The combination of cognitive testing and an informant questionnaire in screening for dementia

RICHARD KNAFELC1, DINA LO GIUDICE2, SUSAN HARRIGAN3, ROSLYN COOK2, LEON FLICKER4, ANDREW MACKINNON5, DAVID AMES3

1Bundoora Extended Care Centre, Bundoora, Victoria, Australia
2National Ageing Research Institute, Melbourne Extended Care and Rehabilitation Service, Victoria, Australia
3Department of Psychiatry, University of Melbourne, Royal Melbourne Hospital, 7th Floor Charles Connibere Building, Grattan Street, Parkville, Victoria 3050, Australia
4University of Western Australia, Royal Perth Hospital, Australia
5Department of Psychological Medicine, Monash University, Mental Health Research Institute, Victoria, Australia

Address correspondence to: D. Ames. Fax: (+61) 3 9342 8954. Email: cain@unimelb.edu.au

Abstract

Aim: a cognitive test and an informant report questionnaire were used together to investigate whether their combined use could improve the accuracy of detecting dementia in a memory clinic, compared with either test used alone. Method: the subjects were 323 patients assessed at a memory clinic. The Mini-Mental State Examination and the short form of the Informant Questionnaire on Cognitive Decline in the Elderly were used. A method of combining the test scores developed by Mackinnon and Mulligan [Am J Psychiatry 1998; 155: 1529–35] was used. Dementia was diagnosed according to Diagnostic and Statistical Manual of Mental Disorders, Third Edition - Revised criteria. Results: logistic regression analysis showed that the combination of the Mini-Mental State Examination and the Informant Questionnaire on Cognitive Decline in the Elderly produced a slightly more accurate prediction of dementia caseness than either test used alone. Using receiver operating characteristic analysis the performance of the combination of the tests according to a weighted sum rule was compared with the performance of either test used alone. This way of combining the tests resulted in a more accurate screening for dementia than when the Informant Questionnaire on Cognitive Decline in the Elderly was used alone. The area under the receiver operating characteristic curve for the Mini-Mental State Examination combined with the Informant Questionnaire on Cognitive Decline in the Elderly was 0.89 compared with 0.82 for the Informant Questionnaire on Cognitive Decline in the Elderly alone (chi-square = 10.63; P= 0.0011). However, there was no improvement in screening performance when the combination of Mini-Mental State Examination and Informant Questionnaire on Cognitive Decline in the Elderly was compared with the Mini-Mental State Examination used alone (area under the receiver operating characteristic curve = 0.89 versus 0.86; chi-square = 3.54; P= 0.060).

Conclusion: the combination of cognitive testing and an informant report according to a weighted sum rule in this population did not result in any advantage over the use of the Mini-Mental State Examination alone. The mixed results of this study contrast with those of Mackinnon and Mulligan.

Keywords: dementia, screening, mini-mental state examination, IQ code

Introduction

Direct testing of the patient’s cognition by a trained rater using an established test is the commonest form of screening for dementia, but informant reports are an alternate method of detecting dementia [1]. Cognitive tests and informant reports in dementia screening do not necessarily measure the same characteristics of a subject [1]. Correlations between the two types of instrument are moderate [2]. Moreover, Jorm et al. [3] in analyses of cognitive test and informant report items have found that there are separate factors in each item domain. The two types of instrument measure rather different things.

Cognitive tests and informant reports have complementary advantages and disadvantages. Cognitive tests
can be influenced by education and pre-morbid ability [4] and require subjects to have intact language, sensory and motor abilities [5]. Informant questionnaires are unaffected by the subject’s education and pre-morbid intelligence [6, 7], but performance may be influenced by both the personality and the affective state of both the subject and informant and the nature of their relationship [8].

Given the complementary aspects of cognitive tests and informant reports, a combination of such instruments might improve the performance of dementia screening. There has been little work on this and a study by Mackinnon and Mulligan [1] is the only published example. They used the Mini-Mental State Examination (MMSE) [9] and the short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) [10, 11] in a Swiss geriatric hospital and memory clinic sample numbering 106.

Using logistic regression to predict Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) [12] dementia caseness, Mackinnon and Mulligan found that the MMSE score alone was a strong and statistically significant predictor of caseness [1]. Prediction of caseness was significantly improved when IQCODE score was added to the equation. Each test gave non-superfluous and complementary information in dementia screening. They used receiver operating characteristic (ROC) analysis to plot curves for each test. Sensitivity, specificity and positive and negative predictive values were calculated for a number of test cut-points. They combined the scores for the MMSE and the IQCODE using three different procedures: the 'or' rule, 'and' rule and the 'weighted sum' rule.

The 'or' rule defines a case of dementia if either the MMSE or IQCODE are positive. Mackinnon and Mulligan calculated sensitivity and specificity across all possible cut-points of the IQCODE while applying the 'or' rule with fixed cut-points for the MMSE. This procedure was repeated for a number of MMSE cut-points and resulted in one ROC curve for each MMSE cut-point. The performance of the tests could be improved with the MMSE at cut-points between 22 and 24 used together with the IQCODE. Using the IQCODE at cut-points greater than 4.0 together with the MMSE at the 23/24 cut-point significantly increased sensitivity compared with using the MMSE alone. Specificity was not reduced significantly by this process.

Mackinnon and Mulligan tested the 'and' rule in the same way. It specifies that only patients positive on both tests are classified with dementia. The performance of most combinations of the MMSE and IQCODE could be matched or improved upon by using the MMSE alone and changing the cut-point.

With the 'weighted sum' rule the scores of the two tests are combined mathematically, yielding a score on a new scale, before the cut-point is calculated. Mackinnon and Mulligan derived the formula for the weighted sum rule from their logistic regression analysis and constructed ROC curves. Test performance expressed in terms of area under the ROC curve was significantly better for the weighted sum of the MMSE and the IQCODE compared with the performance of either the MMSE or the IQCODE used alone. There were no significant differences found between the 'weighted sum' rule and the 'or' rule.

Thus the authors concluded: (i) the informant report and cognitive test instruments used were complementary in screening for dementia in a clinical population. Combining tests yielded more information than either test used alone. (ii) The 'weighted sum' rule enabled a graduated trade-off between the scales. The more evidence that a patient is not impaired on one test, the more evidence of impairment is required on the other test to qualify as a case. (iii) In situations where screening is paramount, the use of the 'or' rule is justified since only sensitivity is increased. (iv) The use of the 'weighted sum' rule may be preferable in cases where both sensitivity and specificity need to be optimised or balanced. (v) Using the 'weighted sum' rule with a cut-point of 0.5 probability of dementia, the authors constructed an easily applied template or 'DemeGraph' upon which the patient's MMSE and IQCODE scores can be plotted to determine whether the presence of dementia is likely.

The combination of tests used by Mackinnon and Mulligan might enhance the efficiency of screening, but their results need to be replicated in other larger populations. The aim of the current study was to answer the following question: using Mackinnon and Mulligan’s statistical model [1], how does the 'weighted sum' combination of MMSE and IQCODE perform in differentiating dementia from non-dementia in another clinical population, compared with the performance of either test used alone? We expected similar results to those of Mackinnon and Mulligan.

Method

Memory clinic and subjects

The subjects attended a Memory Clinic at a geriatric hospital in Melbourne from 1988–1998 [13, 14]. Patients assessed by medical staff at the first visit were tested with a number of instruments including the Cambridge Cognitive Examination (CAMDEX) [15] within which the MMSE is incorporated. Allied health staff saw the carer(s) to obtain an informant history and instruct them in completing the IQCODE. All data for this study were obtained from the initial appointment.

Mackinnon and Mulligan [1] used a French-speaking population tested in French. To enable a direct comparison between our results and theirs, those subjects whose first language was not English were excluded from this study.
Statistical analysis

Group comparisons were made using a range of parametric and non-parametric statistical procedures. Assumptions of parametric analyses were evaluated before analysis, including assessment of the degree of skewness and kurtosis of each variable. Variables were reasonably normally distributed, thus allowing the use of parametric procedures.

Differences on continuous data were assessed using independent samples t-tests, with unequal variances tests being used where appropriate. One-way analysis of variance was used to test for differences between multiple groups, with appropriate post hoc tests performed to identify specific differences. Chi-square tests of significance were used to test for significant associations between cross-tabulated data, and exact tests were used where appropriate. Pearson correlations were used to assess associations between continuous variables. Statistical tests were 2-tailed and results were regarded as significant at or below the 5% probability level.

Logistic regression analysis was used to explore whether using the two tests together resulted in any additional information in the prediction of dementia compared with the use of the tests separately. The following equation was derived from the logistic regression analysis:

\[
\text{Logit (case)} = 1.4508 + 1.5819 \times \text{IQCODE} - 0.3119 \times \text{MMSE}
\]

where ‘Logit (case)’ is the logarithm of the odds of a subject being a case of dementia, which is equal to: log (probability of a case/probability of non-case). ‘Pr (case)’ is the probability of a case of dementia. IQCODE and MMSE are the test score values. A comparison of these equations to those used by Mackinnon and Mulligan [1] reveals that different coefficients have been used here because the derivation of these values is influenced by the characteristics of the sample.

Using ROC analysis, curves were plotted of sensitivity versus 1 minus specificity for all possible cut-off scores of each test. The area under the curve (AUC) for each test was calculated using a non-parametric method. The accuracy of the test in screening for dementia is represented by, and is directly proportional to, the area under the curve. Thus the accuracy of the tests can be simply contrasted by comparing the respective AUCs. Furthermore, a ROC curve was created for the weighted sum of the MMSE and the IQCODE as described in the logistic regression equation, to investigate the performance of combining these tests thus.

Thus ROC curves were constructed for each of the tests and the combination of the tests in the differentiation of the following diagnostic groups in the English-speaking portion of the sample: (i) dementia versus non-dementia; and (ii) dementia versus cognitive impairment other than dementia.

ROC analysis and logistic regression can be regarded as complementary techniques, with some points of difference and some points in common. Logistic regression assesses whether there is a relationship between caseness and one or more predictor variables and gives the optimal equation for predicting probability of caseness. ROC analysis assesses the screening performance of a test by calculating the sensitivity and specificity of each cut-point of the test against caseness and graphically represents the tradeoffs of each cut-point. The use of logistic regression to optimally combine predictors and ROC analysis to represent the cut-point tradeoffs marries the advantages of both techniques. The two techniques will usually yield similar findings. In logistic regression, if a predictor is significantly related to caseness, then the
AUC test of the performance of the screening instrument in ROC is also likely to be significant. However, because the AUC test in ROC analysis is slightly less powerful, being a non-parametric test, it is possible to observe discrepancies in findings resulting from the two techniques.

Results

Subjects

Out of a total of 519 subjects 426 had both the MMSE and IQCODE completed. The 103 non-English speakers were excluded, resulting in a sample of 323 English-speaking subjects, (120 males and 203 females). Some data on 299 of these subjects were utilised in a previous paper which did not address the Mackinnon and Mulligan model [17]. Their ages ranged from 44–93 years (mean = 74.7 years, SD = 8.8). In the sample there were 9 normal subjects, 229 with dementia, 57 with cognitive impairments other than dementia (mainly amnestic disorders (n = 22) and other cognitive impairments not fulfilling DSM-III-R dementia criteria (n = 31) as well as 2 organic mood disorders, 1 organic anxiety disorder, 1 organic personality disorder and 28 with functional psychiatric diagnoses of whom 21 had depression.

Of those in the sample with data for educational status (n = 310), 13.2% had completed schooling at the age of 17 years and older and 31.0% had completed school aged 16 years and older.

MMSE and IQCODE scores

MMSE scores ranged from 4–30 (mean = 20.65; SD = 5.44); IQCODE scores ranged from 2.56–5.00 (mean = 4.22; SD = 0.57). Independent samples t-tests indicated that the dementia group (mean MMSE 18.86) was more impaired than the non-dementia (mean MMSE 25.02) group on MMSE (mean difference = 6.17; t = 12.40; df = 242.06; P < 0.001) and IQCODE (mean dementia group IQ CODE 4.41; non-dementia mean 3.75; mean difference = 0.66; t = −11.14; df = 321; P < 0.001). A plot of MMSE versus IQCODE scores is shown in Figure 1. The Pearson correlation coefficient between the two tests was −0.54 (P < 0.001).

Logistic regression analysis

Table 1 shows the results of the logistic regression analysis predicting DSM-III-R dementia caseness from MMSE and IQCODE scores for the sample (n = 323).

Table 1. Logistic regression of DSM-III-R dementia status against MMSE and IQCODE for the sample (n = 323)

<table>
<thead>
<tr>
<th>Term added</th>
<th>Change in G²</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>119.00</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MMSE²</td>
<td>6.69</td>
<td>0.0097</td>
</tr>
<tr>
<td>IQCODE</td>
<td>26.35</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>IQCODE²</td>
<td>0.07</td>
<td>0.80</td>
</tr>
<tr>
<td>IQCODE × MMSE</td>
<td>0.77</td>
<td>0.38</td>
</tr>
<tr>
<td>IQCODE</td>
<td>95.08</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>IQCODE²</td>
<td>0.12</td>
<td>0.73</td>
</tr>
<tr>
<td>MMSE</td>
<td>51.23</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>MMSE²</td>
<td>5.69</td>
<td>0.017</td>
</tr>
<tr>
<td>IQCODE × MMSE</td>
<td>0.77</td>
<td>0.38</td>
</tr>
</tbody>
</table>

All significance tests have 1 degree of freedom.
Screening for dementia

The scores of $G^2$ change for MMSE plus $G^2$ change for IQCODE in the top half of Table 1 add up to the same value as those in the bottom half of the table.

To determine whether there is a non-linear relationship between dementia caseness and MMSE, the quadratic term $MMSE^2$ was added to the model. A very small but statistically significant $G^2$ change of 6.69 was the result ($P = 0.0097$). A slightly different $G^2$ change of 5.69 (still significant) was seen when the $MMSE^2$ was added into the model after the IQCODE (shown in the bottom half of the table). In Mackinnon and Mulligan’s [1] study the MMSE$^2$ term was small and not significant.

To explore whether there is an interaction occurring between the MMSE and the IQCODE, the multiplicative term ‘MMSE $\times$ IQCODE’ was used. This resulted in a non-significant value implying that there is no such interaction.

Receiver operating characteristic analysis

Figure 2 shows the results of the ROC analysis in differentiating dementia ($n = 229$) from non-dementia (all the other diagnoses including normal cases; $n = 94$) for the sample. The weighted sum rule of combining the MMSE with the IQCODE is derived from the logistic regression equation shown in equations 1 and 2. It was found that the AUC for the weighted sum rule was significantly greater than that for the IQCODE alone (0.89 versus 0.82; chi-square = 10.63; $P = 0.0011$). However there was no significant difference between the AUC for the weighted sum rule and that for the MMSE used alone (0.89 versus 0.86; chi-square = 3.54; $P = 0.060$). Thus for this sample the weighted sum rule holds little or no advantage in dementia screening compared with the MMSE. There was no significant difference between the AUCs for the MMSE and IQCODE (chi-square = 1.85; $P = 0.17$).

The results for the ROC analysis in differentiating dementia ($n = 229$) from cognitive impairment other than dementia ($n = 57$) showed no significant differences between the AUCs for the weighted sum rule and the MMSE or between the MMSE and the IQCODE but there was a significant difference between the area under the curve for the weighted sum rule and that for the IQCODE (0.88 versus 0.82; chi-square = 8.33; $P = 0.0039$).

The results for sensitivity, specificity and positive and negative predictive values for the MMSE, IQCODE and the weighted sum rule in differentiating dementia from non-dementia are shown in Table 2.

Table 2. Performance of individual tests and the combination of tests in dementia screening: differentiation of dementia ($n = 229$) from non-dementia ($n = 94$)

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive predictive value</th>
<th>Negative predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE (23/24 cut-point)</td>
<td>0.84 (0.78–0.88)</td>
<td>0.73 (0.63–0.82)</td>
<td>0.89 (0.84–0.92)</td>
<td>0.65 (0.55–0.74)</td>
</tr>
<tr>
<td>IQCODE ($&lt; 3.6/\geq 3.6$ cut-point)</td>
<td>0.94 (0.91–0.97)</td>
<td>0.47 (0.36–0.57)</td>
<td>0.81 (0.76–0.86)</td>
<td>0.77 (0.64–0.87)</td>
</tr>
<tr>
<td>Weighted sum [Pr (case) ≥ 0.50]</td>
<td>0.91 (0.86–0.94)</td>
<td>0.63 (0.52–0.73)</td>
<td>0.86 (0.81–0.90)</td>
<td>0.74 (0.63–0.83)</td>
</tr>
</tbody>
</table>

The 95% confidence intervals are shown in parentheses next to the respective values.
Discussion

Hypothesis

The main aim was to explore whether the combination of the MMSE and IQCODE in differentiating dementia from non-dementia in a group of patients attending a memory clinic was superior to that of either test used alone. The hypothesis was that results similar to those of Mackinnon and Mulligan [1] would be found. Similar, but not identical results to theirs were obtained.

The study samples

Our sample was three times larger than the Swiss one. It contained proportionately fewer normal individuals and more with cognitive impairments other than dementia than the Geneva cohort. There was a lower rate of secondary school completion at the 16-year-old level in our sample.

Correlations

The correlation coefficient value for the relationship between the MMSE and the IQCODE (−0.54) was similar to Mackinnon and Mulligan’s figure of −0.58, but lower than the correlations of −0.74 to −0.78 found by Jorm and colleagues [7, 10].

Test performance

The results for sensitivity, specificity and positive and negative predictive value are similar to those in other studies [5] including Mackinnon and Mulligan [1]. We found the sensitivity of the MMSE with the 23/24 cut-point was higher than that found by Mackinnon and Mulligan (0.84 versus 0.76 respectively), whilst the specificity in our study was lower (0.73 versus 0.90). For the IQCODE (< 3.6/ ≥ 3.6 cut-point) Mackinnon and Mulligan found a sensitivity of 0.90 (compared with 0.94 for our study) but they had a greater specificity of 0.65 compared with only 0.47 for us. The weighted sum rule (Pr (case) ≥ 0.50) performed better in Mackinnon and Mulligan’s study than in ours — they found a sensitivity of 0.93 and a specificity of 0.85 compared with 0.91 and 0.63 respectively. They found the weighted sum rule (using the Pr (case) ≥ 0.50 cut-point) to perform better than either the MMSE or the IQCODE alone.

Our results were less clear-cut. The weighted sum rule (Pr (case) ≥ 0.50 cut-point) had a higher sensitivity than the MMSE did (23/24 cut point) (0.91 versus 0.84) but lower specificity (0.62 versus 0.73). The IQCODE (< 3.6/ ≥ 3.6 cut-point) had higher sensitivity than either the weighted sum rule or the MMSE alone (0.94) but had much lower specificity (0.47) than either of these. In our clinic the MMSE at the 23/24 cut-point would be an appropriate choice of test where both sensitivity and specificity need to be optimised. However, if sensitivity needs to be maximised at the expense of specificity (in screening for instance) then either the weighted sum rule at the Pr (case) ≥ 0.50 cut-point or the IQCODE at the < 3.6/ ≥ 3.6 cut-point would be appropriate.

In general we found quite high sensitivity and positive predictive values with less impressive specificity and negative predictive values. A possible reason for the lower specificity values in this study could be that there was a larger proportion of cognitive impairment other than dementia in the non-dementia category than in Mackinnon and Mulligan’s study, and fewer normal individuals, making it more difficult for the tests to differentiate non-cases of dementia from cases. Furthermore, in nearly all-screening situations, for higher levels of sensitivity, the specificity falls off disproportionately.

Logistic regression analysis

For the logistic regression analysis, we had similar results to Mackinnon and Mulligan [1]. In both studies MMSE score alone was a strong and statistically significant predictor of dementia caseness and the addition of the IQCODE significantly improved this prediction. It was seen for the model incorporating the MMSE followed by the IQCODE and these results were mirrored in those of the model using the IQCODE followed by the MMSE. This is further evidence that these two tests provide somewhat different and additive information in screening for dementia.

However, there were some points of difference with Mackinnon and Mulligan. The quadratic term of MMSE, although small, did reach statistical significance in our study. This would imply a non-linear relationship in the equation, though this effect would be very small in comparison with the much larger effects of the other significant components in the model.

Our study used DSM-III-R criteria [16] to make diagnoses (as DSM-IV [14] was not published when data collection commenced), whilst Mackinnon and Mulligan [1] used DSM-IV. However, given the minor differences between the two systems with regard to the diagnosis of dementia, it is unlikely that differences in our results could have been entirely explained by this factor.

Receiver operating characteristic analysis

For the ROC analysis, (a non-parametric test which is less powerful than logistic regression) in all of the comparisons there was no significant difference between the performance of the MMSE and the IQCODE as measured in terms of area under the ROC curves. This was also the case for Mackinnon and Mulligan [1].

Some quite substantial AUCs were found (0.86 for the MMSE and 0.82 for the IQCODE in differentiating dementia from non-dementia). Even higher AUCs were found in the more clear-cut differentiation between dementia and normal cases.
Screening for dementia

The combination of MMSE and IQCODE according to the weighted sum rule performed slightly better than either test alone in differentiating dementia from non-dementia. For the other comparisons there was no advantage in using the weighted sum combination in preference to either the MMSE or IQCODE alone.

By the ROC analysis, the combination of the MMSE and the IQCODE using the weighted sum rule performed better than the IQCODE alone in dementia screening. However, the combination did not perform better than the MMSE alone. This is in contrast to Mackinnon and Mulligan’s result in their population. With larger sample sizes more of the results may have been statistically significant.

Improvements in screening performance resulting from combining tests may be quite sensitive to the characteristics of the setting in which screening is undertaken.

There does not appear to be any clear advantage in the everyday combination of these two tests according to the weighted sum rule in our clinic, since the MMSE performed just as well as the combined tests here. However, in populations with a higher percentage of normal subjects (e.g. general practice attenders) it is likely that the weighted sum rule may be of greater utility. Mackinnon and Mulligan’s simple weighted sum rule should be further assessed in populations of older people that include more normal subjects.

Key points
- 323 Australian memory clinic patients were assessed with two dementia screening instruments.
- A combination of an informant report (IQ code) and a direct screen (Mini-Mental State Examination) performed better than the informant report alone, but no better than the direct screen alone.

Acknowledgements
We would like to thank all the staff who worked in the clinic, and the patients and their carers for their help and cooperation.

References

Received 5 April 2002; accepted in revised form 20 March 2003

Downloaded from https://academic.oup.com/ageing/article-abstract/32/5/541/21734 by guest on 18 January 2019