

Effects of education and culture on the validity of the Geriatric Mental State and its AGE-CAT algorithm

MARTIN PRINCE, DAISY ACOSTA, HELEN CHIU, JOHN COPELAND, MICHAEL DEWEY, MARCIA SCAZUFCA and MATHEW VARGHESE
for the 10/66 Dementia Research Group

Background The Geriatric Mental State (GMS) is the most widely used psychiatric research assessment for older persons. Evidence for validity comes from the developed world.

Aims To assess the validity of GMS/AGE-CAT organicity and depression diagnoses in 26 centres in India, China, Latin America and Africa.

Method We studied 2941 persons aged 60 years and over: 742 people with dementia and three groups free of dementia (697 with depression, 719 with high and 783 with low levels of education). Local clinicians diagnosed dementia (DSM-IV) and depression (Montgomery-Åsberg Depression Rating Scale score ≥ 18).

Results For dementia diagnosis GMS/AGE-CAT performed well in many centres but educational bias was evident. Specificity was poor in India and sensitivity sub-optimal in Latin America. A predictive algorithm excluding certain orientation items but including interviewer judgements improved upon the AGE-CAT algorithm. For depression, sensitivity was high. The EURO-D depression scale, derived from GMS items using European data, has a similar factor structure in Latin America, India and, to a lesser extent, China.

Conclusions Valid, comprehensive mental status assessment across cultures seems achievable in principle.

Declaration of interest None.

The Geriatric Mental State (GMS) examination (Copeland *et al*, 1986) is the most widely used comprehensive mental health research assessment for older persons (Copeland *et al*, 2002). It is particularly suited for comparative epidemiological research, given its structured format for identifying, rating and recording symptoms, and the use of the AGE-CAT computerised algorithm (Copeland *et al*, 1986) to generate diagnoses. The purpose of this paper is to assess the extent to which the validity of the GMS, first established in developed countries (Livingston *et al*, 1990; Collinghan *et al*, 1993), extends to poorly educated populations in developing countries. The 10/66 Dementia Research Group recently included the GMS as a key component of an algorithm for the diagnosis of dementia in developing countries, in conjunction with the Community Screening Instrument for Dementia (CSI-D; Hall *et al*, 1993) and the modified CERAD ten-word list learning test (Ganguli *et al*, 1996). Here, using the same data, we assess the performance of the GMS in more detail. Validity is assessed at the level of each of the 26 participating centres. The GMS depression diagnoses are examined in addition to organicity (dementia). Responses to individual items contributing to the depression and organicity diagnostic algorithms are assessed across world regions.

METHOD

Study design

The design of the 10/66 dementia diagnosis pilot study is described in more detail elsewhere (Prince *et al*, 2003). In each centre we aimed to recruit 30 participants into each of four groups: mild to moderate dementia; depression; high level of education; and low level of education. Ethical approval for the studies was obtained in London and in the overseas centres. Recruitment was on the basis of

informed consent, or relatives' agreement where individuals with dementia lacked capacity. All participants were aged 60 years or over. To maintain blindness, independent clinicians established the diagnosis of dementia according to DSM-IV (American Psychiatric Association, 1994) criteria (any dementia subtype) by completing a clinical pro forma, and formally rating dementia severity using the Clinical Dementia Rating (CDR) scale (Morris, 1993). They confirmed the diagnosis of depression with a clinical assessment guided by the Montgomery-Åsberg Depression Rating (MADRS; Montgomery & Åsberg, 1979) with an inclusion criterion of a score of 18 or above. Dementia and depression case groups were recruited either from previous clinical contacts or by local key informant nomination. The two groups with normal cognitive function (low and high levels of education) were recruited from the community. Interviewers were given the participants' name and address but were not told of their diagnosis.

Measures

All study instruments were translated and back-translated by bilingual local investigators and the resulting local language version was reviewed by local key informants to check its face validity. The GMS is a 25–50 min clinical interview generating, from a computerised algorithm (AGE-CAT), nine diagnostic clusters: organicity (dementia and other organic brain syndromes), schizophrenia (and related psychoses), mania, neurotic depression, psychotic depression, hypochondriasis, phobias, obsessional neurosis and anxiety neurosis. A diagnostic confidence level for each syndrome ranges from 0 (no symptoms) to 5 (very severely affected). Levels 3 and greater represent likely cases, a degree of severity warranting professional intervention; levels 1 and 2 are sub-cases. Stage 1 diagnoses are then organised into final stage 2 diagnoses on the basis of precedence determined by a hierarchically structured algorithm. We used the original A3 version of the GMS. A briefer B3 'community' version of the GMS omits those sections that assess syndromes with a low prevalence in the general community: mania, obsessive-compulsive disorder, hypochondriasis and some ratings of hallucinations and delusions. It is possible to generate B3 AGE-CAT diagnoses from A3

data-sets as if the briefer interview had been administered instead.

A subset of 12 GMS items contribute particularly towards the determination of the stage 1 organicity diagnostic confidence level. These comprise tests of cognitive ability (knowledge of date of birth and age; discrepancy between stated date of birth and age; orientation to day, month, year and address; recall of name of interviewer; name of their country's current and previous political leader) and two judgements made by the interviewer (presence of memory deficit, and problems with memory worse than problems with thinking). Twelve symptoms of depression in the GMS (depression, pessimism, wishing death, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment, tearfulness) are used to generate the EURO-D 12-item depression symptom scale (Prince *et al*, 1999). The EURO-D was internally consistent and captured the essence of its parent instrument. Across Europe a two-factor solution seemed appropriate: depression, tearfulness and wishing to die loaded on the first factor (affective suffering); and loss of interest, poor concentration and lack of enjoyment loaded on the second factor (motivation).

Training

All centres were trained in the use of the GMS. M.P. and J.C. trained the Chinese and Indian centres, using English. For Latin America the Brazilian (Portuguese-speaking) and Hispanic 10/66 network coordinators were trained by M.P., using English. They subsequently trained investigators from the 14 Latin American centres using their own languages. Over 2–3 days, each trainee viewed and co-rated two training tapes, completed and rated a supervised training interview and co-rated a further four to six training interviews. This represented a necessary compression of the more conventional 5-day training period for the GMS.

Analyses

Dementia diagnosis

- (a) We estimated the sensitivity (%) for dementia of the GMS-A3/AGECAT stage 2 organicity diagnosis (level 3 or greater), and the false-positive rates (%) for each centre, among those with depression and in the 'high-education' and 'low-education' control groups. We compared these parameters with

those that would have been achieved with the briefer GMS-B3.

- (b) We assessed the validity of the 12 organicity items by estimating the proportions of participants in each region and for each diagnostic group who failed on the ten GMS cognitive test items and who were considered to be impaired on the two objective interviewer assessments.
- (c) We further assessed the 12 organicity items as independent predictors of true dementia status using logistic regression. The collective discriminability of the optimally discriminant items was assessed using predicted probabilities from the logistic regression model. To avoid overprediction the data-set was divided randomly into two halves; the logistic model was developed from the first half (development sample) and applied to the second half (test sample). The sensitivity (%) for dementia of the predictive model (on the test sample) and its false-positive rates (%) for each centre among those with depression and in the high and low education control groups was compared with that of the AGECAT stage 2 organicity diagnosis.

Depression diagnosis

We estimated in each region:

- (a) the sensitivity (%) for depression of the GMS/AGECAT stage 2 depression diagnosis (level 3 or greater) in the depression group, and the proportion of GMS/AGECAT stage 2 depression in each of the three other groups (dementia, high-education and low-education), for which depression status was not a selection criterion;
- (b) the mean EURO-D score and standard deviation for each diagnostic group;
- (c) the internal consistency of the EURO-D scale (excluding those with dementia);
- (d) the factor structure of the EURO-D scale items (using principal components analysis with varimax rotation), comparing the results with those published previously for European centres (Prince *et al*, 1999).

RESULTS

Centres and participants

In all, 2941 persons were interviewed: 746 in India, 336 in China and south-east Asia,

119 in Russia, 74 in Nigeria (Africa) and 1666 in Latin America and the Caribbean. Centres were asked to recruit participants aged 65 years and over. In the event, some centres recruited some participants aged 60–64 years, 207 in all or 7% of the total sample. Of the 2941 participants, 742 were people with dementia, 697 people with depression, 719 high-education controls and 783 low-education controls. In the low-education groups the proportions receiving no or minimal education were 91% for India, 89% for China and 80% for Latin America and the Caribbean. In the high-education groups the proportions completing secondary education were 81%, 99% and 80%, respectively.

Dementia

At regional level, the GMS-A3/AGECAT stage 2 organicity rating has reasonable validity against the gold standard clinical diagnosis of dementia (Table 1). Sensitivity appears to be better in Indian and Chinese centres than in Latin American. The false-positive rate among those with little education is worse in India. However, at centre level the performance is patchy. In Thrissur and Goa in Southern India, although sensitivity is excellent, one-half and two-thirds, respectively, of the least well educated are misdiagnosed. In Latin America in Venezuela, Argentina, Chile and Mexico (Guadalajara) less than half of dementia cases are correctly identified. The GMS-B3 stage 2 organicity rating was identical to the A3 rating in most centres and is therefore not cited here. Where it differed significantly, sensitivity was superior with little or no decline in specificity. Thus, in Guadalajara, sensitivity with version B3 was 50% against 13% for A3, in Chile it was 67% compared with 42% and in Argentina it was 57% compared with 50%.

Item-level analysis showed that all of the cognitive test items and each of the two interviewer objective assessments of the presence of memory deficits discriminated effectively between dementia and the other three groups, in each of the regions (Table 2). Defective orientation to year (more so than to month or day of the week), ignorance of the names of the country's current and previous political leaders were the most educationally biased, particularly in India. Response patterns to the latter items were highly dependent upon local political culture. In Cuba, everyone in the control groups knew that Fidel Castro

Table 1 Discriminability of the Geriatric Mental State (GMS) A3 'organicity' AGECAT stage 2 diagnosis and of the new algorithm derived from GMS organicity items (see also Tables 2 and 3)

		Dementia sensitivity (%)	Depression (FPR%) ¹	High-education controls (FPR%)	Low-education controls (FPR%)
India					
VHS, Chennai	GMS-A3	77	17	0	23
	New algorithm	57	17	0	6
SCARF, Chennai	GMS-A3	80	4	0	17
	New algorithm	100	0	0	0
Goa	GMS-A3	97	47	7	67
	New algorithm	85	9	0	24
Thrissur	GMS-A3	86	19	0	54
	New algorithm	85	0	0	0
Hyderabad	GMS-A3	77	33	0	7
China	New algorithm	88	26	9	0
Beijing	GMS-A3	90	0	0	0
	New algorithm	94	0	0	0
Taipei	GMS-A3	71	9	3	12
	New algorithm	67	0	0	0
Africa					
Anambra	GMS-A3	95	13	0	13
	New algorithm	83	14	0	0
Russia					
Moscow	GMS-A3	61	0	3	3
	New algorithm	84	5	14	6
Latin America					
São Paulo	GMS-A3	67	7	0	20
	New algorithm	73	7	0	7
Botucatu	GMS-A3	77	23	0	13
	New algorithm	77	8	0	0
São Jose	GMS-A3	73	30	7	3
	New algorithm	91	16	0	0
La Habana	GMS-A3	85	3	0	13
	New algorithm	86	6	0	0
Peru	GMS-A3	73	3	7	13
	New algorithm	73	15	10	7
Venezuela	GMS-A3	47	13	0	3
	New algorithm	63	6	0	0
Dominican Republic	GMS-A3	67	11	0	22
	New algorithm	88	14	0	18
Argentina	GMS-A3	50	6	0	0
	New algorithm	92	10	0	0
Chile	GMS-A3	42	18	0	6
	New algorithm	80	0	0	0
Panama	GMS-A3	63	13	3	10
	New algorithm	78	11	0	0
Guatemala	GMS-A3	77	17	0	3
	New algorithm	94	12	7	0
Mexico City	GMS-A3	69	14	0	0
	New algorithm	91	25	0	11
Guadalajara	GMS-A3	13	37	3	3
	New algorithm	100	0	0	0

(continued overleaf)

Table 1 (continued)

		Dementia sensitivity (%)	Depression (FPR%) ¹	High-education controls (FPR%)	Low-education controls (FPR%)
Uruguay	GMS-A3	87	0	6	8
	New algorithm	100	38	0	15
India	GMS-A3	80	30	1	37
	New algorithm	85	12	1	5
LAC ²	GMS-A3	65	14	2	8
	New algorithm	86	13	2	4
China	GMS-A3	80	5	2	6
	New algorithm	79	5	0	2

1. FPR% is the false positive rate (%) or 100-specificity (%).

2. LAC, Latin America and the Caribbean; SCARF, Schizophrenia Research Foundation; VHS, Voluntary Health Services.

was their country's leader, as did 82% of people with dementia. Conversely, in Goa only 3% of people with dementia and 13% of people with low education and no dementia could name Mr Vajpayee. Three cognitive test items – disorientation to address, error in stating age, and confabulation in response to the question 'have you seen me before' – were good discriminators with little educational bias across all three regions. The most effective discriminators were the interviewers' two global assessments. A parsimonious model developed using logistic regression on one random half of the data-set included these most effective and least biased items, and excluded orientation to year and knowledge of past and present political leaders (Table 3). The resulting coefficients were applied to the other 'test' half of the data-set, probabilities of group membership were calculated and a cut-off point of 0.30 (optimal in the development data-set) was applied. The new algorithm was markedly more effective at discriminating between dementia and the other three groups than the AGE-CAT organicity rating, both in every region and in nearly every centre (Table 1). At region level, the false-positive rate in the low-education group in India fell from 37% to 5%, and the sensitivity in Latin America and the Caribbean increased from 65% to 86%.

Depression

The sensitivity of the GMS/AGE-CAT stage 1 diagnosis of depression for the MADRS-defined depression case criterion was close to 90% in each of the three main regions (Table 4). This figure dropped to around

70–80% in AGE-CAT stage 2, mainly because the depressive symptoms had been trumped by the organicity (dementia) ratings in the hierarchical diagnosis. Because dementia was an exclusion criterion for selection into the depression group, this suggested misclassification by the stage 2 AGE-CAT algorithm. The high levels of apparent comorbidity in the dementia case groups are noteworthy, as is the apparent high proportion of those with depression in the high- and low-education control groups in Latin America and the Caribbean compared with Indian and Chinese centres. Case-level depression was neither screened for nor excluded from the high- and low-education or dementia groups, so we could not estimate the specificity of the AGE-CAT depression diagnosis.

The distribution of the EURO-D scale, within diagnostic groups, was similar across the three main regions (Table 4). In each region the mean scores were much higher in the depression group than in the dementia or high- and low-education control groups. Internal consistency (Cronbach's α) was universally satisfactory. For India it was 0.91 (range for centres: 0.87–0.95), for Latin America and the Caribbean it was 0.83 (range for centres: 0.64–0.91) and for China and south-east Asia (both centres) it was 0.88. The other two regions were represented by only one centre each – Anambra in Africa ($\alpha=0.93$) and Moscow in Russia ($\alpha=0.86$). Principal component analysis was attempted for three regions: India; China and south-east Asia; and Latin America and the Caribbean. Two factor solutions were applied in each region following inspection of scree plots. Similar factors were extracted for India and for Latin America and the Caribbean (see

Table 5), conforming to the affective suffering (depression, suicidality, tearfulness) and motivation (enjoyment, interest) factors previously reported for EURO-D (Prince *et al*, 1999). In the Chinese centres all of these items loaded on a single factor, whereas the second factor was characterised by guilt and pessimism.

Dementia

Across the developing-country centres included in this study, the GMS was highly effective at discriminating between dementia cases and high-education controls, therefore the data presented here are entirely consistent with earlier reports of the satisfactory validity of GMS/AGE-CAT when used in well-educated developed-country populations (Livingston *et al*, 1990; Collinghan *et al*, 1993). It was in this context that the GMS was first developed and the AGE-CAT algorithm calibrated. In the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS; 1998) the age-specific prevalence of GMS/AGE-CAT organicity was very similar to that consistently reported from other major European and North American population-based surveys.

In developing countries the GMS is a useful adjunct to dementia diagnosis. Our earlier analyses have demonstrated that it adds to the discriminating power of an algorithm, including informant report of decline in cognitive and functional ability (from the CSI-D) and cognitive testing (from the CSI-D and the CERAD ten-word list learning test) (Prince *et al*, 2003). More detailed findings presented here underline a tendency for the GMS to overdiagnose

Table 2 Discriminability by region of the Geriatric Mental State (GMS) organicity items

Item	Dementia	Depression	High education	Low education
Date of birth incorrect				
India	54	36	7	41
China	22	4	0	2
LAC	47	5	1	2
Age incorrect				
India	45	14	2	15
China	12	0	0	0
LAC	47	5	1	2
Discrepancy between date of birth and age				
India	18	3	0	3
China	37	3	0	2
LAC	49	8	1	4
One or more of above				
India	64	54	8	50
China	61	6	0	3
LAC	61	11	1	6
Day of week incorrect				
India	8	2	0	1
China	41	1	0	0
LAC	30	5	1	1
Month incorrect				
India	35	6	1	4
China	43	2	0	0
LAC	49	6	1	1
Year incorrect				
India	40	11	1	9
China	52	5	1	9
LAC	54	8	1	4
Address incorrect				
India	44	11	1	5
China	28	1	0	1
LAC	49	5	2	6
Claimed to have seen interviewer before				
India	24	2	1	0
China	16	5	0	0
LAC	30	2	1	3
Did not recall interviewer's name				
India	81	31	6	12
China	79	29	8	39
LAC	67	28	7	19
Did not recall country's leader				
India	85	39	12	32
China	44	5	0	0
LAC	54	18	5	12
Did not recall country's past leader				
India	90	50	15	41
China	49	13	0	2
LAC	65	38	18	27
<i>Interviewer's opinion</i>				
Participant has difficulty with memory				
India	52	2	0	0
China	39	2	0	0
LAC	62	6	1	1
<i>Interviewer's opinion</i>				
Problems with memory more prominent than problems with thinking				
India	45	1	1	1
China	26	0	0	0
LAC	65	13	1	7

LAC, Latin America and the Caribbean.

dementia in low-education groups in some but not all centres, and for a relative insensitivity to the presence of dementia in others. Given that the items contributing to the AGE-CAT organicity algorithm can be used to generate an algorithm that is much less educationally biased, one can infer that, in Latin America, AGE-CAT gives more weight to some of those items that we have identified as relatively educationally biased and gives less weight to items that are sensitive to the presence of dementia.

Our data also suggest that the briefer B3 'community' version of the GMS may, paradoxically, be a more valid assessment for dementia than the more comprehensive GMS-A3. In a few centres it would appear that ratings for the sections excluded from B3 (mania, obsessive-compulsive disorder, hypochondriasis and some ratings of hallucinations and delusions) were sub-optimal, giving rise to implausible diagnoses. Thus, in Guadalajara, Mexico, 43% of all dementia true cases were rated by stage 2 as cases of mania. Similar but less extreme problems were noted for some other Latin American centres. Extensive training was provided in all Latin American centres by the regional coordinator but it was not possible logistically after training to supervise directly the conduct of the research in each and every centre. Our collective experience as trainers is that those elements of the GMS-A3 version omitted in the B3 are the most problematic with respect to achieving reliable and accurate ratings, particularly with non-clinical interviewers. Given the low prevalence of these symptoms in community samples it would seem advisable to use the B3 version.

Depression

There is ample evidence from our data for the core validity of the AGE-CAT depression algorithm, at least with respect to its sensitivity to the relatively severe form of depression implied by our independent-clinician inclusion criterion of a MADRS score of 18 or over. It is possible that applying the diagnostic hierarchy in stage 2 may lead to misclassification of depression as dementia. Alternatively, given the typically high rates of dementia incidence in cases clinically diagnosed as depressive pseudodementia, 'false positives' may reflect an incipient dementia process that was not apparent to the independent clinician recruiting the depression cases.

Table 3 Predictive model for the diagnosis of dementia derived from the Geriatric Mental State (GMS) organicity items using logistic regression

	β	s.e.	Wald statistic	d.f.	P	Odds ratio (95% CI)
Age incorrect	0.82	0.35	5.6	1	0.018	2.3 (1.2–4.5)
Day of the week incorrect	1.74	0.48	13.0	1	<0.001	5.7 (2.2–14.6)
Month of the year incorrect	1.63	0.34	22.8	1	<0.001	5.1 (2.6–10.0)
Address incorrect	2.23	0.33	46.6	1	<0.001	9.3 (4.9–17.6)
Claims to have seen interviewer before	1.42	0.44	10.6	1	0.001	4.1 (1.8–9.7)
Does not recall interviewer's name	1.01	0.25	16.5	1	<0.001	2.7 (1.7–4.4)
Interviewer judgement						
Memory impairment	2.26	0.35	40.8	1	<0.001	9.6 (4.8–19.3)
Problems with memory worse than problems with thinking	2.94	0.27	117.8	1	<0.001	18.8 (11.1–32.0)
Constant	-3.82	0.21				

Table 4 Prevalence of Geriatric Mental State (GMS) AGECAAT stage 1 and stage 2 depression diagnoses and EURO-D mean scores (standard deviations) by group and by region

	Dementia	Depression	High education	Low education
India				
EURO-D score	3.6 (3.1)	7.5 (2.5)	0.8 (1.4)	1.2 (1.6)
Stage 1	34%	89%	4%	12%
Stage 2	8%	65%	4%	5%
China and south-east Asia				
EURO-D score	2.6 (2.4)	8.3 (2.1)	1.2 (1.1)	1.8 (1.5)
Stage 1	14%	90%	3%	6%
Stage 2	2%	84%	3%	6%
Latin America and Caribbean				
EURO-D score	4.1 (2.8)	6.9 (2.7)	2.0 (2.0)	2.4 (2.1)
Stage 1	51%	90%	22%	31%
Stage 2	18%	74%	21%	27%
Nigeria (Anambra)				
EURO-D score	0.8 (1.2)	7.2 (2.3)	0.2 (0.6)	0.2 (0.4)
Stage 1	5%	100%	0%	0%
Stage 2	0%	87%	0%	0%
Russia (Moscow)				
EURO-D score	4.4 (2.7)	7.9 (2.3)	1.8 (1.9)	2.5 (2.5)
Stage 1	36%	89%	7%	20%
Stage 2	12%	89%	7%	20%

Misclassification will be more marked in low-education samples, and use of the AGECAAT 'patch' should again remedy this problem. Alternatively, this pitfall may be avoided by using instead the non-hierarchical AGECAAT stage 1 diagnosis; this strategy also permits analysis of comorbidity with dementia, which our data demonstrate to be a phenomenon prevalent in all of the countries and cultures under study.

The EURO-D scale, derived from just 12 GMS items and extensively validated across Europe, would seem to have similar internal validity properties in other cultures. The underlying two-factor solutions for the Indian and Latin American centres were both similar to each other and generally concordant with those derived previously across the 14 EURODEP European centres. The factor solution for

the Chinese centres was somewhat different but difficult to interpret, given the small numbers studied.

DISCUSSION

Implications for future use of GMS/AGECAAT

None of the comprehensive diagnostic assessments in common use in adult populations, whether structured or semi-structured, lay interviewer or clinician administered, has adequately addressed the problems posed by older people with organic conditions. Thus, for research in older populations the GMS remains deservedly popular. However, the GMS on its own was never intended to provide a formal diagnosis of dementia. For such a diagnosis the History and Aetiology Schedule (HAS) informant interview with HAS/AGECAAT or the History and Aetiology Schedule - Dementia Diagnosis and Subtype (HAS-DDS) would have to be used (Copeland *et al*, 2002). Without these, the necessary criteria of cognitive and functional decline cannot be established. Neither is it possible to exclude delirium or stable chronic brain injury as an explanation for cognitive impairment; hence the AGECAAT label of 'organicity' rather than 'dementia'. Empirically, in developed countries the GMS/AGECAAT organicity approximates closely to the clinical construct of dementia. In developing countries and other low-education populations, the focus in the GMS/AGECAAT algorithm upon educationally biased cognitive test items risks overdiagnosis. In the 10/66 Dementia Research Group's previously published diagnostic algorithm (Prince *et al*, 2003) this

Table 5 Principal component analysis of EURO–D scale in each of three world regions

	India (n=414)		Latin America and the Caribbean (n=1242)		China and south-east Asia (n=186)	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
Eigenvalue	6.1	1.0	4.3	1.0	6.0	1.1
Variance (%)	51	9	36	8	50	9
Depression	0.80	0.28	0.76	0.21	0.81	0.28
Suicidality	0.79	0.16	0.28	0.55	0.76	0.32
Guilt	0.04	0.61	–0.01	0.65	0.35	0.68
Interest	0.45	0.75	0.37	0.65	0.80	0.35
Appetite	0.58	0.43	0.56	0.21	0.81	0.05
Fatigue	0.59	0.30	0.49	0.39	0.83	0.15
Concentration	0.65	0.40	0.49	0.39	0.60	0.24
Tearfulness	0.81	0.22	0.79	–0.06	0.38	0.51
Sleep	0.72	0.16	0.48	0.27	0.77	0.11
Pessimism	0.82	0.26	0.48	0.41	0.40	–0.58
Irritability	0.21	0.58	0.22	0.45	0.36	0.55
Enjoyment	0.37	0.74	0.23	0.73	0.79	0.27

Values given in bold type indicate factor loadings ≥ 0.60 .

tendency is corrected through education-fair cognitive assessment and informant history of cognitive and functional decline provided by the CSI–D. The GMS is a key element of the 10/66 algorithm because of its unique ability to discriminate between depression and dementia (Prince *et al*, 2003). If the GMS is to be used alone in developing-country and other low-education populations, then caution is indicated in interpreting the organicity output, which may not map as closely onto clinical dementia as in a developed-country population. Future users of the GMS, particularly in low-education populations, may, where resources permit and when dementia is a principal focus, wish to make use of the 10/66 diagnostic algorithm incorporating the CSI–D and CERAD ten-word list learning test (Prince *et al*, 2003). Others might wish to use the ‘patch’ provided in the form of the logistic regression coefficients included in this paper. Note, though, that we administered the GMS with the two other components of the 10/66 algorithm; thus, the remarkable discriminability of the interviewer judgement of the presence of memory impairment might be explained by global impressions, including information from these assessments. Similar discriminability may not be achieved when the GMS is used on its own. A revised AGE-CAT algorithm will provide a more robust long-term solution and it is towards this goal that we now direct our efforts.

More work is required to clarify the cross-cultural validity of GMS/AGECAT. This certainly should include the predictive validity of the organicity rating for future clinical deterioration. Clinicopathological correlation studies are superficially attractive but problematic. In the UK MRC CFAS study, GMS/AGECAT organicity diagnosis predicted the presence upon autopsy of neuropathological features associated with the most prevalent dementia sub-types – Alzheimer’s disease, vascular dementia, Lewy-body dementia and frontotemporal dementia (Medical Research Council Cognitive Function and Ageing Study, 2001). However, these features were also prevalent among those who did not have an AGE-CAT organicity diagnosis *in vivo*. This may reflect upon the suitability of these pathological indicators as gold standards for clinical dementia diagnosis rather than on the specificity of the GMS/AGECAT algorithm.

APPENDIX

10/66 Dementia Research Group

The 10/66 Dementia Research Group, part of Alzheimer’s Disease International, is a collective of researchers from the developing and developed regions of the world. A full list of members with contact details can be found at <http://www.alz.co.uk/1066>. The following members of the 10/66 Group participated as investigators in this project and can be considered jointly responsible for the develop-

ment of the protocol, the data gathering, data analysis and the preparation of this report.

Coordinating Centre

Professor Martin Prince, 10/66 Coordinator, Institute of Psychiatry, London; Ms Seema Quraishi, 10/66 Administrator, Institute of Psychiatry, London; Professor John Copeland, University of Liverpool; Dr Michael Dewey, Institute of Psychiatry, London.

10/66 India (Regional Coordinator Additional Professor Mathew Varghese)

Bangalore: Professor Mathew Varghese, Dr Srikala Bharath, NIMHANS, Bangalore; Chennai (SCARF): Ms Latha Srinivasan, Dr R. Thara, Schizophrenia Research Foundation; Chennai (VHS): Mr Ravi Samuel, Dr E. S. Krishnamoorthy, Voluntary Health Services; Goa: Dr Vikram Patel, Sangath, Dr Amit Dias, Goa Medical College; Hyderabad: Dr K. Chandrasekhar, Dr M. Ajay Verma, Heritage Hospitals; Thrissur: Assistant Professor K. S. Shaji, Professor K. Praveen Lal, Medical College, Thrissur; Vellore: Professor K. S. Jacob, Dr Arockia Philip Raj, Christian Medical College.

10/66 China and ES Asia (Regional Coordinator Professor Helen Chiu)

China (Beijing): Professor Li Shuran, Dr Jin Liu, Beijing University; China (Hong Kong SAR): Professor Linda Lam, Dr Teresa Chan, Chinese University of Hong Kong; Taiwan (Taipei): Dr Shen-Ing Liu, Mackay Memorial Hospital, Professor P. K. Yip, National Taiwan University Hospital.

10/66 Latin America and Caribbean (Regional Coordinators Dr Daisy Acosta (Dominican Republic) and Dr Marcia Sczufca (Brazil))

Argentina (Buenos Aires): Dr Raúl Luciano Arizaga, Hospital Santojanni (GCBA), Dr Ricardo F. Allegri, Hospital Zubizarreta (GBCA Y CONICET); *Brazil (São Paulo):* Dr Marcia Sczufca, Dr Paulo Rossi Menezes, Universidade de São Paulo; *Brazil (Botucatu):* Dr Ana Teresa de A. R. Cerquerira, Botucatu Medical School, UNESP; *Brazil (São Jose do Rio Preto):* M. Cristina O. S. Miyazaki, Neide A. Micelli Domingos, FAMERP Medical School; *Chile (Santiago) Concepción/Valparaiso):* Dr Patricio Fuentes, G. Hospital Del Salvador, Santiago, Dr Pilar Quoroga, L. Universidad de Concepción, Concepción; *Cuba (Havana):* Dr Juan de J. Llibre Rodriguez, Dr Hector Bayarre Vea, Facultad de Medicina 'Finlay-Albarran', Universidad Medica de la Habana; *Dominican Republic (Santo Domingo):* Dr Daisy Acosta, Universidad Nacional Pedro Henriquez Ureña (UNPHU), Lic. Guillermina Rodriguez, Asociación Dominicana de Alzheimer (ADA); *Guatemala (Guatemala City):* Dr Carlos A. Mayorga Ruiz, Dr Mario Luna de Floran; *Mexico (Mexico City):* Dr Ana Luisa Sosa, Dr Yaneth Rodriguez Agudelo, National Institute of Neurology and Neurosurgery; *Mexico (Guadalajara):* Dr Genaro G. Ortiz, Lab Desarrollo/Envejecimiento, CIBO/IMSS, Dr Elva D. Arias-Merino, Gerontologia, Universidad de Guadalajara; *Panama (Panama City):* Dr Gloriela R. de Alba, Patilla Medical Center Hospital, Dr Gloria Grimaldo, Santa Fe Hospital; *Peru (Lima):* Dr Mariella Guerra, Instituto Nacional de Salud Mental 'Honorio Delgado-Hideyo Noguchi', Universidad Peruana Cautetano Heredia, M. Victor González, Instituto Peruano de Seguridad Social, ESSALUD; *Uruguay (Montevideo):* Dr Roberto Ventura, Dr Nair Raciope, University of Uruguay; *Venezuela (Caracas):* Dr Aquiles Salas, Universidad Central de Venezuela, Faculty of Medicine, Dr Ciro Gaona Yáñez, Fundación Alzheimer's Venezuela.

10/66 Africa

Nigeria (Anambra): Dr Richard Uwakwe, Nnamdi Azikiwe University Teaching Hospital.

10/66 Russia

Moscow (Russia): Professor Svetlana Gavrilova, Dr Grigory Jarikov, Alzheimer's Disease Research Center, Mental Health Research Center of Russian Academy of Medical Sciences.

REFERENCES

- American Psychiatric Association (1994)** *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM-IV). Washington, DC: APA.
- Collinghan, G., MacDonald, A., Herzberg, J., et al (1993)** An evaluation of the multidisciplinary approach to psychiatric diagnosis in elderly people. *BMJ*, **306**, 821–824.
- Copeland, J. R., Dewey, M. E. & Griffith-Jones, H. M. (1986)** A computerised psychiatric diagnostic system and case nomenclature for elderly subjects: GMS and AGE-CAT. *Psychological Medicine*, **16**, 89–99.
- Copeland, J. R., Prince, M., Wilson, K. C., et al (2002)** The Geriatric Mental State Examination in the 21st century. *International Journal of Geriatric Psychiatry*, **17**, 729–732.
- Ganguli, M., Chandra, V. & Gilbey, J. (1996)** Cognitive test performance in a community-based non demented elderly sample in rural India: the Indo-US cross national dementia epidemiology study. *International Psychogeriatrics*, **8**, 507–524.
- Hall, K. S., Hendrie, H. H., Brittain, H. M., et al (1993)** The development of a dementia screening interview in two distinct languages. *International Journal of Methods in Psychiatric Research*, **3**, 1–28.
- Livingston, G., Sax, K., Willison, J., et al (1990)** The Gospel Oak Study stage II: the diagnosis of dementia in the community. *Psychological Medicine*, **20**, 137–146.
- Medical Research Council Cognitive Function and Ageing Study (MRC CFAS) (1998)** Cognitive function and dementia in six areas of England and Wales: the distribution of MMSE and prevalence of GMS organicity level in the MRC CFA Study. *Psychological Medicine*, **28**, 319–335.
- Medical Research Council Cognitive Function and Ageing Study Neuropathology Group (MRC CFAS) (2001)** Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. *Lancet*, **357**, 169–175.
- Montgomery, S. A. & Åsberg, M. (1979)** A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, **134**, 382–389.
- Morris, J. C. (1993)** The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*, **43**, 2412–2414.
- Prince, M. J., Reischies, F., Beekman, A. T. F., et al (1999)** The development of the EURO-D scale – a European Union initiative to compare symptoms of depression in 14 European centres. *British Journal of Psychiatry*, **174**, 330–338.
- Prince, M., Acosta, D., Chiu, H., et al (2003)** Dementia diagnosis in developing countries: a cross-cultural validation study. *Lancet*, **361**, 909–917.

CLINICAL IMPLICATIONS

- The Geriatric Mental State (GMS) and its AGE-CAT computerised algorithm may overdiagnose dementia in developing countries and other low-education populations.
- Testing for orientation to year and knowledge of a country's political leaders are particularly educationally biased.
- The EURO-D scale, derived from GMS depression items, has a common underlying factor structure across several continents.

LIMITATIONS

- Small sample sizes in each centre imply some imprecision in the estimation of sensitivity and false-positive rates.
- Although interviewers were blind to diagnosis, their GMS ratings may have been influenced by knowledge of other cognitive assessments administered in the same sitting.
- We were able to study only the sensitivity and not the specificity of the GMS depression diagnosis.

MARTIN PRINCE, Institute of Psychiatry, King's College, London, UK; DAISY ACOSTA, Universidad Nacional Pedro Henriquez Ureña (UNPHU), Santo Domingo, Dominican Republic; HELEN CHIU, Chinese University of Hong Kong, Hong Kong, SAR; JOHN COPELAND, University of Liverpool, Liverpool, UK; MICHAEL DEWEY, Institute of Psychiatry, King's College, London, UK; MARCIA SCAZUFCA, University of São Paulo, São Paulo, Brazil; MATHEW VARGHESE, National Institute of Mental Health and Neurological Sciences, Bangalore, India

Correspondence: Professor Martin Prince, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK. Tel: +44 20 7848 0136; fax: +44 20 7277 0283; e-mail: m.prince@iop.kcl.ac.uk

(First received 7 May 2003, final revision 27 May 2004, accepted 26 June 2004)