

Relation between education and dementia: the role of test bias revisited

Ben Schmand, Jaap Lindeboom, Chris Hooijer, Cees Jonker

Abstract

Several authors have suggested that dementia screening tests may be biased against low levels of education, whereas others find that a low level of education is a genuine risk factor for dementia. The present paper attempts to reconcile these conflicting views by examining item bias and test bias indices of the mini mental state examination (MMSE). Psychometric calculations and receiver operating characteristics (ROC) analyses of sensitivity and specificity as performed by earlier studies were replicated and extended from the database of the Amsterdam Study of the Elderly. This is a population survey on cognitive decline and dementia (age range 65-84). Subjects with a low level of education (primary school) were compared with better educated subjects (at least some secondary education). Cases were matched by age and sex. The results indicate that the MMSE is not educationally biased as far as item characteristics, reliability, and construct validity are concerned. Yet its predictive validity as a screening test for dementia is educationally biased. This bias will effectively be eliminated with a two point higher cut off score for the subjects whose education extends beyond primary school. Even after such score correction, a low level of education probably remains a genuine risk factor for dementia.

(*J Neurol Neurosurg Psychiatry* 1995;59:170-174)

Keywords: education; dementia; mini mental state examination; test bias

Evidence is accumulating that a low level of education is a risk factor associated with dementia.¹⁻³ This is supposed to be the result of class differences in lifestyle, in brain development, or in mental exercise.¹ Perhaps the onset of symptoms is delayed in the better educated because of a greater "brain reserve".² At the same time, the medical literature on dementia remains doubtful as to whether these findings stem from a bias in dementia screening tests against lower educational levels.¹⁻³⁻⁵ Indeed, many researchers have reported correlations between screening tests and education or socioeconomic state (SES). Age and race or ethnic group correlate with these tests as well. This has generally

been interpreted as a psychometric bias against the less educated, the very elderly, and black or Hispanic subjects. Consequently, corrections of the test scores for these variables have been recommended by some,⁶⁻¹³ but others have objected that such corrections might hamper the study of these variables as risk factors for dementia.¹⁴⁻¹⁵ In this paper, we attempt to reconcile these conflicting views. We focus on the relation between education and the mini mental state examination (MMSE),¹⁶ which is the most widely used dementia screening test.

Several studies have explicitly considered test bias or differential predictive validity of the MMSE. Jorm *et al*¹⁷ found no bias against low education. Because the diagnosis of dementia itself may be biased, they used activities of daily living (ADL) as the predictive criterion. However, ADL is not the best choice as a validity criterion to examine a possible bias of the test as a dementia screening tool. Skills in ADL are influenced by a multitude of factors other than dementia. A better choice would have been the *instrumental* activities of daily living (IADL), which are more closely related to dementia than impaired ADL.¹⁸ Moreover, any group difference in psychometric properties would have implied a test bias, but the opposite is not necessarily true. Thus it remains important to examine the predictive validity for dementia itself. This has been done by several others. From a population survey in Sweden it was concluded that score corrections for education are not necessary.¹⁹⁻²⁰ This conflicts with results of community surveys in China,⁵⁻¹⁰⁻²¹ and with several well conducted American studies that found it necessary to use differential cut off scores.¹¹⁻¹²

To determine more clearly whether or not the MMSE is biased against low levels of education and whether or not score corrections are appropriate, we replicated the study of Jorm *et al*¹⁷ with a much larger sample of community dwelling elderly people and a scale of IADL skills as well as an ADL scale. The replication was extended with analyses of the validity of the MMSE with regard to a dementia criterion—namely, a high score on the organic subscale of the geriatric mental state schedule.²² Finally, we improved on the other studies by matching educational groups for age and sex. This is important as age, sex, and education are interrelated in the elderly population. Without such matching, older women would be overrepresented in the low education group, resulting in an overestimation of the prevalence of dementia in this group.

Amsterdam Study of The Elderly, (AMSTEL project), Department of Psychiatry, Vrije Universiteit, Amsterdam, The Netherlands
B Schmand
J Lindeboom
C Hooijer
C Jonker

Slotervaartziekenhuis, Amsterdam, The Netherlands
B Schmand

Department of Medical Psychology, Free University, Amsterdam, The Netherlands
J Lindeboom

Van Foreest Centrum, Heiloo, The Netherlands
C Hooijer

Correspondence to: Dr B Schmand, AMSTEL project, Prins Hendriklaan 27-29, 1075 AZ Amsterdam, The Netherlands.

Received 14 November 1994 and in revised form 27 January 1995
Accepted 3 April 1995

Materials and methods

SUBJECTS

Subjects were participants in the Amsterdam Study of the Elderly (AMSTEL project). This is a two phase population study on cognitive decline and dementia in community dwelling elderly subjects. Participants ($n = 4051$; 65–84 years of age) were recruited from 30 general practices throughout the city.

The design and sampling methods of the AMSTEL project have been described in detail elsewhere.^{23,24}

MEASUREMENTS

In phase 1, a structured interview was given by trained lay interviewers at the participants' homes. The interview included (in this order) questions on educational history, scales of ADL,²⁵ and IADL,²⁶ the geriatric mental state schedule (GMS),²² and the mini mental state examination.¹⁶ The Dutch adult reading test (DART)²⁷ was given to estimate premorbid verbal intelligence. The DART is the Dutch version of the national adult reading test.²⁸

The ADL and IADL were scored as 0 (complete dependency), 1 (able to do the activity with a little help), or 2 (completely able). The resulting ADL and IADL scales consist of six and eight items respectively, with maximum scores of 12 and 16. The internal consistency of the ADL scale is 0.67; of the IADL scale 0.85 (Cronbach's alpha).

The memory and orientation items of GMS and MMSE are partially overlapping. These items were given once; the remaining MMSE items and the DART were given after completion of the GMS. Thus the GMS measurements were not influenced by performance on the MMSE. Non-response on MMSE items was coded as error, assuming that subjects who refused certain items were unable to do these items.²⁹

The presence of dementia syndromes and depressive symptoms was derived from GMS items. To that end, the GMS was used in conjunction with AGE-CAT, a computer program that mimics the diagnostic decision making process and records the presence of psychiatric syndromes at six levels of diagnostic confidence.³⁰ An organic score of 3 or higher is considered to indicate organic brain disease or dementia.²² Depression was measured at symptom level.

In phase 2, all cases with MMSE scores below 22 and age stratified samples with borderline (22–25) and good MMSE scores (26–30) were invited for follow up. A total of 511 participants were examined again after a median interval of seven (range 1–22) weeks. In this phase, the CAMDEX schedule³¹ was used (including a repeated MMSE). The CAMDEX allows for diagnosing dementia by DSM-III-R criteria.³²

DATA ANALYSES

The calculations concerning the internal validity of the MMSE as performed by Jorm *et al*¹⁷ were replicated. To this end, the sample was divided into a low and a high education group. All subjects with primary education (six years) or less were placed in the low education group ($n = 1692$). All subjects whose education extended beyond primary school were placed in the high education group ($n = 2335$). The level of education was not scorable or unknown in 24 cases; these were dropped from the analyses. Twelve of these subjects had dementia.

To eliminate the influences of age and sex, the education groups were matched for these variables on a case to case basis.³³ Two groups ($n = 2 \times 1535$) were composed. All cases were matched for sex. An exact matching for age was possible in 2×1414 cases. A matching with a deviation of no more than 5 years of age was possible in another 2×121 cases.

The replicated calculations involved (a) slope and intercept of the regression lines predicting ADL on the basis of MMSE in both groups, (b) correlation of MMSE with age in both groups, (c) coefficient of congruence (R_c) of factor loadings of the MMSE items, (d) correlation of item difficulties, (e) internal consistency of the MMSE in both groups. (See Reynolds³⁴ for a description of these methods.) This replication was extended with a regression analysis of MMSE with IADL, the correlations of MMSE with DART-IQ and GMS depression score, and the test-retest reliability of the MMSE. Moreover, the predictive validity of the MMSE for the GMS-AGE-CAT organic criterion was examined by comparing the receiver operating characteristics (ROC)^{35–37} of the MMSE in both education groups. Empirical ROC curves, as well as ROC curves after smoothing by a maximum likelihood estimate³⁵ were considered. Finally, the positive predictive values of the MMSE scores—that is, the percentages of demented cases at a particular score in both groups—were compared.

Results

Table 1 shows demographic characteristics and test scores of the sample before and after matching. Removal of the cases that could not be matched did not alter the characteristics of the sample to a great extent (table 1). The low and high education groups were (of course) significantly different in years of education and DART verbal intelligence. The high education group scored better on IADL and

Table 1 Demographic characteristics and mean (SD) scale scores in the total sample and after matching for age and sex

	Total sample ($n = 4051$)	Sample after matching	
		Low education ($n = 1535$)	High education ($n = 1535$)
Men/women (%)	37.6/62.4	32.5/67.5	32.5/67.5
Age (y)	75.4 (5.7)	76.0 (5.6)	75.8 (5.6)
Education (y)	8.2 (2.5)	5.9 (0.3)	9.8 (2.0)**
Verbal IQ (DART)	97.6 (14.7)	89.0 (11.7)	103.5 (13.7)*
ADL score	11.6 (1.1)	11.6 (1.0)	11.6 (1.0)
IADL score	14.5 (2.9)	14.2 (3.1)	14.7 (2.6)**
MMSE score	26.9 (3.8)	26.0 (4.0)	27.8 (2.8)**
GMS depression	1.8 (1.3)	1.9 (1.4)	1.7 (1.2)**
GMS organicity cases	261 (6.4%)	144 (9.4%)	58 (3.8%)*

* $P < 0.0001$, t test; ** $P < 0.0001$, Mann-Whitney; *** $P < 0.00001$, χ^2 test.

Table 2 Cases with dementia (GMS-AGECAT organic 3 or higher) by level of education before matching (in total sample) and after matching for age and sex

Level of education	Before n (%)	After matching n (%)
Incomplete primary	30 (19.9)	27 (19.6)
Complete primary	136 (8.8)	117 (8.4)
Extended primary	23 (4.8)	19 (5.9)
Lower secondary	26 (3.2)	16 (3.1)
Middle secondary	26 (3.8)	18 (3.9)
Higher levels	8 (2.6)	5 (2.8)

MMSE. The low education group showed more symptoms of depression as measured by the GMS. The prevalence of dementia as defined by the GMS organic scale was very different in the groups (9.4% *v* 3.8%). Table 2 contains the numbers of demented cases at different levels of education. Education seemed to have a strong effect on the prevalence of dementia. GMS-AGECAT and CAMDEX/DSM-III-R diagnoses of dementia in phase 1 and phase 2 respectively agreed in 88.5% of cases. The GMS organicity scale and CAMDEX dementia severity rating correlated (0.65 ($P < 0.0001$; $n = 511$)).

Before matching, education correlated significantly with age (-0.11) and sex (-0.14 ; men coded as 1, women as 2; mean years of education in men: 8.7 SD 2.7; in women: 7.9 SD 2.3; $P < 0.001$ by *t* test). The matching eliminated these correlations (table 3). The correlation of MMSE with education was somewhat increased, and its correlation with ADL and IADL was somewhat lowered by the matching procedure. The pattern of correlations between the remaining variables was not affected. The ADL was not correlated, but IADL was slightly correlated with education.

REGRESSION OF ADL AND IADL ON MMSE

In the low education group, the slope of ADL on MMSE was 0.056 (SEM 0.006) and the intercept was 10.106 (SEM 0.158). In the high education group, these values were 0.113 (SEM 0.009) and 8.497 (SEM 0.249) respectively. Both slope and intercept of the groups were significantly different ($P < 0.0001$).

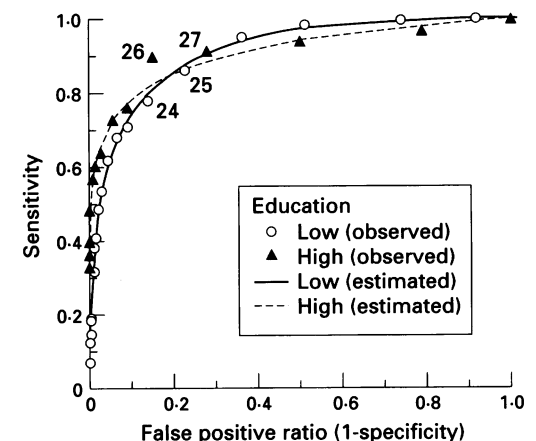
The slope of IADL on MMSE was 0.345 (SEM 0.017) in the low education group; the intercept was 5.236 (SEM 0.458). In the high education group these values were 0.440

(SEM 0.021) and 2.521 (SEM 0.595) respectively. Again, slope and intercept of both groups were significantly different ($P < 0.0001$).

OTHER PSYCHOMETRIC PROPERTIES OF MMSE
Correlations of MMSE scores with age (-0.20), DART-IQ (0.36), and depression (-0.17) were significant ($P < 0.001$), but the differences between pairs of correlations in both education groups were not (Fisher Z). In both groups, a principal components analysis was performed on all 20 items. The coefficient of congruence (R_c) of the item loadings on the first factor was 0.993, which indicates factorial invariance across groups. This first (unrotated) factor explained 23.6% of variance. The correlation between item difficulties in both groups was nearly perfect (0.995). When only the dichotomously scored items were considered, the correlation was 0.98. Internal consistencies (Cronbach's alpha) and test-retest reliabilities of the MMSE were not significantly different in both education groups (overall reliabilities: 0.75 and 0.83 respectively).

PREDICTIVE VALIDITY OF MMSE FOR GMS ORGANIC CRITERION

The figure shows observed operating points and estimated binomial ROC curves for both



Estimated receiver operating characteristic (ROC) curves and empirically observed operating points in low (primary) education and high (secondary) education groups. 24/25 and 26/27 are the optimum cut off points for the low and high education groups respectively.

Table 3 Intercorrelations of demographic characteristics and test scores before matching (lower left) and after matching (upper right)

	Age	Sex	Educ	MMSE	DART	Dep	Org	ADL	IADL
Age									
Sex	0.08								
Education	-0.11	-0.14							
MMSE	-0.22	-0.04	0.16						
DART-IQ	-0.01	-0.12	0.60	0.41					
GMS Depression	0.03	0.15	-0.12	-0.16	-0.10				
GMS Organic	0.17	0.03	-0.10	-0.65	-0.16	0.16			
ADL	-0.15	-0.04	-0.02	0.39	0.09	-0.21	-0.28		
IADL	-0.29	-0.06	0.06	0.52	0.11	-0.26	-0.45	0.60	

Sex coded as 1 = male, 2 = female;

Significance of correlations:

Before matching ($n = 4051$): 0.04 $P < 0.01$; 0.06 $P < 0.001$;

After matching ($n = 3070$): 0.05 $P < 0.01$; 0.06 $P < 0.001$.

Correlations that changed by more than 0.06 as the result of matching are indicated by bold print.

groups. The area under the curve³⁶ was exactly the same in both groups (0.91 (SEM 0.02)). The smoothed curves correspond to the logistic models with the best fit for the observed points (maximum likelihood estimates).³⁵ The optimal cut off points were 24/25 in the low education group and 26/27 in the high education group. Specificity was about equal at these cut off points. The observed sensitivity in the high education group was significantly better than that of the low education group ($Z = 1.974$, $P < 0.05$). The estimated sensitivities at these cut off points were not significantly different, however.

The positive predictive values of MMSE scores in both groups were about the same in the score range from 23 to 26, but were better for the higher educated at lower score levels.

Discussion

The findings of Jorm *et al*¹⁷ were exactly replicated as far as reliability, correlations with age, congruence of factor loadings, and correlation of item difficulties are concerned. The extension with retest reliability and correlations with depression and premorbid verbal intelligence also support the conclusion that the MMSE is education fair.

A different picture emerged when we examined the predictive validity of the MMSE for ADL, IADL, and dementia. Unlike Jorm *et al*¹⁷ we found that the validity of the MMSE for prediction of impairment in ADL and IADL skills is different in low and high education groups. More important were the findings concerning differential predictive validity for dementia. Although the MMSE performed equally well in both groups in terms of area under the curve, it seemed that sensitivity and specificity are comparable only when a two point higher cut off score is used for the better educated than for those with a low level of education (figure).

This converges with the ROC analyses by Uhlmann and Larson,¹¹ who also found a two point score shift between groups with nine years of education or less and 10 years or more; however, it diverges from the Swedish study.^{19,20} This survey is similar to ours in design and sample characteristics. Yet the Swedish investigators concluded that score corrections are not necessary, as the positive predictive value is not influenced by education. The second part of their conclusion is supported by our findings, at least in the critical score range. But a closer look at the Swedish data suggests that they might have concluded too easily that the effect of education is negligible. From the report by Grut *et al*²⁰ it can be read that at the cut off point of 23/24 both sensitivity and specificity are about 90% in the low education group *v* 75% and 95% respectively in the high education group. Raising the cut off for the high education group by two points results in a more comparable sensitivity and specificity of about 85% each.

The score on the GMS organic scale shows

some correlation with education (-0.10 , $P < 0.001$; table 3). This is consistent with findings regarding the DSM-III-R criteria for dementia.³⁸ Thus Jorm *et al*¹⁷ might have been correct in supposing that the dementia criterion itself is biased. It is unlikely that the GMS organic criterion is strongly biased, as its items consist of orientation in time and place, recall of the interviewer's name, and observations on rambling or muddled thinking, perseverations, and indifference or evasiveness to errors. The item most likely to be biased is recall of the name of the prime minister. But even if we were to assume that the GMS organic criterion is moderately biased against low education, the prevalence of dementia seemed to be a factor of 2.5 higher in the low than in the high education group (table 1). Moreover, there is a clear dose-response relation in the present sample (table 2). Such large differences can hardly be explained by a modest bias alone. This finding is even more striking considering that the confounding effect of the most important risk factor for dementia (age) was removed by the matching procedure. A recent study of the incidence of Alzheimer's disease points in the same direction.³ Moreover, the correlation between MMSE and level of education cannot be ascribed to differential prevalences of medical conditions such as poor vision, stroke, Parkinson's disease, diabetes, or depression, as we have shown elsewhere.²⁴ Thus it seems warranted to conclude that, even though dementia screening tests are somewhat biased against low levels of education, there is a real and independent relation between low education and dementia.

The optimal cut off points, which were 24/25 and 26/27 in our study, are not generalisable, but depend on the research purpose or the intended clinical use of the MMSE. Cut off points should be chosen relatively high when good sensitivity is considered more important than a high rate of false positives. On the other hand, a score correction of two points for people with secondary education may be generalised to other populations. Perhaps even larger corrections are necessary for people with university degrees as recommended by Uhlmann and Larson,¹¹ and a bonus for illiterate subjects as suggested by the Chinese studies.^{5,10,21} Both these extremes were hardly represented in our sample.

To summarise: (1) the MMSE is not educationally biased as far as item characteristics, reliability, and construct validity are concerned; (2) its predictive validity as a screening test for dementia is, however, educationally biased; (3) this bias can effectively be eliminated with a two point higher cut off for those whose education extends beyond primary school.

In our opinion, the argument that research into demographic variables may be hampered by score corrections should be turned upside down. The influence of demographic variables as risk factors for dementia can only be studied adequately when their role in normal variation is accounted for.

We are grateful to Dr M Dewey (University of Liverpool) for providing the AGE-CAT computer program, and to Dr L Wouters (University of Amsterdam, Department of Psychiatry) for placing ROC software at our disposal and for his kind assistance in adapting it to our purposes.

We also thank Drs A F Jorm and J Smit for their valuable comments on earlier drafts of this paper.

The AMSTEL project is supported by grants from the Netherlands Health Research Promotion Programme (SGO) and the National Fund of Mental Health (NFGv).

- 1 Mortimer JA, Graves AB. Education and other socioeconomic determinants of dementia and Alzheimer's disease. *Neurology* 1993;43(suppl 4):S39-S44.
- 2 Katzman R. Education and the prevalence of dementia and Alzheimer's disease. *Neurology* 1993;43:13-20.
- 3 Stern Y, Gurland B, Tatemichi TK, et al. Influence of education and occupation on the incidence of Alzheimer's disease. *JAMA* 1994;271:1004-10.
- 4 Jorm AF. *The epidemiology of Alzheimer's disease*. London: Chapman and Hall, 1990.
- 5 Liu HC, Teng EL, Lin KN, et al. Performance on a dementia screening test in relation to demographic variables. *Arch Neurol* 1994;51:910-5.
- 6 Anthony JC, LeResche L, Niaz U, et al. Limits of the 'mini-mental state' as a screening for dementia and delirium among hospital patients. *Psychol Med* 1982;12:397-408.
- 7 Kittner SJ, White LR, Farmer ME, et al. Methodological issues in screening for dementia. *J Chron Dis* 1986;39:163-70.
- 8 Escobar JI, Burnam A, Karno M, et al. Use of the mini-mental state examination (MMSE) in a community population of mixed ethnicity: cultural and linguistic artifacts. *J Nerv Ment Dis* 1986;174:607-14.
- 9 Magaziner J, Bassett SS, Hebel JR. Predicting performance on the mini-mental state examination: use of age- and educational-specific equations. *J Am Geriatr Soc* 1987;35:996-1000.
- 10 Katzman R, Zang MY, Ouang Ya Qu, et al. A Chinese version of the mini-mental state examination: impact of illiteracy in a Shanghai dementia survey. *J Clin Epidemiol* 1988;41:971-8.
- 11 Uhlmann RF, Larson EB. Effect of education on the mini-mental state examination as a screening test for dementia. *J Am Geriatr Soc* 1991;39:876-80.
- 12 Murden RA, McRae TD, Kaner S, Bucknam ME. Mini-mental state exam scores vary with education in blacks and whites. *J Am Geriatr Soc* 1991;39:149-55.
- 13 Ylikoski R, Erkinjuntti T, Sulkava R, et al. Correction for age, education and other demographic variables in the use of the mental state examination in Finland. *Acta Neurol Scand* 1992;85:391-6.
- 14 Berkman LF. The association between educational attainment and mental status examinations: of etiologic significance for senile dementias or not? *J Chronic Dis* 1986;39:171-4.
- 15 Brayne C, Calloway P. The association of education and socioeconomic status with the mini-mental status examination and the clinical diagnosis of dementia in elderly people. *Age Ageing* 1990;19:91-6.
- 16 Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state': a practical method for grading the cognitive status of patients for the clinician. *J Psychiatr Res* 1975;12:189-98.
- 17 Jorm AF, Scott R, Henderson AS, Kay DWK. Educational level differences on the mini-mental state: the role of test bias. *Psychol Med* 1988;18:727-31.
- 18 Barberger-Gateau P, Commenges D, Gagnon M, et al. Instrumental activities of daily living as a screening tool for cognitive impairment and dementia in elderly community dwellers. *J Am Geriatr Soc* 1992;40:1129-34.
- 19 Fratiglioni L, Jorm AF, Grut M, et al. Predicting dementia from the mini-mental state examination in an elderly population: the role of education. *J Clin Epidemiol* 1993;46:281-7.
- 20 Grut M, Fratiglioni L, Viitanen M, Winblad B. Accuracy of the mini-mental status examination as a screening test for dementia in a Swedish elderly population. *Acta Neurol Scand* 1993;87:312-7.
- 21 Hill LR, Klauber MR, Salmon DP, et al. Functional status, education and the diagnosis of dementia in the Shanghai survey. *Neurology* 1993;43:138-45.
- 22 Copeland JRM, Kelleher MJ, Kellett JM, et al. A semi-structured clinical interview for the assessment of diagnosis and mental state in the elderly: The geriatric mental state schedule. I. Development and reliability. *Psychol Med* 1976;6:439-49.
- 23 Hooijer C, Dinkgreve M, Jonker C, et al. Short screening tests for dementia in the elderly population. I. A comparison between AMTS, MMSE, MSQ and SPMSQ. *International Journal of Geriatric Psychiatry* 1992;7:559-71.
- 24 Launer LJ, Dinkgreve MAHM, Jonker C, et al. Are age and education independent correlates of the mini-mental state exam performance of community-dwelling elderly? *J Gerontol* 1993;48:271-7.
- 25 Katz S, Ford AB, Moskowitz RW, et al. Studies of illness in the aged—the index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 1963;185:914-9.
- 26 Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;9:179-86.
- 27 Schmand B, Lindeboom J, van Harskamp F. *De Nederlandse Leestest voor Volwassenen. Handleiding (Dutch adult reading test. manual)*. Lisse: Swets and Zeitlinger, 1992.
- 28 Nelson HE, O'Connor A. Dementia: the estimation of premorbid intelligence levels using the new adult reading test. *Cortex* 1978;14:234-44.
- 29 Fillenbaum GG, George LK, Blazer DG. Scoring non-response on the mini-mental state examination. *Psychol Med* 1988;18:1021-5.
- 30 Dewey ME, Copeland JRM. Computerised psychiatric diagnosis in the elderly: AGE-CAT. *Journal of Micro-computer Applications* 1986;9:135-40.
- 31 Roth M, Huppert FA, Tym E, et al. CAMDEX The Cambridge examination for mental disorders of the elderly. Cambridge: Cambridge University Press, 1988.
- 32 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders, IIIrd revised edition*. Washington DC: American Psychiatric Association, 1987.
- 33 Plugge L, Matcher 1.1. Een programma voor het matchen van proefpersonen. *Psychologie en Computers* 1994;11:28-32.
- 34 Reynolds CR. Methods for detecting construct and predictive bias. In: Berk RA, ed. *Handbook of methods for detecting test bias*. Baltimore: John Hopkins University Press, 1982:199-227.
- 35 Swets JA, Pickett RM. *Evaluation of diagnostic systems: methods from signal detection theory*. New York: Academic Press 1982.
- 36 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.
- 37 Wouters L. *ValidROK, a program for epidemiological analyses*. Amsterdam: Unpublished manuscript, University of Amsterdam, Department of Psychiatry, 1994.
- 38 Henderson AS, Jorm AF, Mackinnon A, et al. A survey of dementia in the Canberra population: experience with ICD-10 and DSM-III-R criteria. *Psychol Med* 1994;24:473-82.