Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 2 (2016) 12-18

Alzheimer's & Dementia

Cognitive & Behavioral Assessment

Using quantile regression to create baseline norms for neuropsychological tests

Ben Sherwood[a], Andrew Xiao-Hua Zhou[b,*], Sandra Weintraub[c], Lan Wang[d]

[a]Department of Biostatistics, Johns Hopkins University, Baltimore, MD, USA
[b]Department of Biostatistics, University of Washington, Seattle, WA, USA
[c]Department of Psychiatry and Behavioral Sciences and Neurology – Ken and Ruth Davee Department, Northwestern University, Chicago, IL, USA
[d]School of Statistics, University of Minnesota, Minneapolis, MN, USA

Abstract

Introduction: The Uniform Data Set (UDS) contains neuropsychological test scores and demographic information for participants at Alzheimer's disease centers across the United States funded by the National Institute on Aging. Mean regression analysis of neuropsychological tests has been proposed to detect cognitive decline, but the approach requires stringent assumptions.

Methods: We propose using quantile regression to directly model conditional percentiles of neuropsychological test scores. An online application allows users to easily implement the proposed method.

Results: Scores from 13 different neuropsychological tests were analyzed for 5413 cognitively normal participants in the UDS. Quantile and mean regression models were fit using age, gender, and years of education. Differences between the mean and quantile regression estimates were found on the individual measures.

Discussion: Quantile regression provides more robust estimates of baseline percentiles for cognitively normal adults. This can then serve as standards against which to detect individual cognitive decline.

Keywords: Alzheimer's disease; Neuropsychological assessment; Cognitive decline; Early detection; Quantile regression

1. Introduction

The Uniform Data Set (UDS) is a central database containing information on participants at contributing Alzheimer's disease centers and Alzheimer's disease research centers funded by the National Institute on Aging. The participants in UDS are self-selecting and may not be representative of the entire population. Using the UDS, Weintraub et al. [1] provided an exploratory analysis of the relationship between neuropsychological test scores and age, years of education and gender. They found that, in almost all cases, age, education, and gender should be taken into account when modeling the expected score of a test. Shirk et al. [2] proposed using these models to estimate percentiles for a given patient by calculating a z-score conditional on the patient's individual characteristics. The z-score was then referenced to the standard normal distribution to estimate the percentile for the subject's performance on a given test.

Natural cognitive decline due to age can be accounted for by adopting a statistical model that incorporates age. This results in an estimate of the percentiles that are dependent on the age of the patient. The statistical models can be made more precise by also including years of education and gender. Ideally, participants can take the test while healthy and create a baseline percentile for the subject in comparison with the peer group that share the same age, education, and

B.S., S.W., and L.W. report no disclosures. A.X.Z. is presently a core investigator and biostatistics unit director at HSR&D Center of Excellence, Department of Veterans Affairs Puget Sound Health Care System, Seattle, WA, USA. This article does not necessarily represent the findings and conclusions of VA HSR&D.

*Corresponding author. Tel.: +1-206-616-2719; Fax: +1-206-543-3286.

E-mail address: azhou@u.washington.edu

http://dx.doi.org/10.1016/j.dadm.2015.11.005

gender. Any future substantial decline from the baseline percentiles would provide a warning sign to take more sophisticated tests to detect cognitive decline. For example, suppose we have a patient who initially scored at the 95th percentile for a given test, meaning 95% of the cognitively healthy population with the same age, gender, and years of education performed worse than the patients on the test. If after 1 or 2 years when taking the test again, their test score dropped to the 80th percentile as calibrated by the same statistical model, this would flag the subject for further medical attention. Even though the patient is still performing relatively well, we would expect them to stay around the same high percentile because the statistical model accounts for decline due to age.

In the literature, mean regression analysis is often used to model the percentiles, which relies on strong modeling assumptions. The popular z-score approach by Shirk et al. [2] adopts mean regression and assumes that the random errors are normally distributed and the variance is constant. The model assumes that the variance of test scores is the same for different genders, years of education, and age. These assumptions are also crucial for the mean regression method to calculate confidence intervals and *P* values for coefficient estimates. For sufficiently large samples sizes, the confidence intervals and *P* values are approximately valid even if the normality assumption is violated. If the constant variance assumption is not satisfied, then the *P* values and confidence intervals will not be correct, although the estimates of the coefficients are not systematically biased. When modeling percentiles, deviations from these underlying assumptions become more problematic. A deviation from either assumption could result in biased percentile estimates that cannot be corrected even if the data set is huge.

We propose a new approach to create baseline norms for neuropsychological tests using quantile regression and provide a web-based application to implement the procedure. Quantile regression was proposed in a seminal article of Koenker and Bassett [3] and aims to directly model the conditional percentiles. We refer to Koenker [4] for a thorough review of quantile regression. In the present article, we applied both quantile regression and mean regression to analyze a cognitively normal subset of the UDS neuropsychological data to create the baseline norms and compared the results.

2. Methods

2.1. Subject selection

To create baseline percentiles for patients without signs of cognitive decline, analysis was limited to cognitively normal adults with complete information on age, gender, and years of education. Specifically, we limit the analysis to participants in the UDS that met the following criteria: (1) a global clinical dementia rating score of 0; (2) a total functional assessment questionnaire score of 0; (3) a normal score from the neuropsychiatric inventory questionnaire (NPI-Q); (4) no missing data for age, gender, or education; and (5) visit dates from January 2005 to February 2013. In addition, we only use data from a subject's first visit. These criteria are similar to those used by Weintraub et al. [1], with the difference that analysis was limited to participants with first visits from September 2005 to August 2007. In our study, there were 5413 participants who met the aforementioned requirements.

Of the 5413 participants, 78% were white, 16% black, 3% multiracial, 2% Asian, and 1% other. Hispanic was not included as a category for race. The age range was 13% younger than 60, 30% between 60 and 69, 37% between 70 and 79, 18% between 80 and 89, and 2% older than 90 at the time of first visit. Females were over-represented (69% of the participants). The breakdown by total years of education was 18% with <12 years, 20% between 13 and 15 years, 24% with 16 years, and 38% with ≥ 17 years of education.

2.2. Neuropsychological tests

We analyzed results from the following UDS neuropsychological tests: mini-mental state examination (MMSE) [5], Wechsler memory scale-revised subtests logical memory A (first story, immediate and delayed recall) [6], digit span forward and backward [6], semantic fluency (animals and vegetables) [7], Boston naming test (BNT; 30 item—odd numbered) [8], Wechsler adult intelligence scale-revised (WAIS-R) digit symbol coding subtest [9], and trail making test (TMT) parts A and B [10]. A review of these tests and how they have been used for the UDS was provided in Weintraub et al. [1]. For delayed recall of logical memory A, participants were asked to remember parts of a story after a delay. In version 1.1 of the UDS, the delay was 30 minutes and for version 1.2 and 2.0, the delay was 20 minutes, but not all participants were delayed at the recommended time. The precise delay time had been recorded for each subject, and this information was used when modeling the performance for the logical memory A delayed recall test [1].

2.3. Data analysis

Both quantile and mean regression models were fit for the aforementioned neuropsychological tests. For the predictors, we considered age, years of education, and gender as single variable models, and also a full model that includes all three variables. To test the assumptions of normality and homoscedasticity, we used a Kolmogorov-Smirnov test for normality [11] and a Breusch-Pagan [12] test for homoscedasticity. Quantile regression does not rely on the assumptions of normality or homoscedastic errors to model the conditional percentiles. All analyses were carried out using R 3.1.1. The quantreg package [13] was used for quantile regression, and the lmtest package [14] was used for the Breusch-Pagan test; otherwise, calculations were implemented using the base software.