*Psychological Medicine*, 1997, **27**, 919–930. Copyright © 1997 Cambridge University Press

# A prospective study of cognitive function in the elderly

A. E. KORTEN,[1] A. S. HENDERSON, H. CHRISTENSEN, A. F. JORM, B. RODGERS, P. JACOMB AND A. J. MACKINNON

*From the NHMRC Psychiatric Epidemiology Research Centre, The Australian National University, Canberra, ACT and the Mental Health Research Institute of Victoria, Parkville, VIC, Australia*

## ABSTRACT

**Background.** We report on the change in cognitive function in a population sample of elderly people who have been examined on two occasions more than 3 years apart.

**Methods.** A sample of 1135 persons aged 70–102 years was interviewed at base-line then re-interviewed 3·6 years later with the Canberra Interview for the Elderly, which included tests of episodic memory and cognitive speed as well as the Mini-Mental State Examination and the National Adult Reading Test (NART).

**Results.** Mortality and loss to follow-up reduced the sample to 736, of whom 614 completed at least one test of cognitive performance on both occasions. Cognitive performance decreased with age, except on the NART. Decline over the follow-up period increased as a function of age in all cognitive measures, except the NART. Change in cognitive scores was close to normal distribution. Incident dementia was associated strongly with age and current level of cognitive performance, but not with rate of decline. Cognitive decline and the risk of incident dementia did not differ by gender.

**Conclusions.** A score indicating possible impairment in the very elderly carries a worse prognosis than for the younger elderly. Decline is almost universal in at least one cognitive area among those over the age of 85.

## INTRODUCTION

There are compelling reasons to know more about cognitive decline in late life. It is uncertain if the dementias and normal ageing form a continuum or are separate entities (Huppert & Brayne, 1994; Roth, 1994; Cooper, 1995); if cognitive decline can be satisfactorily measured on a single scale that does not have equal intervals; or if the phenomenology of its presentation is consistent. Furthermore, most would agree that variability in outcome occurs among those who have declined, but the pattern is far from established; and little is yet known about factors that influence that outcome. Yet many elderly persons want to know if their decline in memory and thinking is likely to herald a dementia. Next, prediction of response to treatment is at an early stage. Lastly, from a public health perspective, cognitive decline is significant because of its high prevalence, the associated co-morbidity and disablement.

The use of prospective longitudinal designs in community samples of elderly persons has several advantages. Assessment of candidate predictor variables can be made before the onset of a person's cognitive impairment, thereby reducing the likelihood of information bias; and the observations on both the cognitive change and the putative risk factors are independent of whether the person has reached treatment, thereby reducing the likelihood of Berkson's bias (Berkson, 1946). While cross-sectional surveys provide evidence of decreasing cognitive performance with increasing age, decline for an

[1] Address for correspondence: Ms A. E. Korten, NHMRC Psychiatric Epidemiology Research Centre, The Australian National University, Canberra, ACT 0200, Australia.



https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

individual cannot be directly assessed. Measures of performance taken on several occasions allows decline to be measured directly.

Using the Mini-Mental State Examination (MMSE) (Folstein *et al*. 1975) in a sample of 215 community residents, Braekhus *et al*. (1995) found that over 40% of those with an initial score of 24 or 25 had developed a dementia after 3 years. In another community sample, Izaks *et al*. (1995) found that the probability of decline in the MMSE score over 3 years was higher in those with a lower initial score. In a large prospective study, Brayne *et al*. (1995) examined 1014 persons on two occasions some 28 months apart. They found that decline became increasingly common with advancing age and was greater in women. Few of the very elderly had a normal score and maintained it. The change in MMSE score was unimodal and approximately normally distributed, with 28% of the sample showing some improvement by the second assessment, but the majority declined. Brayne *et al*. concluded that 'usual ageing includes some decline for the majority of individuals'.

Here we describe a study of a population sample of the elderly who have been examined on two occasions more than 3 years apart. The aim of this part of the study was to describe the pattern of cognitive change across the range of years in later life and its association with the onset of dementia, using a representative sample of both males and females from the population living in the community and in special hostels and nursing homes for the elderly. No selection was made on the basis of cognitive performance. A particular advantage of this study is that, as well as the MMSE, a broader range of cognitive measures was used which tap different cognitive domains.

## METHOD

In 1990–91, a probability sample of 1377 persons aged 70 years and over was drawn from the Electoral Roll for Canberra and the neighbouring town of Queanbeyan. Approximately equal numbers of men and women were sampled in each age stratum (70–74, 75–79 and $\geqslant 80$ years). A further sample was drawn of 143 persons resident in nursing homes and sheltered accommodation for the elderly, supplemented by a further sample of individuals aged 92 and over. All samples were assessed with the Canberra Interview for the Elderly (CIE) (Social Psychiatry Research Unit, 1992; Mackinnon *et al*. 1993). The CIE includes an interview with the respondent and with a close relative or friend as an informant. Interviews with either or both were successfully completed for 945 (69%) of the respondents in the community, on 100 (70%) of the institutional residents and a further 90 persons aged 92 or over. Further account of the methods and findings can be found in Henderson *et al*. (1993, 1994), Christensen *et al*. (1997). In 1994, an attempt was made to re-interview the survivors of this cohort, using the same instrument, again in the hands of professional interviewers trained in the Centre in its administration.

The CIE is a fully-scripted interview that covers both depressive symptoms and cognitive performance. Computer algorithms generate diagnoses according to DSM-III-R, DSM-IV (American Psychiatric Association, 1987, 1994) and ICD-10 (WHO, 1993). In the present report, only a DSM-III-R diagnosis of dementia has been used, because the ICD-10 diagnosis is set at a higher threshold and yields many fewer cases for analysis (Henderson *et al*. 1994). Measures of cognitive function include: (1) the MMSE, respondents were asked both the items 'WORLD spelt backwards' and 'serial sevens', and the better score was included with the other items to give a maximum score of 30; (2) the Symbol Letter Modalities Test (SLMT) (Christensen *et al*. 1994*a*), which is a measure of cognitive speed similar to the Digit Symbol Substitution subtest of the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981), in which subjects are required to read out letters corresponding to geometric symbols as fast as possible. The raw score is the number correct in 90 seconds; (3) the Episodic Memory Test (EMT) (Jorm 1992) consists of four short memory tasks: three-word recall, name and address recall, figure reproduction and face recognition (these are added together to give a score out of 16); and (4) the National Adult Reading Test (NART) (Nelson, 1982), a reading test of 50 irregularly spelled words. The score is the number of correctly pronounced words, it correlates highly with intelligence tests and is relatively stable across the lifespan. Scores on the SLMT and the NART were converted to

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

Table 1. *Response of Wave* 1 *sample at Wave* 2

| | All Wave 1 interviews* *N* | Died between waves *N* (%)† | Refused/no contact Wave 2 *N* (%)† | All Wave 2 interviews* *N* (%)† | MMSE at both waves *N* (%)‡ |
|---|---|---|---|---|---|
| Total | 1135 | 283 (24·9) | 116 (10·2) | 736 (64·8) | 614 (83·4) |
| Age group | | | | | |
| 70–74 | 398 | 59 (14·8) | 46 (11·6) | 293 (73·6) | 273 (93·2) |
| 75–79 | 317 | 57 (18·0) | 37 (11·7) | 223 (70·4) | 186 (83·4) |
| 80–84 | 209 | 63 (30·1) | 15 (7·2) | 131 (62·7) | 104 (79·4) |
| 85–89 | 93 | 30 (32·3) | 10 (10·6) | 53 (57·0) | 36 (67·9) |
| ⩾ 90 | 118 | 74 (62·7) | 8 (6·8) | 36 (30·5) | 15 (41·7) |
| Sex | | | | | |
| Males | 543 | 152 (28·0) | 49 (9·0) | 342 (63·0) | 297 (86·8) |
| Females | 592 | 131 (22·1) | 67 (11·3) | 394 (66·6) | 317 (80·5) |
| Residence at Wave 1 | | | | | |
| Community | 968 | 187 (19·3) | 97 (10·0) | 684 (70·7) | 594 (86·8) |
| Institution | 167 | 96 (61·6) | 19 (12·1) | 52 (33·1) | 20 (38·5) |
| DSM-III-R dementia Wave 1 | | | | | |
| Yes | 78 | 29 (37·2) | 11 (14·1) | 38 (48·7) | 20 (52·6) |
| No | 842 | 170 (20·2) | 65 (7·7) | 607 (72·1) | 544 (89·6) |
| No diagnosis possible | 215 | 84 (39·1) | 40 (18·6) | 91 (42·3) | 50 (54·9) |
| MMSE Wave 1 | | | | | |
| Mean (S.D.) | 26·7 (4·0) | 24·8 (5·6) | 25·8 (4·5) | 27·4 (2·9) | 27·8 (1·9) |
| Valid *N* | 993 | 217 | 100 | 676 | 614 |
| EMT Wave 1 | | | | | |
| Mean (S.D.) | 10·5 (3·4) | 9·0 (4·0) | 9·9 (3·1) | 11·0 (3·1) | 11·5 (2·5) |
| Valid *N* | 949 | 195 | 95 | 659 | 572 |
| SLMT Wave 1 | | | | | |
| Mean (S.D.) | 94·3 (17·9) | 85·8 (19·0) | 92·6 (19·4) | 97·0 (16·5) | 99·6 (14·9) |
| Valid *N* | 939 | 193 | 92 | 654 | 573 |
| NART Wave 1 | | | | | |
| Mean (S.D.) | 111·4 (10·0) | 109·3 (10·3) | 107·1 (10·5) | 112·7 (9·5) | 114·0 (8·5) |
| Valid *N* | 913 | 191 | 90 | 632 | 534 |

* Includes interviews with respondent only, respondent plus informant or informant only.
† Percentages of the total number of Wave 1 interviews.
‡ Percentages of the total number of Wave 2 interviews.

standardized scores (mean = 100, S.D. = 15) based on norms either collected or published from an earlier sample. Scores on the SLMT, EMT and the NART do not suffer from the scaling artefact of the MMSE whose ceiling effect can mean that change in performance of cognitively able persons may not be detected.

For each cognitive scale, a refusal to attempt any item was coded as an error. However, if a subject could not complete an item because of motor or sensory impairment, the item was set at missing. Final scores with missing items were either pro-rated or, where more than a specified number of the items were missing, set at missing. For the MMSE, for example, if items to the value of 8 points or more out of 30 were missing, the total MMSE score was set to missing, while if fewer than 8 points were missing the final score was pro-rated by dividing the observed score by the maximum score on the non-missing items and multiplying this by 30. In this way, 67 subjects at Wave 1 and 50 at Wave 2 had MMSE scores pro-rated to account for missing items.

## Statistical methods

Hierarchical multiple linear regression was used to model change in a cognitive score (MMSE, EMT, SLMT or NART) as a function of sex, age, score at Wave 1 and its interaction with age. To accommodate possible non-linear effects with age, and to facilitate the interpretation of interaction terms, five-year age groups were formed (70–74, 75–59, 80–84, 85–89, ⩾ 90) based on the age at Wave 1, with corresponding dummy variables. Interaction terms were formed as the product of each dummy variable with the Wave 1 measure of cognitive performance. Significant interaction terms indicate a relation-

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

ship with age which is different for subjects with different initial levels of cognitive performance.

The analysis of change was based on conditional regressions. That is, the Wave 2 score was the dependent variable and the Wave 1 score was entered as the first independent variable, followed by age, sex and interaction terms. As the focus of interest is the change score itself, the estimates from the regressions have been used to calculate predicted mean change, adjusting for the covariates. The change score was not used directly in the regressions on Wave 1 score, age and sex: the results may be misleading or difficult to interpret because of the negative correlation between the dependent and independent variables, or regression towards the mean (Plewis, 1985). Logistic regression was used to model the incidence of DSM-III-R dementia between Waves 1 and 2 as a function of age, sex and cognitive scores. Chi-square tests were used to compare groups of respondents by outcome at Wave 2. Unless otherwise stated, results reported as statistically significant were significant at the 0·01 level, while those reported as not significant had a *P*-value greater than 0·05.

## RESULTS

Table 1 shows the response of the initial sample at follow-up. Of the 1135 persons aged 70–102 who were examined in 1990–91, either full or partial interviews were obtained in 1994 for 736 (64·8%). Their mean age was 80·9 (S.D. 3·6), with a range from 73 to 102 years. Of the remainder, 283 (24·9%) were known to have died in the interim, no contact could be made with 26 (2·3%) and 90 (7·9%) refused to participate or their carers declined. Rates for refusal or failure to contact were independent of age-group, sex or residence at Wave 1 (community or institution). But among those not contacted and the refusals at Wave 2, there had been a lower MMSE score and a higher prevalence of DSM-III-R dementia at Wave 1. Mortality increased significantly with age, was significantly higher among males, among those living in nursing homes or hostels for the elderly, and in those with a dementia at Wave 1. Those who had died also had lower scores on all cognitive scales at Wave 1, compared with survivors. Complete data to make a diagnosis of DSM-III-R dementia at Wave 2 were available from 597 respondents. MMSE scores at both waves were obtained on 614 respondents.

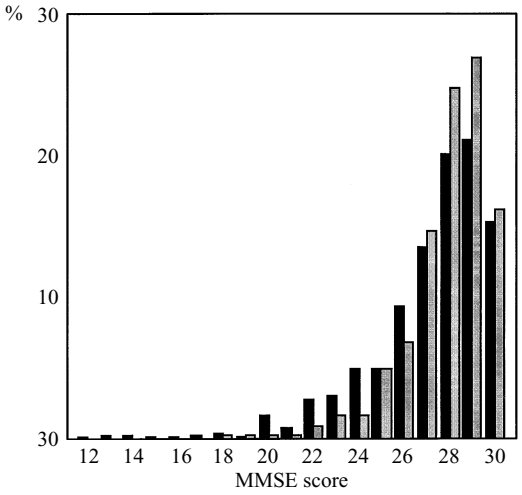



FIG. 1. Distribution of the MMSE scores at Wave 1 (▨) and Wave 2 (■).

### Change in MMSE scores

Among those available for assessment at both waves, the distribution of MMSE scores on the two occasions was highly skewed and showed a distinct downward shift over the 3·6 years (Fig. 1). At Wave 2, 1·8% of the respondents had MMSE scores below 18, a cut-point often taken to denote a high probability of dementia. All of these subjects had scored above 18 at Wave 1. None of the subjects who had scored below 18 at Wave 1 (3·8%) was available for follow-up assessment. It should be noted that these data refer to the distribution of scores across the sample and do not reflect change within individuals. For the latter, the distribution of change in the scores was symmetrical and close to normal (Fig. 2), with mean −0·8 and S.D. 2·4. A change of less than 1 point on the MMSE was recorded by 24% of subjects, a decline of 1 or more points by 48% and an improvement of 1 or more points by 27%. Because of the prorating, changes were not necessarily whole units. Among those scoring 30 at Wave 1, 27% also scored 30 at Wave 2 while a further 23% declined by only 1 unit to score 29.

As in Wave 1 (Henderson *et al.* 1994), the MMSE scores in Wave 2 were lower in the older age groups (Pearson correlation = −0·35), the

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

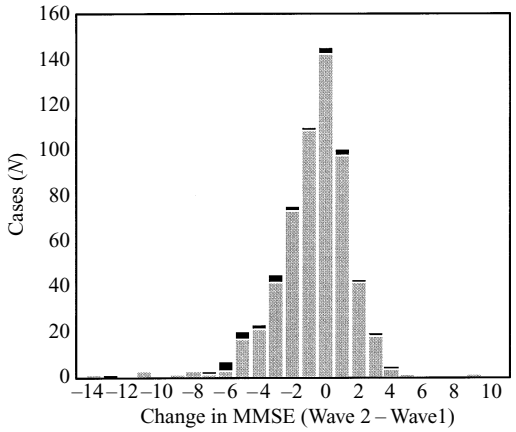



FIG. 2. Distribution of change in MMSE scores (Wave 2 – Wave 1), showing incident DSM dementia cases (■, Yes, Wave 2; ▨, No, Wave 2).

Table 2. *Wave 1 scores and observed mean change (Wave 2 – Wave 1)*

| Age group | *N* | Wave 1 score (S.E.) | Mean change (S.E.) |
|---|---|---|---|
| MMSE | | | |
| 70–74 | 273 | 28·2 (0·1) | −0·3 (0·1) |
| 75–79 | 186 | 27·8 (0·1) | −0·9 (0·2) |
| 80–84 | 104 | 27·3 (0·2) | −1·3 (0·3) |
| 85–89 | 36 | 27·4 (0·3) | −2·3 (0·6) |
| ⩾ 90 | 15 | 26·6 (0·7) | −2·5 (0·7) |
| Total | 614 | 27·8 (0·1) | −0·8 (0·1) |
| SLMT | | | |
| 70–74 | 267 | 103·5 (0·8) | −1·3 (0·7) |
| 75–79 | 169 | 100·5 (1·1) | −4·6 (0·9) |
| 80–84 | 93 | 92·1 (1·4) | −6·3 (1·4) |
| 85–89 | 33 | 91·2 (2·4) | −5·8 (2·5) |
| ⩾ 90 | 11 | 81·7 (4·7) | −0·6 (6·1) |
| Total | 573 | 99·6 (0·6) | −3·3 (0·5) |
| EMT | | | |
| 70–74 | 266 | 12·1 (0·1) | 0·8 (0·2) |
| 75–79 | 169 | 11·6 (0·2) | 0·7 (0·2) |
| 80–84 | 94 | 10·3 (0·3) | 0·2 (0·4) |
| 85–89 | 31 | 11·1 (0·5) | −1·2 (0·6) |
| ⩾ 90 | 12 | 9·0 (1·2) | −1·4 (1·2) |
| Total | 572 | 11·5 (0·1) | 0·5 (0·1) |
| NART | | | |
| 70–74 | 250 | 114·5 (0·5) | 0·4 (0·3) |
| 75–79 | 155 | 114·0 (0·7) | 0·7 (0·4) |
| 80–84 | 84 | 114·0 (0·9) | −0·6 (0·6) |
| 85–89 | 32 | 113·3 (1·3) | −0·1 (0·6) |
| ⩾ 90 | 8 | 118·3 (1·3) | −1·3 (1·1) |
| Total | 529 | 114·1 (0·4) | 0·3 (0·2) |

mean score for those aged 70–74 being 27·9 (S.D. 2·1) and for those aged 90 and above, 24·1 (S.D. 4·5). Decline between waves became significantly greater with age (Pearson correlation between age and change score = −0·24). Table 2 shows the Wave 1 mean scores and the mean decline in MMSE, together with S.E.s, for successive 5-year age groups. Because of increasingly small sample sizes in each age group, S.E.s increase rapidly.

When change in MMSE scores was examined with respect to age, sex and Wave 1 score using a conditional regression approach, 32% of the variance in Wave 2 MMSE scores was explained by the Wave 1 score ($R^2 = 0{\cdot}32$), while age explained an additional 5% of variance ($R^2$-change = 0·05). Neither sex nor sex-by-age interaction was significant and were excluded from subsequent models. The age-by-initial-MMSE interaction terms were significant and explained a further 2% of the variance (Total $R^2 = 0{\cdot}40$). Because an impaired score in a younger respondent represents an extreme value, as does a high score in the very elderly, part of this effect could be due to a regression towards the mean. The significance of the interaction term may also be a result of the ceiling effect, if true decline is being masked more often in the younger age groups than in the very elderly.

The regression equation can be used to estimate the Wave 2 MMSE and hence the mean change in MMSE between waves for given levels of age and Wave 1 score. The regression coefficients indicate that those whose initial levels of MMSE were low but who were in the younger age groups were predicted to show mean improvement, while older subjects with low initial MMSE scores were likely to show further mean decline. Those whose MMSE scores are initially high can only decline, due to the ceiling effect of the MMSE. Among these subjects, the estimated mean decline increases with age. This is illustrated in Fig. 3, which shows the predicted mean change in MMSE and associated 95% confidence intervals by age-group, taking as examples three different initial values of the MMSE (18, 25 and 30). For an individual with a Wave 1 score of 18, the predicted mean decline in MMSE increases rapidly with age. There is a mean improvement in scores for those below age 80 and a large decline in scores for the older age groups. For individuals with an initial MMSE score of 25, decline also increases with age, but less steeply than if the initial MMSE were 18. For an individual with an initial MMSE of 30, there is an estimated mean decline for all ages, and this increases slightly with age.

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

## Change in SLMT scores

The SLMT does not suffer from ceiling effects, unlike the MMSE. The distribution of raw scores was symmetrical, as well as that of change between waves. Between waves, there was a significant mean decline in scores of 3·3 points (S.D. 12·3). Wave 2 scores declined significantly with age (Pearson correlation between Wave 2 score and age = −0·36) and the change between waves increased significantly with age (Pearson correlation between change-score and age = −0·10, $P = 0·015$). Mean changes in the SLMT scores by age group are shown in Table 2. In conditional multiple regressions, the score at Wave 2 was related to the score at Wave 1 ($R^2 = 0·50$) and age ($R^2$-change = 0·03) but was not related to the interaction of the Wave 1 score with age, sex, or sex-by-age interactions. Estimated mean changes in the SLMT scores by age-group, adjusted for initial levels of SLMT, are shown in Table 3. The adjustment is to the average of the Wave 1 SLMT scores across the age groups, i.e. 93·8. The youngest age group, those aged 70–74 at Wave 1, are estimated to show mean improvement in their SLMT scores, for an initial score of 93·8, while those in the older age groups are estimated to show mean decline. It should be noted that the observed Wave 1 mean score for those aged 70–74 was 103·5 and the corresponding predicted mean change is −1·3 (S.E. = 0·7), while those aged 85–90 with an initial score of 103·5 would be predicted to decline 8·9 units (S.E. = 2·1).

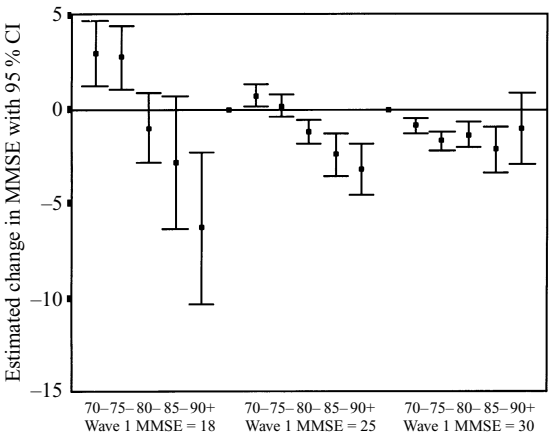



FIG. 3. Predicted mean change in MMSE scores, with 95% confidence intervals, by age group, for different Wave 1 MMSE scores.

Table 3. *Estimated mean change* (*S.E.*) *in SLMT and EMT, adjusted by initial score*

| Age | SLMT* | EMT† |
|---|---|---|
| 70–74 | 1·2 (0·8) | 1·2 (0·2) |
| 75–79 | −2·9 (0·9) | 0·8 (0·2) |
| 80–84 | −6·7 (1·2) | −0·1 (0·3) |
| 85–89 | −6·5 (2·0) | −1·1 (0·5) |
| ⩾ 90 | −3·9 (3·5) | −1·8 (0·8) |

* Initial SLMT score taken as 93·8.
† Initial EMT score taken as 10·8.

## Change in EMT scores

The distributions of the EMT and of the change in EMT scores between waves were symmetrical and, unlike the MMSE, the EMT suffers only minor ceiling and floor effects. Unlike other measures of cognition, however, the EMT showed significant mean improvement between Waves (mean change = 0·51, S.D. = 2·8). This was limited to the younger age groups, as shown in Table 2. Wave 2 scores declined significantly with age (Pearson correlation between Wave 2 score and age = −0·35) and the change between waves increased significantly with age (Pearson correlation between change-score and age = −0·19). In the conditional regression with the score at Wave 2 as dependent variable, the score at Wave 1 (plus its quadratic term) ($R^2 = 0·34$) and age ($R^2$-change = 0·05) were significant, whereas sex and the interaction of age-by-initial-score were not. Estimated mean changes, adjusting for initial levels of the EMT (10·8), are shown in Table 3.

## NART scores

For the NART, distributions of the scores and of the change scores were symmetrical. Overall there was a non-significant improvement in scores of 0·27 (S.D. = 4·6). Scores did not vary significantly with age, sex or Wave 1 score, although those who failed to complete the NART during their Wave 2 interviews had had significantly lower Wave 1 scores (105·6 compared with 114·0, $P = 0·000$). Wave 1 scores and observed mean change according to age-group are shown in Table 2. Because of its stability as a measure of intelligence (Lezak, 1995), the NART is also of interest as a possible predictor of cognitive decline. The correlation of the Wave 1 measure with change scores on the MMSE,

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

Table 4. *DSM-III-R dementia diagnoses at Wave 1 and Wave 2*

| DSM-III-R dementia, Wave 1 | DSM-III-R dementia at Wave 2 | | | | | |
|---|---|---|---|---|---|---|
| | No | Yes | Insufficient data | Died | No contact/ refusal | Total |
| No | 541 | 32 | 34 | 170 | 65 | 842 |
| Yes | 9 | 15 | 14 | 29 | 11 | 78 |
| Insufficient data for diagnosis | 44 | 5 | 42 | 84 | 40 | 215 |
| Total | 594 | 52 | 90 | 283 | 116 | 1135 |

Table 5. *Mean change in cognitive scores by Wave 2 dementia status among dementia-free subjects at Wave 1*

| | DSM-III-R dementia at Wave 2 | |
|---|---|---|
| | No ($N = 541$*) | Yes ($N = 27$†*) |
| MMSE | | |
| Mean change (S.D.) | −0·5 (2·0) | −3·3 (4·0) |
| EMT | | |
| Mean change (S.D.) | 0·6 (2·6) | −1·5 (3·3) |
| SLMT | | |
| Mean change (S.D.) | −2·5 (11·4) | −12·6 (11·3) |
| NART | | |
| Mean change (S.D.) | 0·3 (4·9) | −2·6 (6·1) |

* Maximum *N*. Some missing values for some tests.
† Five cases did not complete any tests at Wave 2.

SLMT and the EMT, adjusted for age and sex, was therefore calculated. None was significant among those completing the tests at both waves.

**Change in DSM-III-R diagnosis of dementia**

Table 4 shows the outcome at Wave 2 with respect to a computer-generated diagnosis of dementia by DSM-III-R criteria. At Wave 1, 78 subjects fulfilled these criteria, of whom 29 died before Wave 2, 11 refused or could not be contacted, and 14 did not provide sufficient information for a diagnosis to be made. Of the remaining 24 cases, 15 retained the diagnosis of dementia. Nine no longer fulfilled the criteria, four because their informants no longer indicated that impairments intefered with usual social activities or relationships with others (Criterion C) and five cases who fulfilled the criteria for intellectual impairment (Criterion B) but just failed to fulfil the criteria for memory impairment (Criterion A). Among all those who were assessed at Wave 2, 38 respondents who were considered by their informants to fulfil criterion C failed on the criteria for cognitive impairment (A and B) while 78 respondents fulfilled both the criteria for memory and intellectual impairment but their informants did not endorse criterion C. This could reflect both the changing circumstances of respondents who were having to face fewer daily demands or it could reflect changing attitudes and expectations of the informants themselves. The data in this survey are insufficient to test these hypotheses.

*Incidence*

Among the non-demented at Wave 1 who provided sufficient information for a Wave 2 diagnosis, there were 32 incident cases. Incidence increased with age ($r = 0{\cdot}21$). Wave 2 scores were significantly lower and decline in all cognitive scores, including the NART, was significantly greater among the incident dementia cases than among the non-cases. Although the MMSE declined 3·3 points, compared with a mean decline of 0·5 points among the non-cases, there is not a strong relationship of change in MMSE score to incident dementia (see Fig. 2), and scores either remained the same or improved for eight of the 26 incident dementia cases with complete MMSE data. Eight out of 27 cases who attempted one or more tests showed improvement in at least one, while at the other extreme, five incident cases, aged between 77 and 86, were so impaired that they were unable to complete any test. Incident dementia cases also had significantly poorer performance on all Wave 1 measures of cognition (see Table 5).

To predict Wave 2 dementia among non-dementia cases at Wave 1, Wave 1 scores on the MMSE, EMT, SLMT and NART were each entered into separate logistic regressions with age and sex. Since sex was non-significant it was excluded from the models. The MMSE, SLMT and EMT, but not the NART, were each significant after adjusting for age, a higher score indicating greater protection against dementia (see Table 6). After adjusting for age, the NART at Wave 1 did not predict subsequent dementia. If age, sex and Wave 1 scores on all of the measures were included in a logistic regression with stepwise selection of independent variables, only age and EMT were selected. To assess whether the amount of decline is a significant predictor of incident dementia, as well as the level reached, the regressions were repeated for

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

Table 6. *Coefficients of logistic regression to predict DSM-III-R dementia at Wave* 2 *from each cognitive score at Wave* 1, *adjusting for age*

| Wave 1 | *B* (S.E.) | OR | *P* |
|---|---|---|---|
| MMSE | −0·21 (0·07) | 0·81 | 0·003 |
| EMT | −0·31 (0·06) | 0·73 | 0·000 |
| SLMT | −0·03 (0·01) | 0·97 | 0·013 |
| NART | −0·03 (0·02) | 0·97 | 0·074 |

each measure with Wave 1 scores entered after age and the corresponding Wave 2 score. After adjusting for age and Wave 2 score, the only significant Wave 1 score was EMT (adjusted OR, Wave 2 EMT = 0·79, $P = 0·001$; adjusted OR, Wave 1 EMT = 0·82, $P = 0·047$). The odds of dementia increased by 1·9 for a unit decrease in the Wave 1 EMT and increased by 2·2 for a decline of 1 unit between waves. An alternative way of viewing this result is that both the level of Wave 1 performance and the rate of change on the EMT are predictors of dementia.

## DISCUSSION

### The sample

The sample was drawn from a defined population and covered a wide age span from 70–102 years. The response rate was close to that found in similar surveys (e.g. Brayne *et al*. 1995). In longitudinal studies, particularly of the elderly, there is non-random attrition of the sample, whereby those persons on whom follow-up data are missing because of death, refusal or non-availability are more likely to have been cognitively impaired at the start of the study. In this study, where the focus is on the prediction of cognitive change with age, the problem is further exacerbated since both mortality and the proportion of subjects providing incomplete data increase with age. Out of an original cohort of 118 subjects aged 90 and above, 63% had died and only 42% of the 36 subjects contacted at Wave 2 provided MMSE scores at both Waves. In contrast, among the cohort aged 70–74, 15% had died and 93% of those contacted provided complete MMSE data at Wave 2. Lower attrition rates might have been achieved with a shorter interval between assessments. The choice of 3–4 years was dictated by the wish to have a sufficiently long period in which to test real decline over and above change associated with random variation (van Belle *et al*. 1990; Tombaugh & McIntyre, 1992) and to avoid practice effects. However, it is likely that those among the very elderly completing the two assessments represent a cognitive élite. Any conclusions about change in cognitive performance of the very elderly, therefore, must be treated with caution; and the present findings are an underestimate of both cognitive decline and incident dementia. A number of respondents provided partial data for one or more of the cognitive tests, which were then adjusted by pro-rating. Analyses were repeated first by excluding those cases from analysis and so lowering the sample size and second by treating any missing items as errors. This maintained the sample size but scores were poorer for those with incomplete data. These analyses did not provide results which were substantively different from those presented here. Models using a Bayesian approach (e.g. Best *et al*. 1996) should provide more insight into the cognitive changes for the cohort as a whole, including those who died or were lost to follow-up. A third wave of assessments will allow explicit modelling of rates of change, including mortality, with respect to initial levels of performance.

### Cognitive tests

A strength of the present study is its use of a range of cognitive tests tapping several areas of cognitive function including verbal ability, memory and cognitive speed. Tests were included that do not suffer ceiling effects that limit the capacity to measure decline in the cognitively able.

### Change over time in MMSE scores

The distribution of the two sets of MMSE scores in Fig. 1 makes it clear that, across the whole sample, there was a decline. The distribution of change in the scores (Fig. 2) was close to normal in this community sample of very elderly persons. Most of the sample changed little over time, while a proportion declined and a minority improved. This pattern is strikingly similar to that found by Brayne *et al*. (1995) in Cambridge. Feskens *et al*. (1994) found a mean decline of 0·3 points in a sample of 377 elderly men over a period of 3 years. On the other hand, in the ECA studies in New Haven, 360 subjects aged 80 and

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

above who were assessed 3 times in 12 months with the MMSE showed overall improvement of 1·6 points; and 1627 subjects aged 60–79 improved by an average of 0·7 points (Bruce *et al*. 1995). In contrast to the Canberra findings, the latter results suggest practice effects, especially given the short time intervals between assesments.

The measurement of change on the MMSE is restricted by scaling artefacts, in particular the ceiling score of 30 achieved by the majority of normal, unimpaired persons. Half of those scoring 30 at Wave 1 scored 29 or 30 at Wave 2. Among these subjects, therefore, a real decline in the level of cognitive functioning could be masked.

The interaction between age and initial level of cognitive impairment, as shown in Fig. 3, carries some relevance for assessing prognosis in clinical practice. An impaired MMSE score in the very elderly carries a worse prognosis than an equivalent score in the younger elderly.

### Change in SLMT

The SLMT tests cognitive speed and requires good sensory functioning. Tables 2 and 3 show that, cross-sectionally, scores decreased rapidly with age. Decline between waves also increased rapidly for the younger age groups, but appeared to level out for the oldest age groups, suggesting that decline on speeded tasks may precede decline in other areas; that subjects reach a floor on this type of test; that decline becomes so rapid among the very elderly that they are no longer able to complete the test; or that they lack the verbal skills necessary to follow the instructions or to name letters. This may also be an effect of measurement error, given the small numbers of very old completing the test.

### Changes in the EMT

On average, the EMT showed improvement. This could be a learning effect. It was mostly due to improved performance on the face recognition task among the younger members of the cohort. For this test, subjects are asked to look at pictures of five faces and say whether they found each pleasant or unpleasant. Later in the interview, a surprise memory task is included, where they are asked to identify those five faces among 15. At Wave 2, they may have remembered the strategy involved in this task from their previous interview and so paid particular attention to the faces when they were next presented. That is, a task involving incidental learning at Wave 1 became one of intentional learning at Wave 2, and it is well recognized that intentional learning is associated with higher recognition rates (Perlmutter & Mitchell, 1982). This could also explain why Wave 1 EMT score and change both predict dementia, while changes on other tests were not predictive. The older age groups showed mean decline; and decline increased rapidly with age. A reason for lack of improvement among the older age groups could be deteriorating sensory function, which has been shown, cross-sectionally, to be associated with age and performance on these same cognitive tests (Christensen *et al*. 1996).

### Changes in the NART

Neither the NART scores nor the change scores varied significantly with age, providing support for the usefulness of the NART as a measure of pre-morbid functioning. It showed greater decline among the incident dementia cases, confirming that it may underestimate IQ in deteriorating AD patients (Crawford *et al*. 1990).

### Differences by sex

Unlike some other studies (Bickel & Cooper, 1994, Brayne *et al*. 1995, Prince *et al*. 1996) but like others (Evans *et al*. 1993, Farmer *et al*. 1995), no sex differences were found among those assessed at both waves for any of the cross-sectional measures of cognitive performance, for the change in scores between waves or for incident dementia. The one exception was the NART which was slightly lower in women compared with men at both waves, although there was no sex difference in the change on the NART. This no doubt reflects the lower educational level of women in the study (Christensen *et al*. 1997). The original sample was selected to have approximately equal numbers of men and women in each age group, although it was drawn from a population in which women were more heavily represented than men among the oldest old (Henderson *et al*. 1993). This, therefore, diminishes the possibility that some age-associated changes are mistakenly construed as associated with sex, a possibility if an analysis were based on broad age-bands, with no men

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

among the oldest old. On the other hand, refusal rates were higher among women and refusal is associated with poorer cognitive performance, so any real sex differences could have been masked.

### Cognitive decline and DSM-III-R dementia

Although mean decline on each cognitive test was indeed greater among those with incident dementia, it is notable that not all respondents with incident dementia showed decline on all of the measures and there was considerable inter-individual variation in the areas which showed decline, irrespective of age. Risk of dementia increased with age, but no one cognitive area consistently deteriorated more rapidly than others, except for the NART, which remained stable. The MMSE is one of the most widely used screening tests for possible dementia, and for this reason deserves particular attention. Although decline was predictive of dementia, and low Wave 1 scores predicted incident dementia, decline did not improve the prediction over that provided by the current score. This was also found for the SLMT. That is, the actual level reached and not the rate at which it is reached is the better predictor. This is similar to the findings of Drachman *et al.* (1990), who found that the level of severity was more important than the rate of progression in the prognosis of Alzheimer's disease. The EMT was found to be the test most strongly associated with incident dementia in this sample, confirming the findings of Masur *et al.* (1994) that tests of delayed recall and visuospatial tasks are the best pre-clinical predictors of dementia. For the EMT, change in cognitive performance as well as current level of performance was related to dementia.

Neither the initial level in cognitive scores nor the observed change was strongly associated with a dementia diagnosis. This could be due partly to selective attrition, and partly because the data do not allow the assessment of the onset of decline. Those who were already declining gradually at Wave 1 will have been lost to follow-up or will have shown greater overall decline than those with a rapid rate of decline that commenced only shortly before Wave 2. Only multiple assessments with short time intervals could overcome this problem.

Prince *et al.* (1996) studied 2584 subjects aged 65–74 and found that NART scores (which they labelled 'intelligence') and education were predictive of decline on the Paired Associate Learning Test, but not of incident dementia of the Alzheimer's type. They used this as an argument to suggest that different aetiologies could be operating for cognitive decline and dementia. However, a number of other investigators have found education to be associated with incident dementia (Katzman, 1993; Bickel & Cooper, 1994). In the present data, the NART at Wave 1 was not related to change in any of the other tests, adjusting for age, while the relationship of education to cognitive decline was not uniform across all cognitive functions. Christensen *et al.* (1997) have found that a higher level of education predicted less decline on tasks tapping crystallized intelligence (MMSE and NART) than for the EMT or the SLMT which involve mental speed and memory. In agreement with Prince *et al.* (1996), however, we found that, after adjusting for age, the NART scores were unrelated to incident dementia.

### Cognitive decline among the oldest old

There is considerable debate about whether decline continues into extreme old age, or whether a plateau exists beyond which no further decline occurs and dementia incidence and prevalence no longer increase (e.g. Wernicke & Reischies, 1994). Like other studies of the elderly, our data do not, as yet, allow us to determine if such a plateau exists, since small samples over age 90 result in estimates of mean decline with large standard errors. Although the decline on the SLMT appeared to flatten after about age 85, both the MMSE and EMT showed increasing decline through to the oldest age group, and the NART remained fairly stable across all age groups. Our data have established that compared with those in younger age cohorts, decline is greater and almost global among the oldest old. Only two respondents out of 51 aged 85 and above failed to show decline on any of the tests attempted and few individuals improved on any of the tests. While there is increasing variability in performance into late old age (Christensen *et al.* 1994*b*), overall mean performance levels are lower. Mortality, sample attrition and the incidence of dementia are associated with poor test performance at all ages, but increase rapidly

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

with age. Accelerated decline among the very elderly may also be a cause of sample attrition. This can be tested only with data from a third assessment.

## CONCLUSION

Across the population of elderly aged 70 and above, there was an overall decline in cognitive performance on a range of tasks over a period of 3·6 years. Change was approximately normally distributed, with most subjects declining or remaining the same, while some improved. Improvements were most apparent for the younger age groups, and are probably due to a combination of measurement error and practice effects. The NART was stable across all ages. Decline on the MMSE, SLMT and EMT became greater with age, so that a poor score at Wave 1 in the older age groups carried a worse prognosis than in the younger, especially for the MMSE. Whether the rate of decline increases or not into the tenth decade, it appears that some impairment is almost universal among the oldest old and that decline is global. The probability of incident dementia is related to poor scores on the cognitive tests, but not to the rate of decline, except possibly for the EMT. There were no differences between men and women on cognitive performance, cognitive decline or incident dementia. A third wave of assessments, planned as a 7-year follow-up, will allow an examination of outcome in terms of both the rate of decline and initial levels of performance.

This study was assisted by a grant from the Australian Rotary Health Research Fund, which is acknowledged with appreciation. Valuable help in the planning and management of the survey was provided by Colleen Doyle, Ruth Scott, Susan Lindsay and Suzanne Dee.

## REFERENCES

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders*, *3rd edn*, *Revised*. American Psychiatric Association: Washington, DC.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*, *4th edn*. American Psychiatric Association: Washington, DC.

Berkson, J. (1946). Limitations of the application of fourfold table analysis to hospital data. *Biometrics Bulletin* **2**, 47–53.

Best, N. G., Spiegelhalter, D. J., Thomas, A., & Brayne, C. E. G. (1996). Bayesian analysis of realistically complex models. *Journal of the Royal Statistical Society*, *Series A* **159**, 323–342.

Bickel, H. & Cooper, B. (1994). Incidence and relative risk of dementia in an urban elderly population: findings of a prospective field study. *Psychological Medicine* **24**, 179–192.

Braekhus, A., Laake, K. & Engedal, K. (1995). A low, 'normal' score on the Mini-Mental State Examination predicts development of dementia after three years. *Journal of the American Geriatrics Society* **43**, 656–661.

Brayne, C., Gill, C., Paykel, E. S., Huppert, F. & O'Connor, D. W. (1995). Cognitive decline in an elderly population – a two wave study of change. *Psychological Medicine* **25**, 673–683.

Bruce, M. L., Hoff, R. A., Jacobs, S. C. & Leaf, P. J. (1995). The effects of cognitive impairment on 9-year mortality in a community sample. *Journal of Gerontology*: *Psychological Sciences* **50B**, P289-P296.

Christensen, H., Jorm, A. F., Henderson, A. S., Mackinnon, A. J., Korten, A. E. & Scott, L. R. (1994*a*). The relationship between health and cognitive functioning in a sample of elderly people in the community. *Age and Ageing* **23**, 204–212.

Christensen, H., Mackinnon, A. J., Jorm, A. F., Henderson, A. S., Scott, L. R. & Korten, A. E. (1994*b*). Age differences and interindividual variation in cognition in community-dwelling elderly. *Psychology and Aging* **9**, 381–390.

Christensen, H., Korten, A., Jorm, A. F., Henderson, A. S., Scott, R. & Mackinnon, A. J. (1996). Activity levels and cognitive functioning in an elderly community sample. *Age and Ageing* **25**, 72–80.

Christensen, H., Korten, A. E., Jorm, A. F., Henderson, A. S. & Jacomb, P. A. (1997). Change in cognitive function and education: differential effects due to type of cognitive test. *International Journal of Geriatric Psychiatry* (in the press).

Cooper, B. (1995). Epidemiology of dementia and cognitive decline in the elderly. *Ballière's Clinical Psychiatry* **1**, 175–184.

Crawford, J. R., Cochrane, R. H., Besson, J. A. & Parker, D. M. (1990). Premorbid IQ estimates obtained by combining the NART and demographic variables: Construct validity. *Personality and Individual Differences* **11**, 209–210.

Drachman, D. A., O'Donnell, B. F., Lew, R. A. & Swearer, J. M. (1990). The prognosis in Alzheimer's disease. 'How far' rather than 'How fast' best predicts course. *Archives of Neurology* **47**, 851–856.

Evans, D. A., Beckett, L. A., Albert, M. S., Hebert, L. E., Scherr, P. A., Funkenstein, H. H. & Taylor, J. O. (1993). Level of education and change in cognitive function in a community sample of older persons. *Annals of Epidemiology* **3**, 71–77.

Farmer, M. E., Kittner, S. J., Rae, D. S., Bartko, J. J. & Regier, D. A. (1995). Education and change in cognitive function. The Epidemiologic Catchment Area study. *Annals of Epidemiology* **5**, 1–7.

Feskens, E. J. M., Havekes, L. M., Kalmijn, S., de Knijff, P., Launer, L. J. & Kromhout, D. (1994). Apolipoprotein $\epsilon$4 allele and cognitive decline in elderly men. *British Medical Journal* **309**, 1202–1206.

Folstein, M. F., Folstein, S. E. & McHugh, P. R. (1975). 'Mini-Mental State': a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* **12**, 189–198.

Henderson, A. S., Jorm, A. F., Mackinnon, A. J., Christensen, H., Scott, L. R., Korten, A. E. & Doyle, C. P. (1993). The prevalence of depressive disorders and the distribution of depressive symptoms in later life: a survey using draft ICD-10 and DSM-III-R. *Psychological Medicine* **23**, 719–729.

Henderson, A. S., Jorm, A. F., Mackinnon, A. J., Christensen, H., Scott, L. R., Korten, A. E. & Doyle, C. (1994). A survey of dementia in the Canberra population: experience with ICD-10 and DSM-III-R criteria. *Psychological Medicine* **24**, 473–482.

Huppert, F. A. & Brayne, C. (1994). What is the relationship between dementia and normal aging? In *Dementia and Normal Aging*. (Ed. F. A. Huppert, C. Brayne & D. W. O'Connor), pp. 3–11. Cambridge University Press: Cambridge.

Jorm, A. F. (1992). Use of informants' reports to study memory changes in dementia. In *Memory Functioning in Dementia* (ed. L. Bäckman), pp. 267–282. Elsevier: Amsterdam.

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press

Izaks, G. J., Gussekloo, J., Dermout, K. M. T., Heeren, T. J. & Ligthart, G. J. (1995). Three-year follow-up of Mini-Mental State Examination score in community residents aged 85 and over. *Psychological Medicine* **25**, 841–848.

Katzman, R. (1993). Education and the prevalence of dementia and Alzheimer's disease. *Neurology* **43**, 13–20.

Lezak, M. D. (1995). *Neuropsychological Assessment*, 3*rd edn*. Oxford University Press: Oxford.

Mackinnon, A., Christensen, H., Cullen, J. S., Doyle, C. J., Henderson, A. S., Jorm, A. F., Korten, A. E. & Scott, L. R. (1993). The Canberra Interview for the Elderly: assessment of its validity in the diagnosis of dementia and depression. *Acta Psychiatrica Scandinavica* **87**, 146–151.

Masur, D. M., Sliwinski, M., Lipton, R. B., Blau, A. D. & Crystal, H. A. (1994). Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. *Neurology* **44**, 1427–1432.

Nelson, H. E. (1982). *The Nelson Adult Reading Test* (*NART*). NFER-Nelson: Windsor.

Perlmutter, M. & Mitchell, D. B. (1982). The appearance and disappearance of age differences in adult memory. In *Ageing and Cognitive Processing* (ed F. I. M. Craik and S. Trehub), pp. 127–144, Plenum Press: New York.

Plewis, I. (1985). *Analysing Change. Measurement and Explanation Using Longitudinal Data*. John Wiley & Sons: Chichester.

Prince, M., Lewis, G., Bird, A., Blizard, R. & Mann, A. (1996). A longitudinal study of factors predicting change in cognitive test scores over time, in an older hypertensive population. *Psychological Medicine* **25**, 555–568.

Roth, M. (1994). The relationship between dementia and normal aging of the brain. In: *Dementia and Normal Aging* (ed. F. A. Huppert, C. Brayne and D. W. O'Connor), pp. 57–76, Cambridge University Press: Cambridge.

Social Psychiatry Research Unit (1992). The Canberra Interview for the Elderly: a new field instrument for the diagnosis of dementia and depression by ICD-10 and DSM-III-R. *Acta Psychiatrica Scandinavica* **85**, 105–113.

Tombaugh, T. N. & McIntyre, N. J. (1992). The Mini-Mental State Examination: a comprehensive review. *Journal of the American Geriatric Society* **40**, 922–935.

van Belle, G., Uhlmann, R. F., Hughes, J. P. & Larson, E. B. (1990). Reliability of estimates of changes in mental status test performance in senile dementia of the Alzheimer type. *Journal of Clinical Epidemiology* **43**, 589–595.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale – Revised Manual*. Psychological Corporation: New York.

Wernicke, T. F. & Reischies, F. M. (1994). Prevalence of dementia in old age: clinical diagnosis in subjects aged 95 years and older. *Neurology* **44**, 250–253.

World Health Organization (1993). *The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic Criteria for Research*. World Health Organization: Geneva.

https://doi.org/10.1017/S0033291797005217 Published online by Cambridge University Press