Aim: To estimate an equivalent to the Modified Mini-Mental State Exam (3MSE), and to compare changes in the 3MSE with and without the estimated scores. Methods: Comparability study on a subset of 405 participants, aged ≥70 years, from the Cardiovascular Health Study (CHS), a longitudinal study in 4 United States communities. The 3MSE, the Telephone Interview for Cognitive Status (TICS) and the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) were administered within 30 days of one another. Regression models were developed to predict the 3MSE score from the TICS and/or IQCODE, and the predicted values were used to estimate missing 3MSE scores in longitudinal follow-up of 4,274 CHS participants. Results: The TICS explained 67% of the variability in 3MSE scores, with a correlation of 0.82 between predicted and observed scores. The IQCODE alone was not a good estimate of 3MSE score, but improved the model fit when added to the TICS model. Using estimated 3MSE scores classified more participants with low cognition, and rates of decline were greater than when only the observed 3MSE scores were considered. Conclusions: 3MSE scores can be reliably estimated from the TICS, with or without the IQCODE. Incorporating these estimates captured more cognitive decline in older adults.

Introduction

As populations in Western countries age, the incidence of dementia is increasing, with the annual number of incident cases of Alzheimer's disease (AD) in the United States expected to double in the next 50 years [1]. As a result, longitudinal assessment of cognitive decline has become an important component of epidemiologic studies. Yet, those participants who are at greatest risk of cognitive decline due to older age, comorbidities or prior low cognitive test scores are least likely to attend follow-up visits, leading to biased estimates of cognitive decline in the population [2–4].

1998, the Cardiovascular Health Study (CHS) performed a comparability study, obtaining scores on the TICS and IQCODE on a subset of participants within 30 days of administering the Modified Mini-Mental State Exam (3MSE) [7], the instrument used for clinic visits. The current report utilizes those data to derive equations to estimate the 3MSE from the other measures and to explore the impact of the estimated values on longitudinal estimates of changes in cognitive function.

**Methods**

The Cardiovascular Health Study

The CHS is a population-based study designed to identify risk factors for cardiovascular disease in individuals aged 65 years or more. In 1989, 5,201 participants were enrolled in 4 field centers in Sacramento County, Calif., Pittsburgh, Pa., Forsyth County, N.C. and Washington County, Md., USA. A supplemental cohort of 687 African-Americans was added in 1992–1993. Invited participants were a random sample of Health Care Financing Administration eligibility lists and persons living in their households. Prospective participants were eligible if they were aged 65 years or older, not wheelchair bound, able to provide informed consent, did not require a proxy respondent, were not under active treatment for cancer and planned to remain in the area for at least 3 years. Participants provided informed consent, and study methods were approved by the institutional review committees at each participating center. Details of the design and recruitment have been published [8, 9].

Participants were contacted every 6 months, alternating between clinic visits and phone calls through June 1999, followed by semi-annual phone calls, which continue through the present time. In 2005–2006, surviving participants were invited to a clinic or home visit as part of an ancillary study to CHS, known as All Stars. The primary goal of the All Stars study was to assess physical and cognitive function in this aging cohort.

Cognitive Assessments

The 3MSE was developed by Evelyn Teng in 1987 [7] to include a greater range of cognitive function and a more sensitive scale than the widely used Mini-Mental State Exam (MMSE) [10]. The 3MSE assesses orientation to place and time, short- and long-term memory, language and construction. It contains components that can only be done in person, including copying a figure and following written directions. Test-retest reliabilities greater than 0.90 have been reported and a score of <80 had a sensitivity and specificity of 0.91 and 0.97, respectively, for dementia [11].

The TICS is a brief telephone interview which was designed to identify cognitive impairment in AD patients [5]. It consists of 11 items with a maximum score of 41. Although modeled after the MMSE, only 2 items are identical to the MMSE. The version used in CHS had a maximum score of 40, because it did not include a question regarding the state in which the respondent lived.

The IQCODE is a 26-item questionnaire given to a proxy to assess changes in an older person's cognitive performance compared to 10 years prior [6]. Each item was rated on a 5-point scale from 1 (much better) to 5 (much worse), and the ratings were averaged over the 26 items to produce a total score with the same range of values, but on a continuous scale. A score of 3 indicated no change over time, and scores from 3 to 5 suggested increasing cognitive decline.

At the in-person visits, cognitive function was measured annually from 1990 to 1999 and in 2005–2006 using the 100-point 3MSE. Beginning in 1996, attempts were made to collect the TICS or IQCODE for participants who did not have an in-person visit.

Comparability Study

During follow-up in CHS, it became apparent that there was little or no cognitive decline among participants returning to the clinic for cognitive function tests. Investigators speculated that those undergoing decline were no longer coming to the clinic, and sought instruments for determining cognitive function among this pool of participants. The TICS and IQCODE were evaluated against the 3MSE in a subset of participants during the annual clinic visits in 1996–1998.

Participants were randomly selected to receive the 3MSE, the TICS and the IQCODE (by proxy) according to strata defined by education (< 10 or ≥ 10 years) and prior 3MSE scores (< 70, 70–78 or ≥ 78). Equal numbers were planned for each stratum for a total of 312 participants in the study. After administration of the 3MSE during the annual clinic visit, identified participants were informed of the comparability study and invited to participate. Those who agreed were called within the next 30 days for administration of the remaining 2 tests. When selected subjects refused to participate, an attempt was made to identify another subject in the same stratum to serve as a replacement. When that was not possible, additional participants outside the stratum were selected in order to reach the targeted sample size. In many cases, either the TICS or IQCODE (but not both) was obtained, and the clinics continued recruiting. The final subgroup of participants who had either the TICS or IQCODE within 30 days of the 3MSE totaled 405 participants, 364 with a TICS score and 331 with an IQCODE, representing 290 participants with both tests.

Statistical Methods

Linear regression was used to predict 3MSE from TICS and/or IQCODE. For each test, regression models were developed on a randomly selected study sample of 70% of those with the TICS score and 2/3 of those with the IQCODE, and tested on the remaining participants, forming the validation sample. The demographic measures of age, sex, race and education were considered as covariates. We used locally weighted scatterplot smoothers [12] and tested higher order terms and interactions in order to determine the optimal relationship of the TICS and/or IQCODE with the 3MSE. Once the form of the model was determined and validated, the full sample was used to estimate the coefficients for the final model [12]. R², a measure of the total variability explained by the model, was used to assess fit, and the Pearson correlation coefficient between the observed and predicted 3MSE was calculated to quantify their linear association.

Receiver operating characteristic (ROC) curves [12] and percent correctly predicted were used to explore the ability of the TICS score, the IQCODE and both combined to classify participants with or without low cognitive ability in accordance with a 3MSE score <80. Since higher scores of the IQCODE indicate
greater cognitive decline, the scores were subtracted from 5 to rescale for the ROC curve.

In order to assess the impact of missing 3MSE scores on estimates of change in cognitive function over time, we used random effects models [13] to estimate change, with and without estimating the missing 3MSE scores from the TICS and/or IQCODE. The statistical model may be represented by the following equation for the i-th subject at time j: score_{ij} = \beta_1 + \beta_2 \times t_j + b_{1i} + b_{2i} \times t_j + e_{ij}. The parameters of interest are \beta_1 and \beta_2, the intercept and slope of the average trajectory, with allowance for random deviations about the intercept and slope for individuals (b_{1i}, b_{2i}). We also explored the impact of the missing 3MSE scores on estimates of Apoe4, a known risk factor for decline in 3MSE [14], by adding terms for age, Apoe4 and Apoe4 \times time in the above model. Analyses were done with STATA, version 10.

**Results**

**Comparability Study**

The final sample included a total of 405 participants; 290 with all 3 cognitive function measures, 74 with only TICS and 3MSE, and 41 with only IQCODE and 3MSE. A greater percentage of more educated and cognitively able participants than the 52 originally planned for each stratum were enrolled (table 1). The mean age at the time of the study was 80.5 years, with a range of 70–102 years; 62% were women and 28% were African-American.

There was a quadratic relationship between the TICS and 3MSE scores in the study sample (p = 0.005 for the quadratic term). None of the demographic variables (age, sex, race or education) was associated with 3MSE once TICS was considered. TICS and its square alone accounted for 70% of the total variability in 3MSE scores in the study sample. Using the coefficients from this model to estimate 3MSE in the validation study produced an R² of 0.67. The Pearson correlation between the observed and predicted 3MSE in the full sample was 0.82. Results were similar in subgroups defined by sex or race. Results were less consistent among participants aged 80 years or older. The correlation between the observed and predicted 3MSE scores in this group was 0.71, compared with a correlation of 0.89 among participants younger than 80 years.

Among the 222 participants in the IQCODE study sample, there was 1 outlier and influential point which was dropped from the analysis. There was a quadratic relationship between 3MSE and IQCODE (p = 0.004 for the quadratic term). The model explained 48% of the variability in 3MSE scores, and when predicted 3MSE scores were computed, none were >88. The inability to predict higher scores with the quadratic model led us to adopt a linear model, which explained 46% of the variability, allowed predicted scores to exceed 88, and increased the percent correctly predicted above and below 80 in the validation sample. When applied to the validation set, R² was 0.42, indicating that less than half of the variability in 3MSE scores was explained by the IQCODE (fig. 3a). The mean 3MSE values were less well predicted in the validation sample than in the study sample, and sensitivity for classifying 3MSE scores <80 in the validation sample was 68%, lower than for the TICS (fig. 3b; table 3). The final model was 3MSE = 161.42–24.70 IQCODE, with an R² of 0.44. The predicted values were not able to capture the full range or variability of the observed 3MSE scores, and the area under the ROC curve was 0.75 (fig. 2), indicating only modest ability to classify participants without low cognition. An IQCODE score ≤3.30 corresponded to a 3MSE score ≥80. The Pearson correlation between the observed and predicted 3MSE in the full sample was 0.66.

### Table 1. Study participants by prior 3MSE scores and education level

<table>
<thead>
<tr>
<th>Education</th>
<th>Prior 3MSE score</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;70</td>
<td>70–78</td>
</tr>
<tr>
<td>≤10 years</td>
<td>50 (12.4)</td>
<td>45 (11.1)</td>
</tr>
<tr>
<td>&gt;10 years</td>
<td>51 (12.6)</td>
<td>44 (10.9)</td>
</tr>
<tr>
<td>Total</td>
<td>101 (24.9)</td>
<td>89 (22.0)</td>
</tr>
</tbody>
</table>

Data represented as the number of participants, with percentages in parentheses.
Although the IQCODE by itself explained less than half the variability in the 3MSE, it remained associated with 3MSE ($p < 0.001$) when added to the model containing TICS and TICS$^2$ based on all participants with both TICS and IQCODE in the 30-day window ($n = 290$). The model for predicting 3MSE from both scores was \(3MSE = 57.83 + 2.66 \text{TICS} - 0.0238 \text{TICS}^2 - 8.49 \text{IQCODE} \). This model explained 72.4\% of the total variability in 3MSE scores and correctly predicted 84\% of 167 scores $\geq 80$, and 79\% of 123 scores $< 80$. The correlation between the observed and predicted scores was 0.85 (fig. 4). The range of the predicted scores was 21–100. The mean difference, observed – predicted was 0.034 ± 9.7, with a range of −33 to 41. There was no improvement in area under the ROC curve over the model that included only TICS and TICS$^2$ (fig. 2).

**Fig. 1.** 3MSE and TICS score in validation sample. **a** Observed scores with fitted regression line. **b** Observed and predicted scores.
Longitudinal Analysis

The number of participants with available 3MSE scores in CHS declined over time due to death, but also due to illness or disabilities that prevented participants from coming into the clinic. Table 4 shows the number and percent of surviving participants with 3MSE scores for the final 5 years of annual clinic visits in the CHS, the last 4 of which correspond to the years when the TICS and IQCODE were administered, and for the All Stars in-person visit in 2005–2006. The table also shows the total numbers with any cognitive function test for each year, and the percent with a 3MSE or predicted 3MSE score <80. In 1998–1999, the final year of annual clinic visits, the estimation of 3MSE from TICS and/or IQCODE resulted in an additional 518 (16.5%) participants with cognitive function data when compared against the number with observed 3MSE scores, and, for the All Stars visit, the increase was 46% over the 1,135 with observed 3MSE scores. For each year from 1994–1999, the percent of participants estimated to score <80 on the 3MSE increased by adding each of the TICS and IQCODE, indicating that, as expected, those experiencing cognitive decline were

![Fig. 2. ROC curves for predicting 3MSE ≥80 from TICS score, IQCODE score or both.](image)

Table 2. 3MSE, observed and predicted by TICS

<table>
<thead>
<tr>
<th>Study sample (n = 261)</th>
<th>Validation sample(^1) (n = 103)</th>
<th>All(^2) (n = 364)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>observed</td>
<td>predicted</td>
</tr>
<tr>
<td>3MSE</td>
<td>78.5 ± 18.2</td>
<td>78.4 ± 15.2</td>
</tr>
<tr>
<td>Range</td>
<td>0–100</td>
<td>27–102</td>
</tr>
<tr>
<td>≥80(^3)</td>
<td>152</td>
<td>86%</td>
</tr>
<tr>
<td>&lt;80(^3)</td>
<td>109</td>
<td>85%</td>
</tr>
</tbody>
</table>

3MSE scores are presented as means ± SD.

\(^1\) Based on model developed in study sample.

\(^2\) Based on final model using all data.

\(^3\) Entry in observed column = number of participants above or below cut-point; in predicted column = % correctly predicted.
less likely to come to the clinic. In 2005–2006, participants had the option of an in-person home or clinic visit, and the proportion scoring <80 did not increase when the 3MSE was predicted from the TICS for those who chose a phone visit rather than an in-person visit. The proportion did increase when IQCODE was included.

In a longitudinal analysis of change between 1994 and 1999, we considered only participants with at least 2 cognitive function scores. Using the observed 3MSE, 4,118 CHS participants met that criterion; using TICS and/or IQCODE to estimate 3MSE increased the number to 4,462. Importantly, \(214 = 62\%\) of the additional participants were missing the last 1, 2 or 3 years of 3MSE scores. This is important because missing values straddled by observed values can be interpolated as the mean of the observed values, whereas missing values at the end of fol-

**Fig. 3.** 3MSE and IQCODE in validation sample. 

- **a** Observed scores with fitted regression line. 
- **b** Observed and predicted scores.
We estimated missing 3MSE values from the TICS only and then from TICS and/or IQCODE each year, using both scores when available. The observed average change per year in 3MSE scores was −0.68 (95% CI: −0.74 to −0.61), p < 0.001. When missing 3MSE scores were estimated from TICS scores, the average change per year was −0.78 (95% CI: −0.84 to −0.72), p < 0.001, and after estimating with both TICS and IQCODE, the average change per year was −0.84 (95% CI −0.91 to −0.78), p < 0.001. The regression lines with confidence bounds indicated at selected points are shown in figure 5. Estimates of the intercept were similar for each model, so there was substantial overlap of confidence intervals initially, followed by a separation of the lines over time. Differences in estimates of rates of decline by Apoe4 status, adjusted for age, are shown in table 5. For all of the models the effect of Apoe4 on slope was significant, p < 0.001, so the impact of missing 3MSE scores was not so severe as to miss an association with an important predictor. However, the estimated slopes were steeper in the models that included 3MSE values computed from TICS and/or IQCODE, and the difference in slopes between those with and without the Apoe4 allele were greater in the fully estimated model than in the observed model.

Since the IQCODE was designed to capture decline in cognitive function over time, we compared the yearly change in 3MSE scores with the IQCODE. Figure 6 shows a smoother scatterplot on the graph of individual 3MSE annual slopes by IQCODE. Note that the relationship between the individual 3MSE annual slopes and IQCODE is flat for values below 3, the point at which no change was reported, and decreases linearly for values >3. There remains a lot of variability around the line, due in part to the 5-year time frame for the measurement of change in 3MSE and the 10-year reporting period for the proxy. Several of the participants with no change in 3MSE during the last 5 years of the study already had low scores, and the proxy correctly reported a decline over a longer period of time. Even so, the linear correlation between the IQCODE and the annual rates of change in 3MSE scores was −0.694.

**Discussion**

Data from a comparability study in CHS support the utility of using alternate measures of assessing cognitive function in participants who are unable or unwilling to return for a clinic visit. The TICS correlated highly with 3MSE scores, and discriminated well between those who would be classified cognitively intact and those who would not, based on 3MSE scores. The IQCODE alone was not a good predictor of 3MSE, leaving more than half the variability unexplained. It did track well with decline in cognitive function over time, and adds information beyond that collected by the TICS to improve precision of the estimated 3MSE score.

Other researchers have compared the TICS to the 30-point MMSE. Brandt found a correlation of 0.94 between the 2 scales in a sample of AD patients, but the correlation when the higher scoring controls were included was not reported [5]. Jarvenpaa et al. [15] reported a correlation of 0.86 between the 2 measurements in a study of 30 AD patients and 26 healthy controls, and Ferrucci et al. [16] administered both the MMSE and TICS face to face and found high comparability. Carpenter et al. [17] reported

<table>
<thead>
<tr>
<th>Study sample (n = 221)</th>
<th>Validation sample1 (n = 110)</th>
<th>All2 (n = 331)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3MSE observed</td>
<td>76.0 ± 20.5</td>
<td>72.4 ± 24.4</td>
</tr>
<tr>
<td>3MSE predicted</td>
<td>76.0 ± 13.9</td>
<td>76.2 ± 14.2</td>
</tr>
<tr>
<td>Range</td>
<td>0–100</td>
<td>0–100</td>
</tr>
<tr>
<td>≥803</td>
<td>125</td>
<td>53</td>
</tr>
<tr>
<td>&lt;803</td>
<td>96</td>
<td>57</td>
</tr>
</tbody>
</table>

3MSE scores are presented as means ± SD.

1 Based on model developed in study sample.

2 Based on final model using all data.

3 Entry in observed column = number of participants above or below cut-point; in predicted column = % correctly predicted.

---

**Table 3. 3MSE, observed and predicted by IQCODE**
a correlation of 0.73 between the MMSE and TICS in 40 community-dwelling subjects. They found the percent of correct responses on the TICS to be generally lower than on the MMSE, primarily due to the larger memory component in the TICS, which includes a 10-item recall list, compared with a 3-item list in the MMSE, for which prompting is provided. A similar observation is reflected in our results, where a TICS score of 25/40 = 62.5% is equivalent to a MMSE score of 80%.

The IQCODE has also been compared against the MMSE. In 15 published studies, the correlation between the IQCODE and the MMSE ranged from –0.78 to –0.37 [18]. The IQCODE performed at least as well as the MMSE in screening for dementia [19], and correlated with change in MMSE over 7–8 years (r = –0.47) [20], which is less strong than the correlation of –0.694 we observed between the 3MSE and IQCODE.

It is not unexpected that the IQCODE is not a good substitute for the 3MSE at a single point in time. It was designed to assess change over a long period of time, and reflects the perceptions of an informant, and therefore depends on how well the informant knows the participant. Even with a spouse as proxy, Jorm et al. [21] showed that the results may be influenced by the relationship of the proxy to the participant.
As expected, including 3MSE values estimated from the TICS and/or IQCODE led to an estimate of greater decline in cognitive function over time. An analysis of data from the Cambridge City over 75 Cohort demonstrated that drop out at later stages in the study was associated with greater cognitive decline at earlier stages [3]. Similarly, in the Australian Longitudinal Study of Aging, information gathered from proxies revealed that those participants who did not return were at higher risk of cognitive decline and dementia [4].

Our method does not account for the uncertainty of the prediction, but neither does the commonly used method of last value carried forward, which has the disadvantage of not allowing for change at the end of the observation period, which is very likely, especially in older participants. Due to the correlation among measurements within a person, the random effects model implicitly imputes missing observations along the observed trajectory, yet our estimates of greater decline per year using the predicted 3MSE scores suggest that participants with missing 3MSE scores cannot be assumed to follow their prior score trajectory. That is, the probability of missing scores depends on unobserved quantities, so the data are not missing at random. Multiple imputation is another method for filling in missing data, but it too relies on the assumption that the data are missing at random. Brayne et al. [3] have shown that this assumption does not hold, and that the probability of being missing is related to the cognitive function score that would have been observed if the instrument had been administered. In the case of such informative drop-out, imputation methods must be accompanied by sensitivity analyses, which require additional assumptions and produce a range of possible results.

![Fig. 5. Regression of 3MSE on study year for observed values (solid line), adding TICS (dashed line) and then adding IQCODE (dotted line), with confidence bands at selected points.](image)

<table>
<thead>
<tr>
<th>Apoe4 allele</th>
<th>Observed 3MSE</th>
<th>+ TICS</th>
<th>+ IQCODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>–0.50 (–0.57 to –0.42)</td>
<td>–0.61 (–0.69 to –0.54)</td>
<td>–0.65 (–0.73 to –0.58)</td>
</tr>
<tr>
<td>Yes</td>
<td>–0.85 (–0.98 to –0.73)</td>
<td>–0.97 (–1.10 to –0.84)</td>
<td>–1.11 (–1.24 to –0.98)</td>
</tr>
</tbody>
</table>

95% CI are given in parentheses.
Our regression-based method was validated internally by splitting the study participants into 2 groups; 1 for model development and 1 for validation. We tried more complex models that improved $R^2$ in the study sample, but resulted in large reductions in $R^2$ in the validation sample, suggesting overfitting. An advantage of the simpler model is that no other data are needed beyond the test scores in order to predict the missing 3MSE score.

Our study is not without limitations. The originally planned sampling scheme was unfilled, largely due to a smaller pool for selection among those with low prior 3MSE scores and a lack of participation among members of this group. This is consistent with the motivation for the study; i.e. those who perform poorly are less likely to participate in follow-up visits. Study scores over time on the 3MSE in some participants vary greatly, perhaps due to the participant’s mood or interest in the exam on a particular visit. Variability within an individual not attributable to true decline reduces our ability to accurately predict scores. The 3MSE was administered up to 6 times before the comparability study. Participants may have learned the questions, leading to higher scores on the 3MSE, which would not carry over to the TICS. Finally, our version of the TICS did not include the question regarding in which state the participant lived, so other investigators wishing to use our function may need to re-score the TICS if their instrument included that question.

Conclusions

A comparability study done in CHS has shown that scores on the TICS can reliably be used to estimate scores on the 3MSE. The IQCODE is not a reliable estimate of a single 3MSE score, but improves the precision of the estimated 3MSE when used along with the TICS, and correlates well with change in 3MSE scores over time. The formula for converting TICS and/or IQCODE scores to 3MSE scores was internally validated and could be used by other researchers with similar study populations of individuals aged 70 years and older to more accurately assess decline in cognitive function over time.

Acknowledgment

The research reported in this article was supported by contract numbers N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133 and grant number U01 HL080295 from the National Heart, Lung, and Blood Institute, with additional contribution from the National Institute of Neurological Disorders and Stroke. Additional support was provided through R01 AG-15928, R01 AG-20098 and AG-027058 from the National Institute on Aging, R01 HL-075366 from the National Heart, Lung and Blood Institute, and the University of Pittsburgh Claude D. Pepper Older Americans Independence Center P30-AG-024827. A full list of principal investigators and institutions can be found at http://www.chs-nhlbi.org/pi.htm.
References