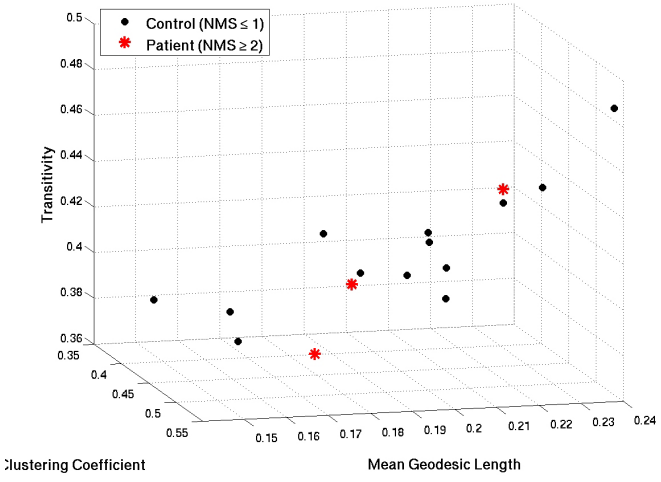
Predicting neurological outcome in neonatal encephalopathy: A machine learning and network analysis approach

Etay Ziv[1], Olga Tymofiyeva[1], Sonia L Bonifacio[2], Patrick S McQuillen[2], Donna M Ferriero[2], A James Barkovich[1], Duan Xu[1], and Christopher P Hess[1]
[1]*Department of Radiology & Biomedical Imaging, UCSF, San Francisco, CA, United States,* [2]*Department of Pediatrics, UCSF, San Francisco, CA, United States*

Introduction: Neonatal encephalopathy represents a heterogeneous group of conditions associated with life-long developmental disabilities. The ability to predict outcome early on in the perinatal period could potentially have a significant impact on subsequent treatment. Structural connectivity networks of the brain can be constructed using diffusion MRI [1] and various network properties have been measured on such networks derived from developing pediatric brains [2]. We hypothesize that networks derived from patients who have poor outcome may have different structure than those who have good outcome. Here we present an unbiased approach to enumerate a large set of network properties and using a combination of unsupervised and supervised learning, we demonstrate surprisingly good discrimination between good and poor neurological outcome.

Methods: Diffusion MRI was performed on 16 term-born babies at the age of six months who had encephalopathy at birth. The babies were scanned on a 3T GE EXCITE MR scanner using SE EPI with a FOV=24 cm, 128x128 matrix, min TE, 30 directions, b-value=700 s/mm^2. Construction of structural networks was performed including preprocessing [3], tractography [4], parcellation via equal area sphere partitioning [5], and finally assembly of the connectivity matrix. We next constructed a mapping of each network to a high-dimensional space of network properties, including traditional measures (clustering coefficient, characteristic path length, and transitivity) as well as all of the subgraphs with up to 7 edges (for a total of 130 features) [6]. Principle component analysis (PCA) was performed on the feature space, which was then used as the input space to a supervised learning algorithm. The neuromotor score (NMS) [7] assessed for each baby at the age of 6 months by pediatric neurologists was then binarized (NMS>1 was considered abnormal) and used as the class label and a support vector machine (SVM) [8] was trained and tested using leave-one-out cross-validation (LOOCV).

Results: Fig. 1 shows 3 typical network properties (clustering coefficient, characteristic path length, and transitivity) for the two classes with complete overlap. Fig. 2 shows the two classes mapped to the first three principle components of the subgraph feature space with notable separability except for one outlier. An SVM with linear kernel trained on this space with LOOCV results in 87.5% testing prediction accuracy of poor neurological outcome.



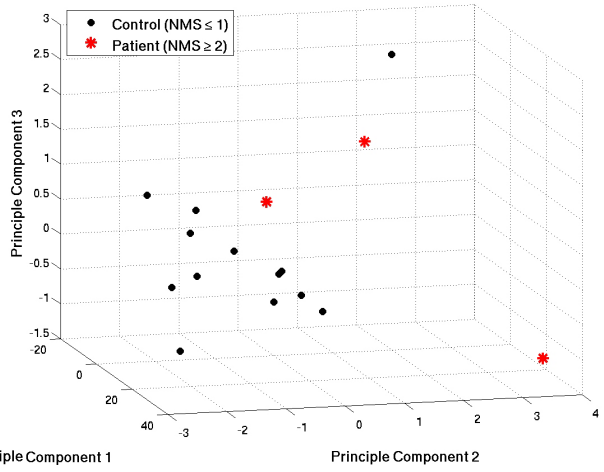


Fig.1 Non-separability in 3 network measures. **Fig. 2** Near-complete separability in the first three principle components of subgraph feature space.

Discussion: We report a high prediction accuracy of poor neurological outcome using unsupervised and supervised learning algorithms performed in a large network measure space. While the sample size is small and the number of patients low, the results of the cross-validation are suggestive and motivate a larger study. We also highlight the approach is easily generalized to any neurological disease which one may hypothesize to affect the connectome.

Acknowledgements: This work is supported by NIH grants R01EB009756 and P50NS035902.

References: [1] Hagmann et al (2007) PLoS ONE 2:e597. [2] Yap et al (2011) PLoS ONE 6(9):e24678. [3] Smith et al (2004) NeuroImage, 23(S1):208-219. [4] Wang et al (2007) Proc ISMRM, #3720. [5] Leopardi (2006) ETNA 25:309-327. [6] Middendorf et al (2005) PNAS 102(9):3192-7. [7] Hajnal et al (1999) Pediatr Neurol 21:788-93. [8] Cristianini (2000) Support Vector Machines, Camb Univ Press.