Dementia diagnosis in developing countries: a cross-cultural validation study

*Martin Prince, Daisy Acosta, Helen Chiu, Marcia Scazufca, Mathew Varghese, for the 10/66 Dementia Research Group**

Summary

Background Research into dementia is needed in developing countries. Assessment of variations in disease frequency between regions might enhance our understanding of the disease, but methodological difficulties need to be addressed. We aimed to develop and test a culturally and educationally unbiased diagnostic instrument for dementia.

Methods In a multicentre study, the 10/66 Dementia Research Group interviewed 2885 people aged 60 years and older in 25 centres, most in Universities, in India, China and southeast Asia, Latin America and the Caribbean, and Africa. 729 had dementia and three groups were free of dementia: 702 had depression, 694 had high education (as defined by each centre), and 760 had low education (as defined by each centre). Local clinicians diagnosed dementia and depression. An interviewer, masked to dementia diagnosis, administered the geriatric mental state, the community screening instrument for dementia, and the modified Consortium to Establish a Registry of Alzheimer's Disease (CERAD) ten-word list-learning task.

Findings Each measure independently predicted a diagnosis of dementia. In an analysis of half the sample, an algorithm derived from all three measures gave better results than any individual measure. Applied to the other half of the sample, this algorithm identified 94% of dementia cases with false-positive rates of 15%, 3%, and 6% in the depression, high education, and low education groups, respectively.

Interpretation Our algorithm is a sound basis for culturally and educationally sensitive dementia diagnosis in clinical and population-based research, supported by translations of its constituent measures into most languages used in the developing world.

Lancet 2003; **361:** 909–17
See Commentary page 888

*Members who contributed as investigators are listed at end of article

Institute of Psychiatry, King's College, London, UK (Prof M Prince MD)**; Department of Psychiatry, Universidad Nacional Pedro Henriqueuz Ureña, Santo Domingo, Dominican Republic** (D Acosta MD)**; Department of Psychiatry, The Chinese University of Hong Kong, Shatin, Hong Kong** (Prof H Chiu FRCPsych)**; Institute of Psychiatry, University of Sao Paulo, Sao Paulo, Brazil** (M Scazufca PhD)**; and National Institute of Mental Health and Neurosciences, Bangalore, India** (M Varghese MD)

Correspondence to: Prof Martin Prince, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK (e-mail: m.prince@iop.kcl.ac.uk)

Introduction

The 10/66 Dementia Research Group was established to encourage good quality research into dementia in developing countries.[1] Differences in prevalence and incidence between populations can be interpreted with confidence only if they result from common, standardised procedures that are culturally sensitive.[2] Whichever criteria are used, three conditions must be met for a diagnosis of dementia to be made: decline in at least two domains of cognitive function, including memory; interference with social or occupational functioning; and the absence of an alternative explanation, such as depression, for these characteristics.

In the standard two-stage method for dementia diagnosis, cognitive screening instruments exclude most people who do not have dementia. In developing countries, low levels of education, literacy, and numeracy can result in cognitively unimpaired people screening positive for dementia.[3,4] Culturally and educationally sensitive screening instruments exclude items that test arithmetical ability or require reading or writing skills. Instruments can be adapted to different cultural circumstances[4,5] or new instruments can be devised.[6] Interviewing informants about decline in the patient's cognitive and functional abilities has been shown in different cultures to be at least as effective as cognitive testing and is not biased by educational level.[6–11] The community screening instrument for dementia (CSI 'D')[6] combines culturally sensitive cognitive testing of the patient and an informant interview into a predictive algorithm that has been extensively validated[12] in Cree American Indians,[6,13] Nigerians in Ibadan, and African-Americans in Indianapolis.[14] It has achieved 83% specificity at 87% sensitivity for a diagnosis of DSM-III-R (diagnostic and statistical manual of mental disorders) dementia.[6]

One-stage comprehensive diagnostic procedures allow information on other psychiatric diagnoses to be obtained in a similar way to normal clinical practice. Given the high attrition rate between stage-one and stage-two interviews in developing countries, one-stage procedures reduce bias in the assessment of prevalence and causes, and simplify statistical analysis.[2] Two instruments are commonly used: geriatric mental state (GMS/AGECAT),[15,16] which has attained the greatest popularity; and Cambridge examination for mental disorders in the elderly schedule (CAMDEX).[17] GMS/AGECAT has been widely used to diagnose case-level depression in the community,[18] and its diagnostic validity for organic disorder (dementia) is well established in developed countries.[16] However, difficulties have been noted with this part of the algorithm in developing countries. The instrument has few cognitive items and mainly tests domains of orientation and short-term memory. These tests are likely to be poor discriminators of disease in people with low levels of education; the positive predictive value for GMS-defined

organic disease against a clinical gold standard diagnosis was 57% in a population-based study in India.[19] Also, GMS does not include interviews with informants.

Administration of CSI 'D' with GMS should address some of these difficulties. We have done a pilot multicentre study to develop and test an integrated, one-stage, culturally and educationally sensitive dementia diagnostic instrument based on GMS and CSI 'D'.

Methods

Participants

In every centre, we aimed to recruit 30 participants into each of four groups: mild to moderate dementia (DSM-IV dementia and clinical dementia rating [CDR] scale,[20] mild [1] to moderate [2] severity criteria); depression, defined as Montgomery Asberg depression rating scale (MADRS)[21] score of 18 or higher, but with no evidence of dementia; high levels of education, defined by each centre, with no evidence of dementia; and low levels of education, defined by each centre, with no evidence of dementia.

All participants were aged at least 60 years, were living in the community, and had an informant available for interview. To ensure that interviewers were not aware of participants' diagnoses, participants were recruited and diagnosed by local clinicians who were not involved in subsequent assessments.

We sought to identify people with depression and dementia without relying on previous contact of patients with health services. If there was no alternative, centres recruited patients through contact with health services. Independent clinicians established the diagnosis of dementia by completing a clinical checklist, and formally rated dementia severity using the clinical dementia rating scale. They confirmed the diagnosis of depression with a clinical assessment guided by the Montgomery Asberg depression rating (entry criterion: score of 18 or higher). The two groups with normal cognitive function (low and high education) were recruited either from the general community, or from older people's organisations.

Centres were advised to define high and low education to achieve as much distinction as possible between the two groups while ensuring that adequate numbers of eligible people could be identified. We sought to exclude dementia from these two groups, and from the depression group, by key informant reports of normal functioning, rather than by direct clinical assessment. If possible, participants were interviewed in their own homes to maximise masking. Interviewers were given names and addresses, but were not told diagnoses.

Ethics approval was obtained in the UK from South London and Maudsley NHS Trust/Institute of Psychiatry Ethical Committee (Research) and in other centres from university or hospital ethics committees. All participants gave informed consent, or relatives gave consent on behalf of people with dementia that rendered them unable to give consent. Whether consent was written or verbal varied between centres and was dependent on local practice and the literacy of participants.

Procedures

Centres joined the 10/66 Dementia Research Group between September 1998 and mid-2000; most were University-based and led by clinicians, neurologists, psychiatrists, psychologists, or physicians. Staff at all centres were trained in the study protocol, data handling, data entry, and use of GMS and CSI 'D'. Training at the Chinese and Indian centres was done by MP and John Copeland in English. For Latin America, the Brazilian (Portuguese speaking) and Hispanic network coordinators were trained by MP in English. The coordinators trained investigators from the Latin American centres in their own languages. For the GMS, training took 2–3 days: each trainee viewed and co-rated two training tapes, completed and rated a supervised training interview, and co-rated four to six training interviews. Interviewers were clinicians (psychiatrists, neurologists, and primary-care doctors), paraclinicians (psychologists, nurses), or lay people (psychology graduates, social scientists, and social survey interviewers).

All study instruments were translated, back-translated, and assessed for acceptability and conceptual equivalence. Translations were done locally, by investigators fluent in English (the language of the instruments) and in the local language or languages to be used in the study. The local version was reviewed by local informants.

A research assistant, masked to participants' diagnoses, administered four assessments to every participant. First, CSI 'D', which is a 32-item cognitive test administered to the participant (20 min) and a 26-item informant interview about the participant's daily functioning and general health (15 min). Three summary scores can be generated from CSI 'D': cognitive score (COGSCORE), an item-weighted total score from the participant cognitive test; informant score (RELSCORE) an unweighted total score from the informant interview; and discriminant function score (DFSCORE), a weighted score combining COGSCORE and RELSCORE. COGSCORE and DFSCORE have validated cut-off points for probable and possible cases of dementia.

Second, the animal naming verbal fluency task[22] from the Consortium to Establish a Registry of Alzheimer's Disease (CERAD), which can be extracted from the cognitive test component of CSI 'D', but is given little weight in COGSCORE. Participants are encouraged to name as many different animals as they can in 1 min.

Third, an adapted CERAD ten-word, list learning task.[23] Six words were taken from the original CERAD English language list:[24] butter, arm, letter, queen, ticket, and grass. Pole, shore, cabin, and engine were replaced with corner, stone, book, and stick, which were deemed more culturally appropriate. In the learning phase, the list is read out to the participant, who is immediately asked to recall the words they remember. This process is repeated three times, giving a total learning score out of 30. After 5 min, the participant is asked to recall the ten words, giving a delayed recall score out of 10.

Fourth, GMS/AGECAT, which is a 25–40 min clinical interview that generates, from a computerised algorithm (AGECAT), symptom scores in nine diagnostic clusters: organic brain syndrome (dementia); schizophrenia; mania; neurotic and psychotic depression; and obsessional, hypochondriac, phobic, and anxiety neuroses. Scores of 3–5 denote probable cases, 1 and 2 denote subcases, and 0 denotes no or negligible relevant symptoms. These stage-one diagnoses are organised into final stage-two diagnoses on the basis of a hierarchy imposed by a structured algorithm.

Statistical analysis

We estimated the means of the CERAD ten-word immediate and delayed recall scores, and the animal naming score, and of the CSI 'D' COGSCORE, RELSCORE, and DFSCOREs, by centre, region, and group status. We used ANOVA to account for the variance in these measures by group, and having adjusted

	Dementia	Depression	High education controls	Low education controls	Total
India					
Vellore	22	24	11	33	90
Chennai (VHS)	30	30	30	30	120
Chennai (SCARF)	30	25	30	30	115
Goa	30	30	30	30	120
Thrissur	29	21	14	28	92
Bangalore	21	19	32	31	103
Hyderabad	30	30	30	30	120
Subtotal	192	179	177	212	760
China and southeast Asia					
China (Beijing)	30	30	30	30	120
China (Hong Kong SAR)	30	30	30	30	120
Taiwan (Taipei)	31	32	30	34	127
Subtotal	91	92	90	94	367
Africa					
Nigeria (Anambra)	20	16	10	30	76
Latin America					
Argentina	30	33	33	34	130
Brazil (São Paulo)	30	30	30	30	120
Brazil (Botucatu)	30	30	30	30	120
Brazil (São Jose do Rio Preto)	30	30	30	30	120
Chile	27	23	22	28	100
Cuba	40	29	30	31	130
Dominican Republic	30	30	30	30	120
Guatemala	30	30	30	30	120
Mexico (Mexico City)	29	29	29	34	121
Mexico (Guadalajara)	30	30	30	30	120
Panama	30	30	30	30	120
Peru	30	30	29	31	120
Uruguay	30	31	34	26	121
Venezuela	30	30	30	30	120
Subtotal	426	415	417	424	1682
Overall total	729	702	694	760	2885

VHS=voluntary health services. SCARF=Schizophrenia Research Foundation.

Table 1: **Numbers of participants in each group in every centre**

for group, the variance independently accounted for by centre and region. The centre and region effects were estimated in two separate models since centre is a subclassification of region. We assessed the overall ability of the above measures as the area under the receiver-operator characteristics curve (sensitivity plotted against 1 minus specificity) in discriminating dementia from depression, dementia from low education, and dementia from all non-cases. We estimated the sensitivity (%) for dementia of the CSI 'D' COGSCORE and DFSCORE using previously derived item weights and cut-off points,[6,14] and their false positive rates (%) in participants with depression and in the high and low education control groups, by region. The distribution of GMS/AGECAT diagnoses were assessed by group status (dementia, depression, high and low education control groups) and by region.

We identified sources of bias in two ways. First, as the proportion of variance in the predictive measure accounted for by group status that was accounted for by the differences between the high and low education groups. Two models were compared: model 1, dementia versus depression versus high education controls versus low education controls; and model 2, dementia versus depression versus controls (grouping low and high education controls). The relevant statistic was:

$$\frac{\text{(variance accounted for by model 1} - \text{variance accounted for by model 2)}}{\text{variance accounted for by model 1}}$$

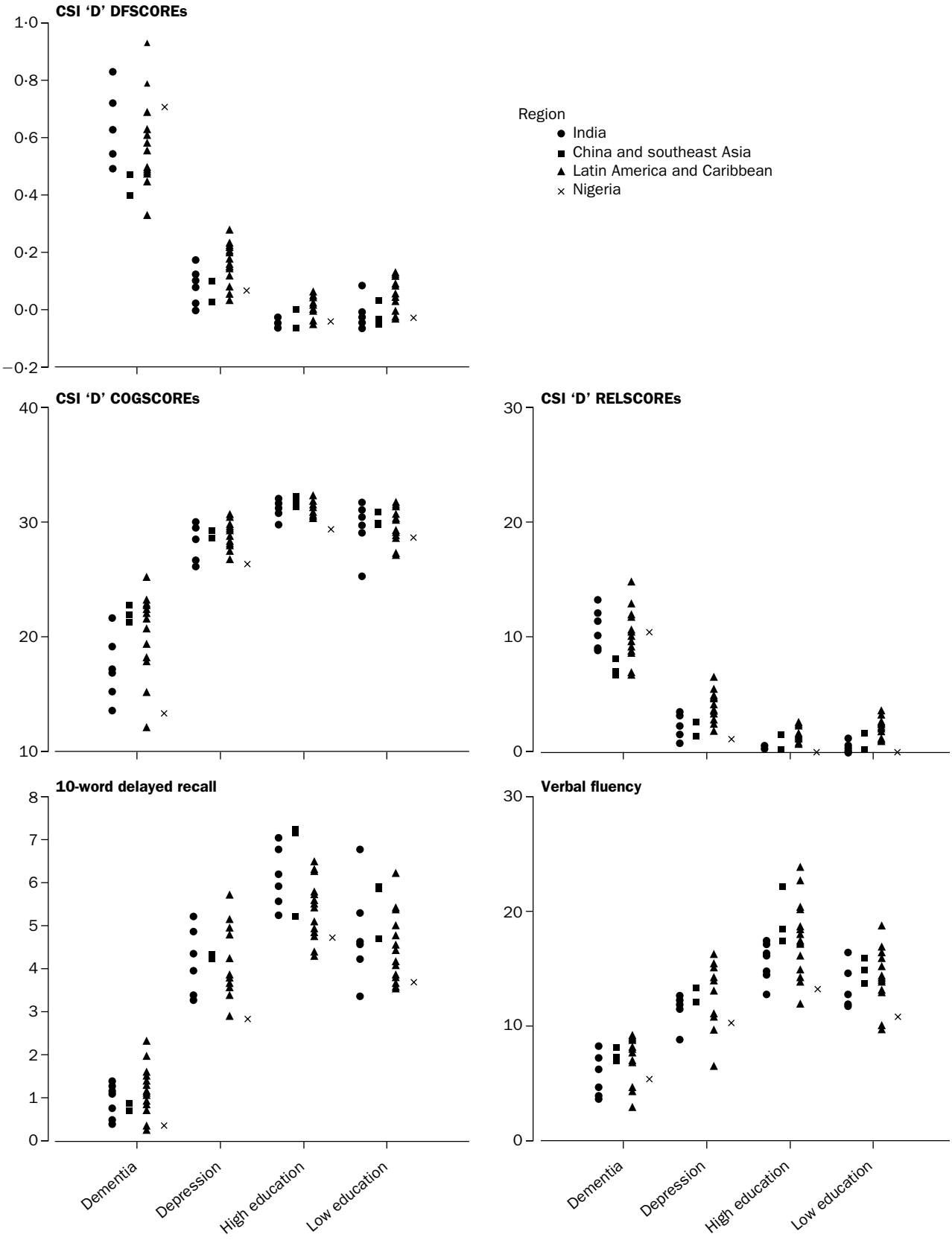
The ideal result for an educationally unbiased measure is 0%. An F statistic for the significance of the effect of education was also calculated.

Second, sources of bias were assessed as independent predictors of the diagnostic test (GMS/AGECAT dementia, or probable dementia by COGSCORE or DFSCORE) after adjustment in a logistic regression model for the independent clinicians' gold standard of true diagnosis. Any independent associations show bias, a systematic tendency to misdiagnose dementia.

We assessed the overall discriminatory ability of the most discriminant items, tests, or both using predicted probabilities derived from another logistic regression model in which the gold standard diagnosis was simultaneously regressed on all the independently predictive tests. The effects of centre and region were not included in this model. To avoid the overprediction implicit in the circularity of this procedure, we divided the dataset in half using random numbers generated by SPSS. The logistic model was calibrated on the first half (development sample), and then applied to the second half (test sample). The performance of the calibration model (applied to the test sample) was compared with that of the individual assessments.

Role of the funding source

The sponsors of the study had no role in the study design, data collection, data analysis, data interpretation, writing of the report, or decision to submit the paper for publication.



Centre means for continuous scales, by region and case group

Results

Table 1 shows the numbers of interviews completed in each centre. 2885 people were interviewed: 760 in India, 367 in China and southeast Asia, 76 in Nigeria, and 1682 in Latin America and the Caribbean. 729 had dementia, 702 depression, 694 were high education controls, and 760 low education controls. We did not gather data on response rates; however as with most such research in developing

Test	Region	Dementia *vs* low education	Dementia *vs* depression	Dementia *vs* all non-cases	% of group effect due to education	p for effect of education
Animal naming	India	89 (86–93)	82 (78–87)	88 (85–92)	6·0	<0·0001
	China and SE Asia	85 (79–91)	77 (70–85)	87 (82–91)	14·5	<0·0001
	Latin America	84 (82–87)	79 (76–82)	85 (83–87)	12·9	<0·0001
	All centres	85 (84–87)	80 (77–82)	85 (84–87)	11·2	<0·0001
10-word list learning Immediate recall	India	87 (84–91)	84 (79–88)	89 (86–92)	11·7	<0·0001
	China and SE Asia	91 (87–95)	83 (77–90)	91 (88–94)	9·0	<0·0001
	Latin America	84 (81–87)	81 (78–84)	86 (84–88)	8·7	<0·0001
	All centres	86 (84–88)	82 (79–84)	87 (86–89)	9·3	<0·0001
10-word list learning delayed recall	India	90 (87–93)	86 (82–90)	91 (88–93)	7·2	<0·0001
	China and SE Asia	96 (93–98)	87 (81–93)	94 (92–97)	2·9	0·002
	Latin America	88 (86–90)	84 (82–87)	88 (87–90)	4·0	<0·0001
	All centres	90 (88–91)	85 (83–87)	90 (89–91)	4·7	<0·0001
Cognitive score	India	93 (90–96)	89 (85–93)	93 (91–95)	1·7	0·0005
	China and SE Asia	95 (92–98)	88 (83–94)	95 (92–97)	2·2	0·006
	Latin America	89 (87–91)	85 (82–87)	90 (88–92)	1·7	<0·0001
	All centres	91 (89–92)	86 (84–88)	91 (90–92)	1·7	<0·0001
Informant score	India	98 (97–100)	93 (90–96)	97 (95–98)	0·0	0·65
	China and SE Asia	97 (95–99)	89 (84–95)	95 (92–97)	0·0	0·99
	Latin America	95 (93–96)	87 (84–89)	93 (91–94)	0·3	0·04
	All centres	96 (95–97)	88 (87–90)	94 (93–95)	0·2	0·15
Discriminant function score	India	99 (97–100)	95 (92–97)	97 (96–99)	0·1	0·09
	China and SE Asia	98 (97–100)	92 (87–96)	97 (95–98)	0·3	0·37
	Latin America	96 (94–97)	90 (87–92)	94 (93–96)	0·6	0·004
	All centres	97 (96–98)	91 (89–93)	96 (95–96)	0·3	0·002

Table 2: **Discriminatory ability (areas under ROC curves, with 95% CIs) of individual tests by region**

countries,[2] very few people refused to participate. Despite the lack of fixed criteria, the distribution of education in the low and high education groups was similar across regions. Thus, in the low education groups, the proportions receiving no, or minimal, education were 91% for India, 89% for China and southeast Asia, and 80% for Latin America and the Caribbean. In the high education groups, the proportions completing secondary education were 81%, 99%, and 80%, respectively.

Group status accounted for most variance in the CSI 'D' scores (COGSCORE 48%, RELSCORE 64%, and DFSCORE 66%). However, after adjustment for the effect of group, CSI 'D' scores varied significantly between centres (COGSCORE 8% [p<0·001], RELSCORE 14% [p<0·001], and DFSCORE 12% [p<0·001]) and between regions (COGSCORE 2% [p<0·001], RELSCORE 6% [p<0·001], and DFSCORE 3% [p<0·001]) (figure). Thus, most centre variance was within rather than between regions. However, in general, more impaired scores were returned from Latin American centres, which was mainly accounted for by RELSCORE (informant interview scores) rather than COGSCORE (cognitive test scores). Informants in Latin American and Caribbean countries, and to a lesser extent Chinese-speaking countries, were more likely to complain of cognitive and functional impairment in relatives with depression than were informants from India.

CSI 'D' showed good discriminatory ability (tables 2 and 3); however, there were difficulties distinguishing dementia from depression, and, to a lesser extent, dementia from the low education control group. CSI 'D' functioned better in India than in Chinese or Latin American regions. In most centres and all regions, CSI 'D' DFSCORE combining CSI 'D' COGSCORE and

	Dementia sensitivity (%)	Depression FPR (%)	High education controls FPR (%)	Low education controls FPR (%)
CSI 'D' COGSCORE				
India	96	2	5	24
China and SE Asia	90	34	0	15
Latin America	90	34	7	27
Nigeria	100	56	10	35
All centres	92	33	6	25
CSI 'D' DFSCORE				
India	96	13	1	2
China and SE Asia	87	11	2	2
Latin America	96	39	5	13
Nigeria	100	19	0	0
All centres	95	29	4	9
GMS/AGECAT				
India	83	25	2	33
China and SE Asia	71	7	1	5
Latin America	65	14	2	8
Nigeria	95	13	0	14
All centres	71	15	2	13
CSI 'D', GMS, and 10-word list algorithm				
India	99	10	3	1
China and SE Asia	95	10	0	2
Latin America	92	20	3	10
Nigeria	100	14	0	7
All centres	94	15	3	6

FPR=false-positive rate.

Table 3: **Discriminatory ability of individual tests and combined algorithm**

	Dementia	Depression	High education controls	Low education controls
Dementia case				
India	118 (83%)	34 (25%)	2 (2%)	48 (33%)
China and SE Asia	65 (71%)	6 (7%)	1 (1%)	5 (5%)
Latin America	274 (65%)	57 (14%)	8 (2%)	34 (8%)
All centres	476 (71%)	99 (15%)	11 (2%)	91 (13%)
Dementia subcase				
India	6 (4%)	0 (0%)	3 (2%)	5 (3%)
China and SE Asia	7 (8%)	2 (2%)	1 (1%)	2 (2%)
Latin America	16 (4%)	3 (1%)	2 (1%)	4 (1%)
All centres	29 (4%)	5 (1%)	7 (1%)	13 (2%)
Depression case				
India	12 (9%)	95 (70%)	3 (2%)	5 (3%)
China and SE Asia	3 (3%)	77 (84%)	3 (3%)	6 (6%)
Latin America	74 (18%)	304 (74%)	85 (20%)	117 (28%)
All centres	89 (13%)	490 (75%)	91 (14%)	128 (19%)
Depression subcase				
India	3 (2%)	0 (0%)	8 (6%)	11 (8%)
China and SE Asia	3 (3%)	2 (2%)	7 (8%)	21 (22%)
Latin America	15 (4%)	20 (5%)	71 (17%)	63 (15%)
All centres	22 (3%)	22 (3%)	86 (13%)	96 (14%)
Other case or subcase				
India	1 (1%)	5 (4%)	21 (16%)	10 (7%)
China and SE Asia	6 (7%)	5 (5%)	15 (17%)	21 (22%)
Latin America	36 (9%)	25 (6%)	116 (28%)	109 (26%)
All centres	43 (6%)	35 (5%)	154 (24%)	149 (22%)
Healthy				
India	2 (1%)	2 (2%)	92 (72%)	67 (46%)
China and SE Asia	7 (8%)	0 (0%)	63 (70%)	39 (42%)
Latin America	4 (1%)	3 (1%)	134 (32%)	97 (23%)
All centres	13 (2%)	5 (1%)	301 (46%)	216 (31%)

*Discrepancies between regional and "All centres" totals are accounted for by the 75 Nigerian participants who were too few in number to be analysed separately, but were included nevertheless in the "All centres" analyses.

Table 4: **Number (%) within each group accorded GMS/AGECAT (stage 2) diagnoses by region***

RELSCORE was better than COGSCORE alone. Sensitivity for dementia was marginally affected, but the false-positive rate in the depression and especially the low education control groups was much reduced.

The ten-word list learning task was a moderate predictor of dementia status (figure and table 2). Performance on delayed recall was generally more predictive than was performance on immediate recall, and was less affected by education. The animal naming task was a poor predictor of dementia status (figure and table 2). There was a substantial educational effect in all regions, and performance in participants with depression was markedly impaired.

	Odds ratio (95% CI)	β coefficient ($\log_e$ odds ratio)
GMS/AGECAT		
Healthy	1·0	0·0
Dementia case	4·8 (1·6–14·1)	1·57
Dementia subcase	4·7 (0·9–24·5)	1·55
Depression case	0·5 (0·2–1·6)	–0·64
Depression subcase	0·5 (0·1–2·0)	–0·67
Other case or subcase	1·4 (0·4–4·6)	0·34
10-word list learning (delayed recall)		
7–10	1·0	0·0
5–6	8·8 (4·5–464)	2·18
4	13·1 (2·9–284)	2·56
1–3	28·5 (1·2–149)	3·35
0	46·7 (0·8–94·4)	3·82
CSI 'D' COGSCORE		
>31·84	1·0	0·0
30·67–31·83	0·8 (0·1–4·9)	–0·23
28·62–30·66	2·4 (0·5–11·3)	0·87
23·70–28·61	4·0 (0·9–18·2)	1·38
0–23·69	16·5 (3·3–81·4)	2·80
CSI 'D' RELSCORE		
0	1·0	0·0
0·5–1·5	6·7 (0·7–66·6)	1·91
2–5	10·1 (1·1–87·7)	2·31
5·5–12	64·8 (7·8–536)	4·17
>12·5	293 (34·5–2486)	5·68

Table 5: **Calibration model (logistic regression) derived from the development dataset including the four assessments independently predictive of dementia status (n=1413)**

	GMS/AGECAT dementia diagnosis OR (95% CI)	CSI 'D' cognitive score OR (95% CI)	CSI 'D' discriminant function score OR (95% CI)	Final algorithm OR (95% CI)
True dementia case	28 (20–33)	60 (43–83)	189 (67–303)	188 (127–279)
Male sex	0·7 (0·6–0·9)	0·8 (0·6–1·0)	1·0 (0·5–1·8)	0·9 (0·6–1·1)
Age	1·04 (1·02–1·05)	1·07 (1·05–1·08)	1·04 (1·03–1·06)	1·07 (1·05–1·09)
Education				
Tertiary	1·0	1·0	1·0	1·0
Secondary	0·6 (0·3–0·9)	0·9 (0·5–1·4)	1·0 (0·6–1·6)	0·8 (0·4–1·5)
Primary	1·2 (0·8–1·9)	2·1 (1·3–3·3)	2·4 (1·5–3·9)	1·7 (0·9–3·3)
Minimal	1·2 (0·8–1·9)	2·9 (1·8–4·4)	1·5 (0·9–2·4)	1·8 (1·0–3·2)
None	3·3 (2·0–5·3)	10·0 (6·3–15·8)	2·6 (1·6–4·3)	3·4 (1·8–6·2)
Region				
India	1·0	1·0	1·0	1·0
China and SE Asia	0·2 (0·1–0·3)	0·5 (0·3–0·7)	0·5 (0·3–0·9)	0·5 (0·3–0·9)
LAC	0·3 (0·2–0·4)	1·1 (0·8–1·4)	3·1 (2·1–4·5)	1·3 (0·9–2·0)

OR=odds ratio. LAC=Latin America and Caribbean.

Table 6: **Independent predictor variables in GMS/AGECAT dementia diagnosis and CSI 'D' probable dementia screening**

Overall, GMS/AGECAT was reasonably accurate in assigning stage-2 (hierarchical) diagnoses (table 4). 505 (75%) people with dementia were identified as dementia cases or subcases. 512 (78%) people with depression were identified as depression cases or subcases. However, 91 (13%) low education controls were misdiagnosed as dementia cases, as were 99 (15%) people with depression. The 91 (14%) high education controls and 128 (19%) low education controls identified as cases of depression might show a genuine high prevalence of depression in these communities, since people with depression were not excluded from these groups. The discriminatory ability of GMS dementia stage-2 hierarchical diagnosis varied between centres. Sensitivity was adequate to good, with 64–97% of true dementia cases correctly identified. Sensitivity tended to be lower in Latin American and Caribbean centres. There was, however, a tendency for misdiagnosis, especially in the low education control group. This tendency was most marked in some Indian centres.

In the development half of the sample, GMS/AGECAT dementia diagnosis, CSI 'D' COGSCORE and RELSCORE, and delayed recall of the ten-word list were each independently associated with true dementia caseness (table 5). For GMS/AGECAT, dementia caseness and subcaseness increased the odds of having dementia four-fold, whereas depression caseness or subcaseness halved the odds. A predicted probability of more than 0·25 produced the best sensitivity and specificity. When the β coefficients from this model were applied to the assessment characteristics of the test half of the sample, and the same probability cut-off point applied, the algorithm was an improvement on each measure used separately (table 3). Consistent results were obtained across the three major regions studied. We stratified analyses by interviewer status (clinician, paraclinician, or lay interviewer), but there was no clear effect on discriminatory ability of the algorithm, or of any of its components.

The proportion of bias by educational level is shown in table 2. The proportions did not vary substantially between regions. For GMS/AGECAT dementia diagnosis, the least educated participants were the most likely to be allocated the diagnosis irrespective of their true case status (table 6). Dementia diagnosis was also marginally associated with older age and male sex. Participants from India were more likely to receive the diagnosis than people from other regions. For CSI 'D', after adjustment for the gold standard, scoring above the cut-off point signifying probable dementia on both COGSCORE and DFSCORE was associated with increasing age, lower levels of education, and region. The effect of education was more prominent for COGSCORE than for either GMS/AGECAT or DFSCORE. For DFSCORE, the effect of region was mainly due to overdiagnosis in Latin America by comparison with India and China. The overall predictive algorithm was as successful as DFSCORE in keeping bias by education to a minimum, and was more culturally sensitive than any of its component measures. However, it remained minimally biased by education and age.

Discussion

Our predictive algorithm, derived from CSI 'D', GMS/AGECAT, and the modified CERAD ten-word list-learning task was a considerable improvement on each measure used separately. We are optimistic that a one-stage culturally and educationally sensitive dementia assessment schedule, based on these three measures, is an achievable goal. Application of the combined algorithm yielded 94% sensitivity, with specificity of 85% in participants with depression, 97% in those with high education, and 94% in those with low education.

The individual measures were robust in assessment of dementia despite the low educational status of many study participants. However, low levels of education present a real difficulty in dementia diagnosis. Injudicious use of inappropriate assessment methods is likely to lead to overdiagnosis of dementia. Confounding of depression with dementia might also be an under-recognised source of diagnostic bias. All three measures were biased by study region, suggesting that identification of a completely culture-free harmonised measure could be challenging. Bias might be explained by several factors: systematic differences in the way in which measures are being administered or coded; systematic differences in the way in which participants are responding to interviews; systematic differences in misclassification of caseness at the point at which participants were recruited into the dementia, depression, and high and low education groups, or in the severity of dementia cases who were recruited.

We attempted to keep methodological artefacts to a minimum by careful attention to translation and rigorous training procedures. However, inevitably with the number of centres included, error will have occurred. Arguably, most of this error will have been random rather than systematic, tending to reduce the discriminatory ability of the measures.

The assessments had complementary discriminating characteristics. GMS was biased by education, with a high level of misdiagnosis in participants with low levels

of education who did not have dementia. This tendency was balanced by CSI 'D', which is less educationally biased than GMS. However, CSI 'D' was biased with respect to depression, in particular in the informant section since relatives of people with depression tended to rate them as cognitively or functionally impaired, a tendency most marked in Latin America. The comprehensive assessment of mental-health status in GMS and capacity for generating hierarchical differential diagnoses seemed to compensate for this effect.

The discriminatory ability of the combined algorithm was best in China and southeast Asia and India and worst, although still perfectly acceptable, in Latin America. The discrimination of cases of dementia from depression evidently poses a continuing challenge, with 15% of depression cases overall being misclassified by the algorithm as probable dementia. Although this figure is substantial, it should be acceptable for population-based research, in which the prevalence of major depression in elderly people (roughly equivalent to the MADRS cutpoint of ⩾18 selected for this study) has been consistently reported to be as low as 1%.[25–27] Because of the typically high rates of dementia incidence in cases clinically diagnosed as depressive pseudodementia,[28] false-positives could indicate an incipient dementia process that was not apparent to the clinician who recruited participants with depression.

Our algorithm is available for use in population-based research, and would seem a valid approach regardless of whether the component interviews are administered by clinical or lay interviewers, after appropriate training. It is especially suited for studies based in low education populations in the developing world, or those designed to make valid comparisons across countries and cultures. The gold standard in our study was the local clinicians' diagnosis of dementia syndrome according to DSM-IV and CDR mild-to-moderate dementia. Further validation in population-based research against similar clinical criteria would therefore be unnecessarily repetitive.

We intend to focus on the prospective validity of the algorithm-based outcome, against future cognitive and functional decline and postmortem neuropathology. The 10/66 Dementia Research Group will apply the algorithm in population-based studies, using harmonised protocols in Latin America, India, and China. To enable allocation of dementia subtype (eg, Alzheimer's disease, vascular dementia, Lewy body dementia) the dementia algorithm assessments (CSI 'D', GMS, and CERAD ten-word list learning) will be complemented by an extended informant interview about the onset and course of dementia (the dementia diagnosis and subtype modification of the history and aetiology schedule [HAS-DDS], 15–20 min)[29] and a brief structured neurological assessment.[30] The overall interview burden for this one-stage dementia diagnostic package is around 40–60 min for the participant, and 15–40 min for the informant. This procedure has been tested in our Havana centre and was acceptable. Further details can be obtained from the 10/66 website at http://www.alz.co.uk/1066.

Contributors

The idea for this study was proposed and the design developed at a 10/66 Group Consensus meeting attended by many of the listed members. The design was finalised by M Prince. Centres were trained by M Prince, J Copeland, M Scazufca, and D Acosta. D Acosta (Latin America), H Chiu (China and SE Asia), M Varghese (India), and M Scazufca (Brazil) coordinated research efforts in their respective regions. Data cleaning, collation, and preliminary analysis was done by S Quraishi and M Dewey. M Prince completed data analyses and prepared the first draft of the paper. All authors made contributions to this and subsequent drafts.

The 10/66 Dementia Research Group

A full list of members is available at http://www.alz.co.uk/1066. The following members participated as investigators in this project, and were jointly responsible for development of the protocol, data gathering, data analysis, and the preparation of this report.

Coordinating centre
M Prince, S Quraishi (Institute of Psychiatry, London); J Copeland (University of Liverpool); M Dewey (University of Nottingham).
10/66 India
Bangalore—M Varghese, S Bharath (NIMHANS). Chennai—L Srinivasan, R Thara (SCARF); R Samuel, E S Krishnamoorthy (Voluntary Health Services). Goa—V Patel, A Dias (Sangath). Hyderabad—K Chandrasekhar, M A Verma (Heritage Hospitals). Thrissur—K S Shaji, K P Lal (Medical College). Vellore—K S Jacob, A P Raj (Christian Medical College).
10/66 China and southeast Asia
H Chiu. China (Beijing)—L Shuran, J Liu (Peking University). China (Hong Kong SAR) L Lam, T Chan (Chinese University of Hong Kong). Taiwan (Taipei)—S-I Liu (Mackay Memorial Hospital); P K Yip (National Taiwan University Hospital).
10/66 Latin America and Caribbean
Argentina (Buenos Aires)—R L Arizaga (GCBA; Hospital Santojanni); R F Allegri (GBCA Y CONICET; Hospital Zubizarreta). Brazil (São Paulo)—M Scazufca, P R Menezes (University of São Paulo Medical School). Brazil (Botucatu)—A Teresa de A R Cerqueira (UNESP; Botucatu Medical School). Brazil (São José do Rio Preto)—M C O S Miyazaki, N A M Domingos (São José do Rio Preto Medical School). Chile (Santiago; Concepción; Valparaíso)—P Fuentes (Hospital Del Salvador, Santiago); P Quiroga (Universidad de Concepción, Concepción). Cuba (Havana)—J de J Libre Rodríguez, H B Vea (Facultad de Medicina " Finlay-Albarran", Universidad Medica de la Habana). Dominican Republic (Santo Domingo)—D Acosta (Universidad Nacional Pedro Henriquez Ureña); G Rodriguez (Asociación Dominicana de Alzheimer). Guatemala (Guatemala City)—C A M Ruiz, M L de Floran. Mexico (Mexico City)—A L Sosa, Y R Agudelo (National Institute of Neurology and Neurosurgery). Mexico (Guadalajara)—G G Ortiz (CIBO/IMSS; Lab Desarrollo/Envejecimiento); E D Arias-Merino (Gerontología, Universidad de Guadalajara). Panama (Panama City)—G R de Alba (Paitilla Medical Center Hospital); G Grimaldo (Santa Fe Hospital). Peru (Lima)—M Guerra (Instituto Nacional de Salud Mental "Honorio Delgado-Hideyo Noguchi", Universidad Peruana Cauetano Heredia); M Víctor González (Instituto Peruano de Seguridad Social; ESSALUD). Uruguay (Montevideo)—R Ventura, N Raciope (University of Uruguay). Venezuela (Caracas)–A Salas (Universidad Central de Venezuela, Faculty of Medicine); C G Yánez (Fundación Alzheimer's Venezuela).
10/66 Africa
Nigeria (Anambra)—R Uwakwe, F Olebuezie (Nnamdi Azikiwe University Teaching Hospital).

Conflict of interest statement

None declared.

Acknowledgments

We thank the research workers and clinicians who volunteered their time. We also thank K Hall and H Hendrie for permission to use the CSI 'D' and for their advice and support. The research was done largely without funding support. Alzheimer's Disease International and Friends of Alzheimer's Disease International funded network meetings and training sessions; in particular, we thank N Graham and E Rimmer for their encouragement and commitment. The following centres received direct support for the pilot research programme: São Paulo and Botucatu (Brazil)—FAPESP grant 1998/12727-0 and CNPq grant 301330/96-4 (São Paulo); Taipei (Taiwan)—National Science Council of Taiwan grant NSC-89-2314-B-195-029; Vellore (India) and Santo Domingo (Dominican Republic)—WHO.

References

1 The 10/66 Dementia Research Group. Dementia in developing countries: a consensus statement from the 10/66 Dementia Research Group. *Int J Geriatr Psychiatry* 2000; **15:** 14–20.
2 The 10/66 Dementia Research Group. Methodological issues for population-based research into dementia in developing countries: a position paper from the 10/66 Dementia Research Group. *Int J Geriatr Psychiatry* 2000; **15:** 21–30.
3 Chandra V, Ganguli M, Ratcliff G, et al. Studies of the epidemiology of dementia: comparisons between developed and developing countries. *Aging Milano* 1994; **6:** 307–21.
4 Ganguli M, Ratcliff G, Chandra V, et al. A Hindi version of the MMSE: the development of a cognitive screening instrument for a largely illiterate rural elderly population in India. *Int J Geriatr Psychiatry* 1995; **10:** 367–77.

5 Anzola-Perez E, Bangdiwala SI, Barrientos de Llano G, De La Vega ME, Dominguez O, Bern-Klug M. Towards community diagnosis of dementia: testing cognitive impairment in older persons in Argentina, Chile and Cuba. *Int J Geriatr Psychiatry* 1997; **11:** 429–38.

6 Hall KS, Hendrie HH, Brittain HM, et al. The development of a dementia screening interview in two distinct languages. *Int J Methods Psychiatric Res* 1993; **3:** 1–28.

7 Ritchie K, Fuhrer R. A comparative study of the performance of screening tests for senile dementia using receiver operating characteristics analysis. *J Clin Epidemiol* 1992; **45:** 627–37.

8 Jorm AF, Scott R, Cullen JS, MacKinnon AJ. Performance of the informant questionnaire on cognitive decline in the elderly (IQCODE) as a screening test for dementia. *Psychol Med* 1991; **21:** 785–90.

9 Law S, Wolfson C. Validation of a French version of an informant-based questionnaire as a screening test for Alzheimer's disease. *Br J Psychiatry* 1995; **167:** 541–44.

10 Morales JM, Gonzalez Montalvo JI, Bermejo F, Del Ser T. The screening of mild dementia with a shortened Spanish version of the "informant questionnaire on cognitive decline in the elderly". *Alzheimer Dis Assoc Disord* 1995; **9:** 105–11.

11 Fuh JL, Teng EL, Lin KN, et al. The informant questionnaire on cognitive decline in the elderly (IQCODE) as a screening tool for dementia for a predominantly illiterate Chinese population. *Neurology* 1995; **45:** 92–96.

12 Hall KS, Gao S, Emsley CL, Ogunniyi AO, Morgan O, Hendrie HC. Community screening interview for dementia (CSI 'D'); performance in five disparate study sites. *Int J Geriatric Psychiatry* 2000; **15:** 521–31.

13 Hendrie HC, Hall KS, Pillay N, et al. Alzheimer's disease is rare in Cree. *Int Psychogeriatr* 1993; **5:** 5–14.

14 Hendrie HC, Osuntokun BO, Hall KS, et al. Prevalence of Alzheimer's disease and dementia in two communities: Nigerian Africans and African Americans. *Am J Psychiatry* 1995; **152:** 1485–92.

15 Copeland JRM, Dewey ME, Griffith-Jones HM. A computerised psychiatric diagnostic system and case nomenclature for elderly subjects: GMS and AGECAT. *Psychol Med* 1986; **16:** 89–99.

16 Copeland JR, Dewey ME, Saunders P. The epidemiology of dementia: GMS-AGECAT studies of prevalence and incidence, including studies in progress. *Eur Arch Psychiatry Clin Neurosci* 1991; **240:** 212–17.

17 Roth M, Tym E, Mountjoy CQ, et al. CAMDEX: a standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *Br J Psychiatry* 1986; **149:** 698–709.

18 Copeland JRM, Hooijer C, Jordan A, et al. Depression in Europe: geographical distribution among older people. *Br J Psychiatry* 1999; **174:** 312–21.

19 Rajkumar S, Kumar S, Thara R. Prevalence of dementia in a rural setting: a report from India. *Int J Geriatric Psychiatry* 1997; **12:** 702–07.

20 Morris JC. The clinical dementia rating (CDR): current version and scoring rules. *Neurology* 1993; **43:** 2412–14.

21 Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; **134:** 382–89.

22 Goodglass H, Kaplan E. Assessment of dysphasia and related disorders. Philadelphia: Lea and Febiger, 1983.

23 Ganguli M, Chandra V, Gilbey J. Cognitive test performance in a community-based non demented elderly sample in rural India: the Indo-US cross national dementia epidemiology study. *Int Psychogeriatrics* 1996; **8:** 507–24.

24 Welsh KA, Butters N, Mohs RC, Beekly D, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): V—a normative study of the neuropsychological battery. *Neurology* 1994; **44:** 609–14.

25 Weissman MM, Leaf PJ, Tischler GL, et al. Affective disorders in five United States communities. *Psychol Med* 1988; **18:** 141–53.

26 Bland RC, Newman SC, Orn H. Prevalence of psychiatric disorders in the elderly in Edmonton. *Acta Psychiatr Scand Suppl* 1988; **338:** 57–63.

27 Henderson AS, Jorm AF, MacKinnon A, et al. The prevalence of depressive disorders and the distribution of depressive symptoms in later life: a survey using Draft ICD-10 and DSM-III-R. *Psychol Med* 1993; **23:** 719–29.

28 Kral VA, Emery OB. Long-term follow-up of depressive pseudodementia of the aged. *Can J Psychiatry* 1989; **34:** 445–46.

29 Copeland JR, Dewey ME, Wood N, Searle R, Davidson IA, McWilliam C. Range of mental illness among the elderly in the community: prevalence in Liverpool using the GMS-AGECAT package. *Br J Psychiatry* 1987; **150:** 815–23.

30 Broe GA, Akhtar AJ, Andrews GR, Caird FI, Gilmore AJ, McLennan WJ. Neurological disorders in the elderly at home. *J Neurol Neurosurg Psychiatry* 1976; **39:** 361–66.