Using Multivariate Analysis for Comparing Locomotor Patterns of Sprague-Dawley Rats from Different European Breeders

Theresa Andreasson
theresa.andreasson@neurosearch.se

Jenny Börjesson
jenny.borjesson@neurosearch.se

Elisabeth Ljung
elisabethl.jung@neurosearch.se

Katarina Rydén Markinhuhta
katarina.ryden@neurosearch.se

Marianne Thorngren
marianne.thorngren@neurosearch.se

NeuroSearch Sweden AB
Arvid Wallgrens Backe 20, 413 46 Göteborg, Sweden

ABSTRACT

Our aim was to characterize the locomotor activity pattern of Sprague-Dawley (SD) rats from European suppliers in both normal (n=44-91 per supplier) and hypoglutamatergic state (n=17-28 per supplier). Spontaneous locomotor pattern of male SD rats weighing 220-250g from Taconic, BK Scanbur, Harlan and Charles River Laboratories were tested in motility meter boxes. Both vehicle (NaCl) and MK-801 (0.7mg/kg) treated animals were tested for one hour. A locomotor pattern was generated by calculating the time series of x, y (horizontal activity) and z (vertical activity) coordinates into eleven main variables sampled at 25 Hz. Each main variable was studied at seven frequencies from 25 Hz to 0.25 Hz and pooled into 15 min bins, generating a locomotor pattern matrix of each animal consisting of 308 variables.

Multivariate Analysis (MVA) by means of Principle Component Analysis (PCA)[1, 2] was used for locomotor characterisation. PCA is a method that preserves as much (linear) information as possible in the data while projecting the observations onto a lower dimensionality (usually 2-3 PC's), reducing noise and making it possible to both investigate clusters among the studied objects as well as correlations among the variables. The first PC will always contain the largest amount of variation (i.e. information) among the PC's and all subsequently derived PC's will describe less of the total variation in data. Higher PC's with small eigenvalues (i.e. little variance) that are unstable (insignificant according to cross validation [3, 4]) are often regarded as unstructured noise. All derived PC's are orthogonal to each other (i.e. only containing variance not taken into account of by previously derived PC's), which also implies that they are linearly independent and uncorrelated.

As is the case with all statistic methods, MVA works better on certain types of data, for example normally distributed data is preferred as well as a decent number of observations especially if we want to look at group differences and have many variables. However, PCA is relatively distribution independent and can be performed on both short and wide as well as long and thin data matrices and can in most cases handle missing data in an acceptable way. PCA is also unbiased in the sense that the model doesn't take group belongings into account in the calculation but reveal those directions in data that contains the largest variation. The PC's are often interpreted as latent structures that reveal the "true" nature of the data in form of variances, correlations, similarities and dissimilarities even though they in a strict theoretical way are not statistically independent unless all variables follow a Gaussian distribution (normal distribution).

PCA is most commonly interpreted with mainly two closely related plots, the score-plot and the loading-plot. The score plot shows the projection of objects onto the principal components. Objects located close to each other are similar with respect to our measured data. Objects located on the opposite side of origo have opposite characteristics and if they are located in a perpendicular direction to each other (origo is the reference) have independent characteristics. Thus, a score plot is used to find clusters or patterns in the distribution of the objects in our model. Loadings are derived as the eigenvectors of the correlation matrix sorted by the eigenvalues. Loadings reveal the inherent correlation structure present in our data. Variables located close to each other have a positive correlation and variables located on the opposite direction of origo (approximately fitting a straight line) have a negative correlation. If lines from two different variables that cross each other in origo have an

approximate perpendicular angle they are uncorrelated and linearly independent. The scores and loading plots are tightly connected and can be approximately interpreted by overlaying the graphs.

The experiments was approved by the Ethics Committee for Animal Experiments in Gothenburg (Göteborgs djurförsöksetiska nämnd)

Author Keywords
Phenotype, locomotor activity, multivariate analysis, rat, hypoglutamatergic, MK-801, PCA.