

Model based cluster analysis to determine unique patient states in traumatic brain injury

T Bylinski, C Hawthorne, M Shaw.

Introduction:

The international BrainIT database contains information on 261 TBI patients¹, including physiological and outcome data. We aimed to identify distinct groups of TBI patients that can be distinguished by their physiological trajectories over time by applying cluster analysis, a form of unsupervised data mining², to the BrainIT database. We hypothesised that the resulting groups would have a good or poor outcome, as measured by the extended Glasgow outcome score (eGOS). This would give an indication of physiological trajectories associated with a good or poor prognosis.

Methods:

After cleaning the BrainIT database, 155 patients remained. We applied the general linear mixed model, a form of partitional cluster analysis, to these patients' data in order to cluster them based on 24-hour trajectories of 5 physiological variables: CPP, ICP, HR, SaO₂, and mean BP. Three variants of this cluster model were run: Algorithm 1 clustered patients using minute-by-minute physiological data, Algorithm 2 used minute-by-minute data with 6 outlying patients removed, and Algorithm 3 used hour-by-hour data with the same 6 outlying patients removed as in Algorithm 2. The resulting clusters' physiological trajectories were represented graphically. Each cluster was also paired to its admissions data, such as first GCS and outcome data, as measured by eGOS.

Results:

Algorithm 1 identified six outliers with distinct and extreme physiological trajectories and eGOS. (n=149, eGOS 5; n=3, eGOS 7; n=2, eGOS 1; n=1, eGOS 1). Algorithm 2 did not separate patients into clusters. Algorithm 3 revealed three clusters (n=65, eGOS 5; n=58, eGOS 5; n=26, eGOS 6) with similar physiological trajectories and outcomes, although the cluster with an eGOS of 6 was the oldest and had a higher CPP and mean BP.

Conclusions:

Model-based cluster analysis of TBI patients based on physiological trajectories is possible. We were able to identify outliers with distinct eGOS scores and physiological trajectories based on minute-by-minute physiology. The model is not able to identify groups of patients with distinct outcomes based on hour-by-hour clustering with outliers removed.

1. Brain IT Group. *Brain Monitoring with Information Technology. About.* [online] Available from: <http://www.brain-it.eu/about>. [accessed 16/04/2015]
2. Jain AK, Murty MN, Flynn PJ. Data Clustering: A Review. *ACM Comput Surv.* 1993; 31(3).