Reproducibility of qualitative assessments of temporal lobe atrophy in MRI studies

S. Sarria-Estrada*, C. Acevedo, R. Mitjana, L. Frascheri, S. Siurana, C. Auger, A. Rovira

Unidad Ressonància Magnètica (I.D.I.), Servei de Radiologia, Hospital Universitari Vall d’Hebron, Barcelona, Spain

Received 1 December 2013; accepted 23 April 2014

Abstract

Objective: To determine the reproducibility of the Scheltens visual rating scale in establishing atrophy of the medial temporal lobe.

Material and methods: We used coronal T1-weighted inversion recovery sequences on a 1.5 Tesla MRI scanner to study 25 patients with clinically diagnosed Alzheimer's disease or mild cognitive decline and 25 subjects without cognitive decline. Five neuroradiologists trained to apply the Scheltens visual rating scale analyzed the images. We used the interclass correlation coefficient to evaluate interrater and intrarater agreement.

Results: Raters scored 20 (80%) of the 25 patients with mild cognitive decline or Alzheimer’s disease between 2 and 4; by contrast, they scored 21 (84%) of the 25 subjects without cognitive decline between 0 and 1. The interrater agreement was consistently greater than 0.82, with a 95% confidence interval of (0.70–0.9). The intrarater agreement ranged from 0.82 to 0.87, with a 95% confidence interval of (0.56–0.93).

Conclusion: The Scheltens visual rating scale is reproducible among observers, and this finding supports its use in clinical practice.

© 2013 SERAM. Published by Elsevier España, S.L.U. All rights reserved.

Reproducción del valoración cualitativa de la atrofia del lóbulo temporal por RM

Resumen

Objetivo: Determinar la reproducibilidad de la escala visual de Scheltens para establecer la atrofia del lóbulo temporal medial.

Material y métodos: Reunimos 25 pacientes con diagnóstico clínico de enfermedad de Alzheimer leve o deterioro cognitivo leve (DCL) y 25 sujetos sin deterioro cognitivo. Todos fueron estudiados con RM 1,5 Tesla utilizando secuencias de inversión recuperación ponderadas en T1 en el plano coronal. Cinco neurorradiólogos fueron entrenados para aplicar la escala de


* Corresponding author.
E-mail address: ssarria@idi-cat.org (S. Sarria-Estrada).

2173-5107 © 2013 SERAM. Published by Elsevier España, S.L.U. All rights reserved.
Scheltens y analizaron las imágenes. Se utilizó el coeficiente de correlación intraclase para valorar el grado de acuerdo inter e intraobservadores.

Resultados: El 80% de los pacientes con deterioro cognitivo leve o enfermedad de Alzheimer obtuvieron puntuaciones entre 2 a 4, mientras que 21 de los 25 controles sanos (84%) fueron puntuados entre 0–1. La concordancia entre observadores fue consistentemente mayor de 0,82, con un intervalo de confianza del 95% (0,7–0,9). La concordancia intraobservador varió entre 0,82 y 0,87, con un intervalo de confianza del 95% (0,56–0,93).

Conclusión: La clasificación de Scheltens es un método reproducible entre observadores, lo que apoya su uso en la práctica clínica.

© 2013 SERAM. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Introduction

The aging of population has increased significantly the prevalence of neurodegenerative conditions, especially Alzheimer’s disease (AD). Mild cognitive impairment (MCI) is a clinical syndrome that allows us to categorize people with cognitive impairment who do not meet the criteria for dementia. Even though it is well known that most people in this group advance to AD in about 9.5 year the percentage of these patients is not easy to predict and estimate. The rate of advancement of AD depends on many variables including demographic, genetic and the presence of other conditions in the moment of diagnosis. Establishing early the diagnosis of AD and distinguishing this condition from other causes of cognitive impairment is relevant to be able to establish prognosis and administer possible therapies. Thus there is a great scientific interest in trying to identify the biomarkers capable of predicting the risk of patients to advance from MCI to AD.

The structural studies of MR in patients diagnosed with AD or MCI usually show atrophy of the entorhinal cortex and hippocampus. Establishing the degree of atrophy of these temporal structures is relevant to be able to predict the risk of conversion to AD but it is also most important to assess the volumetric changes produced in relatively short periods of time. Scheltens et al. developed a system of visual categorization to assess the degree of atrophy of the medial temporal lobe (MTL) which applied cross-sectionally and longitudinally in an adequate clinical setting will allow us to identify all those individuals with a high risk of advancing to AD. The system is reproducible, does not require image processing software and is easy to learn.

In this study we try to validate the qualitative analysis of the atrophy of the MTL as a reproducible method among neuroradiologists previously trained to apply Scheltens’ visual rating scale. The hypothesis claims that after proper training the reproducibility among different evaluators is good which will allow us to apply the scale to other centers.

Material and methods

Brain MRIs were performed in 25 patients diagnosed with mild AD or MCI and in 25 controls without cognitive impairment (13 men and 12 women in each group). Of the 50 subjects studied, 13 (26%) had not gone to school, 18 (36%) had gone to primary school and 19 (38%) had gone to secondary school or higher studies. The hospital ethics committee approved carrying out the study. Since this study was carried out using MRIs performed within the clinical practice and the risk of damage was not greater than usual a waiver from informed consent was requested to the authors of the study—which was later approved by the PR (IDI) ethics committee 130/2013.

Patients were considered MCI or AD based on the NINCDS-ADRDA criteria and classified both neuropsychologic and neurologically based on how serious dementia was according to the Blessed Dementia Rating Scale, the Mini-Mental State Examination (MMSE) and the Global Deterioration Scale-Functional Assessment Staging. The MRIs were taken no more than 2 months after the clinical diagnosis as part of the protocol for the diagnosis of AD (Table 1). The control group consisted of patients without symptoms of cognitive impairment confirmed in the clinical history.

After reviewing the clinical history patients with evidence of another neurologic conditions different than AD were excluded including traumatic, demyelinating, neoplastic and cerebral-vascular causes; toxic and metabolic clinical presentations too. All patients were studied using one Symphony 1.5 Tesla MRI-equipment (Siemens, Erlangen, Germany) through a standardized protocol including T1-weighted inversion-recovery sequences (with coronal cuts perpendicular to the longitudinal axe of the temporal lobe), with a cut thickness of 4 mm, matrix of 230 × 100, TR = 7.000, TE = 72, TI = 350 and one plane resolution of 0.8 × 1 mm. Aside from this sequence for analytical purposes 5 cuts spreading from the prelaminar region to the IV ventricle floor were taken into consideration. The cuts obtained in each and every patient were analyzed through Scheltens et al.’s visual rating scale that categorizes the atrophy of the temporal lobe internal structures. This scale uses a scoring from 0 to 4 (Fig. 1) based on the size of the hippocampal gyrus, the amplitude of the choroidal fissure and the lateral ventricle temporal horn. This scale was applied independently in each temporal lobe. The scale was applied independently by 5 neuro-radiologists without information on the age and clinical data of patients (SSE, CA, LF, SSM, RW, with 6–10 years of experience in neuroradiology). To be able to stratify the MTL atrophy common examples of each and every degree of the scale were used. Prior to the analysis the 5 observers were trained in Scheltens’ scale. Once they were familiar with it they immediately
started applying it on the study subjects. One week after the analysis, 2 of the observers repeated the assessment of patients without information on their prior scores or the scores assigned by the neuro-radiologists to study the inter-observer concordance.

The quantitative variables were compared to the variance analysis test and qualitative variables to the $\chi^2$ test. To measure the degree of concordance in the scoring of the inter and intra-observer MLT degree of atrophy the interclass correlation coefficient (ICC) was applied using the SPSS 17.0 software for Windows. The degree of concordance with the ICC was classified as very good (>0.9); good (0.9–0.8); moderate (0.8–0.65) and poor (<0.65).

Results

No significant differences were found in age, gender degree of schoolarization between the two groups studies but in the MMSE score (Table 1). Eighty per cent of patients with MCI or AD scored between 2 and 4 in Scheltens’ scale while 21 of the 25 controls (84%) scored between 0 and 1 (Fig. 1). Both inter and intra-observer concordance was found in the range of 0.80–0.90 suggestive of the good degree of agreement for the test. The inter-observer assessment varied between 0.84 and 0.86 for the right and left temporal lobe, respectively (95% CI, 0.7–0.9). In the intra-observer assessment the ICC ranged between 0.82 and 0.87 (95% CI 0.56–0.93).

Discussion

The atrophy of the temporal lobe medial structures during the course of AD has been widely documented both in anatomopathologic and neuro-radiologic studies. Our study confirmed that the Scheltens’ scale designed for the subjective assessment of the MTL atrophy is reproducible which allows elevated degrees of inter and intra-observer concordance after learning how to use it correctly.

Even though not all patients with MTL atrophy advanced to dementia, both the MRIs obtained in the index and subsequent evaluations can be an important biomarker in the diagnostic process of the disease. Different studies have proven that measuring the volume of the temporal lobe medial region (manually or automatically) can help us distinguish patients with mild AD from healthy controls with

<table>
<thead>
<tr>
<th>Score</th>
<th>Amplitude of the choroidal fissure</th>
<th>Amplitude of the temporal horn</th>
<th>Height of hippocampus</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>1</td>
<td>Slightly widened</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>2</td>
<td>Moderately widened</td>
<td>Slightly widened</td>
<td>Slightly reduced</td>
</tr>
<tr>
<td>3</td>
<td>Significantly widened</td>
<td>Moderately widened</td>
<td>Moderately reduced</td>
</tr>
<tr>
<td>4</td>
<td>Significantly widened</td>
<td>Significantly widened</td>
<td>Significantly reduced</td>
</tr>
</tbody>
</table>


![Figure 1](image_url) Scale of degree of atrophy of the medial temporal lobe (Scheltens et al.). The visual assessment of the temporal lobe atrophy was performed in 5 coronal cuts obtained parallel to the floor of the IV ventricle through T1-weighted inversion-recovery sequences. The cuts obtained in our study were analyzed with the Scheltens' scale including scores from 0 to 4.
a 95% specificity and a 85% sensitivity. However these kinds of measures require segmentation programs that are time-consuming which makes it hard to be able to implement them in the clinical practice. In our study we used the scale proposed by Scheltens et al.\(^\text{9}\) to visually establish the degree of MTL atrophy in patients with AD and MCI and in a subset of patients without cognitive impairment and we confirmed that it can be reproducible. The method used is simple and can be easily transferred to the standard clinical practice.

This study has been designed to evaluate the degree of concordance among neuro-radiologists but not the capacity of discrimination among patients with and without cognitive impairment. As it would be expected the higher scores belonged to patients with MCI or mild AD that in turn supports the use of MRIs as a biomarker to predict the risk of conversion to AD.\(^\text{3}\)

The main limitations of this study include the inclusion of a small sample of patients with AD and MCI and also the fact that the cross-sectional nature of the study does not let us assess the clinical/prognostic value of the Scheltens’ scale. On the other hand all studies were performed by experienced neuro-radiologists from a single institution so we would still have to determine if the scale is also reproducible among radiologists of various institutions without exclusive dedication to neuro-radiology.

As a conclusion the visual analysis of the degree of atrophy of the medial structures of the temporal lobe through Scheltens’ scale is a reproducible easy method to implement. The good degree of inter and intra-observer concordance obtained after an adequate learning supports its use as an analytical method of the degree of atrophy of the temporal lobe in clinical practice.

Ethical responsibilities

Protection of humans and animals. Authors confirm that all proceedings and experiments followed relate to the committee of responsible human experimentation ethical rules and regulations in compliance with the World Medical Association and the Declaration of Helsinki.

Confidentiality of data. Authors confirm that they have followed the protocols of their working centers about the publication of data from patients.

Right to privacy and informed consent. Authors confirm that in this report there are no personal data of patients.

Author

1. Manager of the integrity of the study: SSE, CA, AR.
2. Original Idea of the Study: AR, SSE, CA.
3. Study Design: CA, SSE, AR.
4. Data Mining: CA, SSE, RM, LF, SS, CA.
5. Data Analysis and Interpretation: CA, SSE, RM, LF, SS, CA.
7. Reference Search: CA, SSE, AR.
8. Critical review of the manuscript with intellectually significant remarks: SSE, CA, RM, LF, SS, CA, AR.
9. Approval of final version: SSE, CA, RM, LF, SS, CA, AR.

Conflict of interests

Authors declare no conflict of interests.

References