Multinomial regressions to identify phenotypes in the Obsessive-Compulsive Disorder

Guaraci L. Requena*
Institute of Mathematics and Statistics, University of Sao Paulo (IME-USP), Sao Paulo, Brazil - guaraci@ime.usp.br

Roseli G. Shavitt
Department and Institute of Psychiatry, University of Sao Paulo (IPq - HCFMUSP), Sao Paulo, Brazil

Peggy M.A. Richter
Sunnybrook Health Sciences Centre, Department of Psychiatry, University of Toronto, Toronto, Canada

Gwyneth Zai
Centre for Addiction and Mental Health, Department of Psychiatry, University of Toronto, Toronto, Canada

Carlos A. B. Pereira
Institute of Mathematics and Statistics, University of Sao Paulo (IME-USP), Sao Paulo, Brazil

Obsessive-Compulsive Disorder (OCD) is a psychiatric disorder characterized by intrusive thoughts (obsessions) and repetitive behaviors (compulsions), which affects about 2% of the population worldwide. Research in this area is under continuous development, with many interesting results being found, including genetic findings. In order to measure OCD severity, there are many interviews. The most used in the field is the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) and, most recently, the Dimensional Yale-Brown Obsessive-Compulsive Scale (DY-BOCS) has also been used by some authors. The DY-BOCS measures OCD severity discriminating five different and homogeneous symptom dimensions, plus one heterogeneous dimension. These measures reveal the OCD phenotype, which is of interest for genetic studies. On the other hand, even the Y-BOCS being the most used instrument to study OCD, it does not offer this phenotypical distinction. Our main objectives are 1. To build a technique to compare these two instruments in order to rewrite the DY-BOCS data in the same format of an existing Y-BOCS data set; and 2. To propose a multinomial logistic regression to model and to predict the phenotype (i.e., the most severe symptom dimension among the five homogeneous dimensions of the DY-BOCS) in a subject who has only the Y-BOCS data. This is an original study that aims to develop a statistical procedure to extract dimensional severity ratings in OCD from the existing Y-BOCS severity scores for use in genetic and other neurobiological research.

Keywords: OCD; DY-BOCS; Y-BOCS; multinomial logistic regression.