Post-hoc principal component analysis on a largely illiterate elderly population from North-west India to identify important elements of mini-mental state examination

Sunil Kumar Raina, Vishav Chander, Sujee Raina¹, Ashoo Grover²
Departments of Community Medicine and ¹Medicine, Dr. R. P. Government Medical College, Tanda, Himachal Pradesh, ²Division of Non-Communicable Diseases, ICMR, New Delhi, India

ABSTRACT

Background: Mini-mental state examination (MMSE) scale measures cognition using specific elements that can be isolated, defined, and subsequently measured. This study was conducted with the aim to analyze the factorial structure of MMSE in a largely illiterate, elderly population in India and to reduce the number of variables to a few meaningful and interpretable combinations. Methodology: Principal component analysis (PCA) was performed post-hoc on the data generated by a research project conducted to estimate the prevalence of dementia in four geographically defined habitations in Himachal Pradesh state of India. Results: Questions on orientation and registration account for high percentage of cumulative variance in comparison to other questions. Discussion: The PCA conducted on the data derived from a largely illiterate population reveals that the most important components to consider for the estimation of cognitive impairment in illiterate Indian population are temporal orientation, spatial orientation, and immediate memory.

Key words: Illiterate, mini-mental state examination, North-west India, post-hoc, principal component analysis

Introduction

The mini-mental state examination (MMSE) is a sensitive, valid, and reliable 30-point scale, extensively being used in clinical and research settings to measure cognitive impairment.[1,2] It is also being used to estimate the severity and progression of cognitive impairment and to follow the course of cognitive changes in an individual. The MMSE is not intended to provide a diagnosis for any particular nosological entity.[3] The scale (MMSE) examines the functions including registration, attention and calculation, recall, language, ability to follow simple commands, and orientation.[3,4]

However, questions in past have been raised on the sensitivity and specificity of MMSE. To improve sensitivity and specificity of MMSE scale, studies have suggested modifications to make it useful for estimating cognitive impairment across different population groups.[5‑10] Some studies have worked on changing the cut-off scores. Not much success has been achieved in this approach. It was observed that changing the cut-off affects the sensitivity and specificity of the test, increasing one while decreasing the other.[5‑7] Studies have also made efforts to compensate for various ages and education levels with some assuming that MMSE is valid only if the person has 9 or more years

Address for correspondence:
Dr. Sunil Kumar Raina, Department of Community Medicine, Dr. R. P. Government Medical College, Tanda, Himachal Pradesh, India.
E-mail: ojasrainasunil@yahoo.co.in

How to cite this article: Raina SK, Chander V, Raina S, Grover A. Post-hoc principal component analysis on a largely illiterate elderly population from North-west India to identify important elements of mini-mental state examination. J Neurosci Rural Pract 2016;7:44-7.
of schooling.\textsuperscript{[8–10]} However, problem stays the same, it affects the sensitivity and specificity of the test, increasing one while decreasing the other. Attempts have also been made to identify the importance of individual items in the scale. For this, authors have attempted multivariate procedures to differentially weigh the existing MMSE items.\textsuperscript{[8,11]}

The various functions measured by the scale are based on the assumption that cognition is made up of specific elements that could be isolated, defined, and subsequently measured. However, these individual elements are not entirely independent, but are all interrelated. For reasons related to culture and language, the factorial structure of the MMSE in dementia might be different among different populations. This study was conducted with the aim to analyze the factorial structure of MMSE in an elderly population in India and to reduce the number of variables to a few meaningful and interpretable combinations. Extensive PubMed search has revealed no investigations of the factorial structure of the MMSE in India.

Methodology

This study is a secondary data analysis of the data generated by a research project conducted to estimate the prevalence of dementia in four geographically defined habitations in Himachal Pradesh state in North-west India. In this project, 500 individuals above 60 years of age were included from each geographical site giving a target sample size of 2000. A total of 32 (1.6\%) out of a total of 2000 elderly individuals were classified as demented after clinical evaluation. The details are provided somewhere else.\textsuperscript{[12]}

For the purpose of this study, the 11 subtests composing the MMSE were considered independently [Table 1]. The MMSE is composed of 11 major items; temporal orientation (5 points), spatial orientation (5 points), immediate memory (3 points), attention/concentration (5 points), delayed recall (3 points), naming (2 points), verbal repetition (1 points), verbal comprehension (3 points), writing (1 points), reading a sentence (1 points), and constructional praxis (1 points).

The MMSE has a maximum score of 30, with five different domains of cognition analyzed: (1) Orientation; contributing a maximum of 10 points, (2) memory; contributing a maximum of 6 points, (3) attention and calculation, as a measure of working memory; contributing a maximum of 5 points, (4) language; contributing a maximum of 8 points, and (5) design copying; contributing a maximum of 1 point.

For the purpose of principal component analysis (PCA) on MMSE responses from all the 2000 individuals, above 60 years of age, the questions on MMSE were considered. Partial scores (sub‐totals) associated to each factor were computed for each subject and their distribution was examined as a function of the MMSE total score.

PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called as principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (i.e., accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to (i.e., uncorrelated with) the preceding components. The principal components are orthogonal because they are the eigenvectors of the covariance matrix, which is symmetric.

Results

The total variance among the different components of MMSE has been provided in Table 2. The results show that component 1–4 accounts for almost 66\% of the cumulative variance which is a large percentage. In other words, questions on orientation and registration account for high percentage of cumulative variance in comparison to other questions. A further look at Table 2 shows that there is a drop in the variance of 11.10 (29.325–18.226) from component 1 to component 2. The drop in variance from component 2–3 is 7.6 (18.226–10.569). However, the drop reduces when we move from component 3 onward to component 11. A look at Figure 1 explains results similar to that shown in Table 1. A sharp drop from one
Eigen value to the next at component 4 serves as another indicator of how many components are important and which components to extract. The extraction of factors on PCA reveals that the most important components to consider are component 1–3. The remaining factors explain a very small proportion of the variability and are likely unimportant in our setup.

The remaining factors explain a very small proportion of the variability and are likely unimportant. The least variance is shown by component 11 (constructional praxis) probably reflecting on its less importance in the assessment of cognitive impairment in our setup.

**Discussion**

The current study was conducted with the aim to identify the important components of MMSE which may be useful for identifying cognitive impairment in our setup. With a large number of variables, the dispersion matrix may be too large to study and interpret properly. With a predominantly illiterate, elderly population, extraction of factors with less variance may reduce the total number of variables used to identify the cognitive impairment. This is because MMSE and its modifications use learnt language as a medium of assessment.

This becomes important, keeping in view, the fact that majority of elderly Indians are illiterate and it becomes difficult to navigate through the various components of MMSE. Cullum et al. in their study using a stepwise regression analysis with highly normal subjects aged 50–80, reported that recall of three words and orientation to time correlated 87 with total score.[13] Galasko et al. employing a logistic regression model with Alzheimer’s disease patients found that the sum of the scores for recall of three words and orientation to place resulted in sensitivity and specificity levels that were similar to those produced by the total MMSE scores.[5] Magaziner et al. generated a series of prediction equations for different age groups and educational levels.[14] However, studies from other authors question the generalization of these equations.[15,16] The current study used PCA to arrive at relatively more important components of MMSE. The idea was to reduce the number of components of MMSE to make it user friendly. The PCA conducted on the data derived from a largely, illiterate population reveals that the most important components to consider for the estimation of cognitive impairment in illiterate Indian population are component 1–3. In other words, the three most important components are; temporal orientation, spatial orientation, and immediate memory. The remaining factors explain a very small proportion of the variability and are more likely unimportant in Indian setting. The argument is further strengthened by the fact that the least variance is shown by component 11 (constructional praxis) probably reflecting on its less importance in the assessment of cognitive impairment in our setup. Studies in past have reported on the

---

**Table 2: Explaining the total variance among the different components of MMSE**

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial eigen values</th>
<th>Extraction sums of squared loadings</th>
<th>Rotation sums of squared loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Percentage of variance</td>
<td>Cumulative (%)</td>
</tr>
<tr>
<td>2</td>
<td>2.005</td>
<td>18.226</td>
<td>47.550</td>
</tr>
<tr>
<td>3</td>
<td>1.163</td>
<td>10.569</td>
<td>58.119</td>
</tr>
<tr>
<td>4</td>
<td>0.866</td>
<td>7.870</td>
<td>65.989</td>
</tr>
<tr>
<td>5</td>
<td>0.748</td>
<td>6.800</td>
<td>72.789</td>
</tr>
<tr>
<td>6</td>
<td>0.688</td>
<td>6.254</td>
<td>79.043</td>
</tr>
<tr>
<td>7</td>
<td>0.674</td>
<td>6.124</td>
<td>85.167</td>
</tr>
<tr>
<td>8</td>
<td>0.542</td>
<td>4.930</td>
<td>90.097</td>
</tr>
<tr>
<td>9</td>
<td>0.435</td>
<td>3.954</td>
<td>94.052</td>
</tr>
<tr>
<td>10</td>
<td>0.362</td>
<td>3.289</td>
<td>97.341</td>
</tr>
<tr>
<td>11</td>
<td>0.292</td>
<td>2.659</td>
<td>100.000</td>
</tr>
</tbody>
</table>

Extraction method: Principal component analysis. MMSE: Mini-mental state examination

**Figure 1: Scree plot showing individual components**
performance of study participants on constructional praxis.

A study from Sydney, reported that one-third of the community-dwelling people without dementia (CDR-0) performed poorly on the test.\[17\] This could be as Katzman\[18\] reported that if one does not learn to copy as a child he or she will not be able to complete the task when required at a later point in life. The same appears to be true for questions on recall, writing, and reading a sentence.

**Financial support and sponsorship**

Nil.

**Conflicts of interest**

There are no conflicts of interest.

**References**