman, et al., 1992 Diagnostic paradigm (battery of neuropsycho-logical tests)	Subjects selected from volunteers who met inclusion criteria for a study that aims to attain a high degree of accuracy and consistency in the diagnosis of dementia in heterogenous populations (Stern, 1992). Cut-scores based on performance of 172 previously studied subjects (Stern, 1992). Physician diagnosis based on semi-structured medical and psychiatric history, physical & neurological exam, short Blessed, and 3 ADL measures. Diagnostic paradigm diagnosis based on DSM-III-R criteria (poor performance on 2 of 3 areas assessed).	N=430; 37.5% White (W), 29.1% Black (B), 33.4% Latino (L); Washington Heights & Inwood (northern Manhattan); $\bar{x}$ education: W: 13 years, B: 9.9 years, L: 7.2 years; <6 years education: 23.5%, college education: 25.6%. Response rate: 6.4% of subjects refused to complete one or more of the tests, 14.6% unable to complete one or more of the tests (Stern, 1992) Note: partial information included.	Categories for each method: 0=normal; 1=cognitively disordered; 2=demented 9 Subjects (2%) with maximum disagreement between methods (7 cases physician found normal, paradigm found demented); Physician found: %(n): 0=68.6% (295), 1=21.6% (93), 2=9.8% (42) battery found: 0=60.5% (260), 1=29.1% (125), 2=10.5% (45) 71.8% (n=309) agreement between physician & battery: 0=53.5% (230), 1=12.3% (53), 2=6.0% (26). among cases of disagreement: 63.6% (77 of 121) more impaired by battery- 71% (86 of 121) of these in 0 vs 1 categories, 67.4% (58 of the 86) rated 0 by physician. & 1 by battery. 21.5% (26 of 121) 1 vs 2, 46% (of 26) 2 by battery & 1 by physician; 54% (14) 2 by physician and 1 by battery	Further refinement of the paradigm could raise reliability to an acceptable level (reliability probably underestimated since those who did not represent a challenge for the diagnostician were excluded from the sample). A composite diagnosis (paradigm and diagnostic) is most reliable.
Teng et al., 1994 Cognitive Abilities Screening Instrument (CASI). (It includes: HDSS-CE; MMSE-CE; 3MS-CE)	CASI-E & CASI-J were simultaneously tested in LA and Seattle & in Osaka and Tokyo. No match on age and education done for patients and controls. All patients met the DSM-III-R criteria for dementia; the diagnosis was made by physicians based on history, functional assessment, clinical examinations, and neuropsychological testing results excluding the scores on the CASI. LA: patients from geriatric neurobehavioral & dementia referral center, control: spouses, community residents Osaka: patients from neuropsychology clinic of national medical research center; control also patients at clinic. Seattle: patients & control subscribers of a health maintenance program. Tokyo: patients from outpatient clinic of a geriatric hospital; control: students at community college for aged.	LA: patients: n=62, control: n=50; Osaka: patients: n=23, control: n=61; Seattle: patients: n=71, control: n=86; Tokyo: patients: n=52, control: n=38 % Female / Age: $\bar{x}$, sd: LA=58% / D: 74.2, 7.8; C: 70, 13.3 Osaka=27% / D: 70.7, 8.4; C: 63.6, 9.7 Seattle=63% / D: 78.3,6.0; C: 77.5, 6.4 Tokyo=66% / D: 78.1, 6.7; C: 73.3,4.9 Education: $\bar{x}$:, sd: LA= D: 12.1,4.1; C: 12.7, 3.5 Osaka=D: 11.2, 4.0; C: 11.9, 3.4 Seatle=D: 12.5, 3.0; C: 13.8, 2.9 Tokyo=D: 8.4, 2.9; C: 11.1, 4.3	USA samples: [LA] Cut-score / sens / spec; [Seattle] Cut-score / sens / spec CASI 78 .91 .91; 86 .94 .94 CASI-Short 23 .95 .94; 25 .94 .94 HDSS-CE 26 .84 .85; 28 .86 .88 MMSE-CE 22 .86 .86; 24 .91 .93 3MS-CE 76 .92 .92; 83 .94 .94 Japan samples: [Osaka] Cut-score / sens / spec; [Tokyo] Cut-score / sens / spec CASI 71 .95 .94; 76 .93 .93 CASI-Short 21 .89 .88; 21 .89 .92 HDSS-CE 24 .93 .93; 25 .92 .92 MMSE-CE 23 .92 .92; 21 .94 .94 3MS-CE 68 .93 .94; 74 .93 .93	In assessing dementia, care should be taken to consider the “ecological validity” of the test for the study population; test items, cutoff points, and even cognitive domains may need to be adjusted accordingly.
Tsai, et al., 1989 HDS	Can be completed in approximately 10 minutes. Scores of 31-32.5= nondemented, 22-30.5= subnormal, 10.5-21.5= early dementia, 0-10= demented. Sample: all inhabitants of specific urban and rural communities in Shanghai aged 60 years and older (over 60 according to text). Instrument administered by 2 psychiatrists and 2 nurses, clinical evaluation by the psychiatrists. DSM-III criteria for dementia.	N=2573; 63.3% female; According to score: 35.3% non-demented, 46.3% subnormal, 16.1% early dementia, 2.3% demented; age group/n: 60-64/384, 65-69/792, 70-74/560, 75-79/441, 80-84/307, 85+/89; education level: illiterate: 67.1% Response rate: 84%	96.3% diagnostic agreement between clinical and instrument. 80.2% sensitivity, 99.4% specificity (Note: determined from data presented). HDS score ≤15: 100% diagnosed as demented, score of 16-21: 94.6% diagnosed as demented, score of 22-27: 8.7% diagnosed as demented, score of 28+: 0% diagnosed as demented. HDS level / % diagnosed with dementia: normal / 0%, subnormal / 7.1%, early dementia / 97.1%, demented / 100%.	Assessment of dementia by HDS corresponds well with clinical evaluation. It is a useful national and international assessment tool.

Reference/ Instrument	Method of data collection/ Clinical criteria	Sampling method/Sample characteristics/ Response rate	Method of analysis: Results by group / Validity (Criterion)	Recommendations of investigators
Wilder, et al., 1995 MSQ; SPMSQ; CARE-Diag; Blessed MIC; MMSE	Subjects were screened and given a dementia/non-dementia diagnosis from a North Manhattan probability sample of 65 and older residents (transcultural pilot study (TPS) & NMAP Reporting Registry). Criterion diagnosis: DSM-III-R criteria for dementia by clinical core team of neurologists, physicians and neuropsychologists.	N=795; Black: n=299, Latino: n=355, White: n=136. age range (%): 65-74: B: (25.1); L: (41.7); W: (17.6) 75-84: B: (46.2); L: (36.6); W:(30.9) 85+: B: (28.8); L: (21.7); W:(51.5) education: (%): <5: B: (17.6) L: (52.8); W: (11.4) 5-11: B:(60.0); L: (37.1); W: (40.9) ≥ 12: B:(22.4); L: (10.1); W: (47.7)	Differences in area under the curve (AUC) by race-ethnicity/ by education: Blessed MIC: .057 / .127; CARE Dementia .040 / .094; Kahn Goldfarb MSQ: .044 / .112; MMSE: .080 / .131; SPMSQ: .068 / .176; CARE Homogeneous: .043 / .115; "Culture fair": .048 / .118 Specificity when sensitivity set at 90% derived from ROC curves with criterion diagnosis: Racial-ethnic groups / Educational groups B L W T <5 5-11 ≥12 Blessed MIC: . 72 .47 .74 .60 .33 .55 .80 CARE Dementia .49 .58 .68 .58 .45 .53 .78 Kahn Goldfarb MSQ: .57 .55 .62 .57 .42 .54 .77 MMSE: .54 .38 .59 .44 .24 .51 .74 SPMSQ: .46 .42 .62 .48 .27 .46 .74 CARE Homogeneous: .60 .59 .71 .62 .43 .62 . 77 "Culture fair": .60 .53 .69 .58 .40 .57 .83	Shorter scales perform as well as longer ones, are more consistent across cultural and educational groups and can be more easily modified to improve performance in culturally diverse populations.