Introduction

Knowledge translation from large clinical data sets is frequently hindered by traditional means of statistical evaluation. Many clinical features co-occur in distinct patient clusters, rather than exhibiting linear relationships.

The objective of this project is to demonstrate how a topological exploration of three distinct clinical databases revealed or confirmed novel subgroups with associated features that may inform medical decision-making.

Methods

Retrospective evaluation of three patient care data sources was approved by the IRB of the University of Texas:

• Specialized Programs of Translational Research in Acute Stroke (SPOTRIAS) thrombolyis registry. (n = 999)
• An institutional database of medication safety alerts for hyperkalemia. (n = 952)
• Pediatric Emergency Care Applied Research Network (PECARN) Traumatic Brain Injury (TBI) study in minor trauma. (n = 42,430)

Topological visualizations of each data set were created in Iris (Ayasdi Inc., Palo Alto, CA). Iris uses similarity measures as dimension reduction to create topological network visualizations.

Results/Discussion

Networks and clusters of interest were identified in exploratory fashion. Categorical data were compared using Pearson’s chi-square, whilst continuous data were compared using Student’s t-test. A predictive model was built for PECARN TBI using a hold-out test set. Topological cluster group membership was used to predict presence of TBI.

Conclusions

Exploratory topological visualization using Iris is a valuable adjunct to traditional methods of statistical evaluation. Correlating identified clusters with clinical knowledge may inform medical decision-making.

Building predictive models based on topological representations of data holds promise. The non-linear co-occurrence of features, with multiple distinct subgroups of patients clustered around an outcome of interest, may support the use of patient-similarity models for outcome prediction.

Acknowledgements

This project was supported by the Agency for Healthcare Research and Quality grant T32HS017586. The sponsor had no role in project approval or design.

Data obtained from public use data sets under data sharing agreement with PECARN. The data provider had no role in project approval or design. Available from: pecarn.org/studydatasets.

Software from Ayasdi was received free of charge via academic licensing agreement.

Contact information: ryan.p.radecki@uth.tmc.edu
Follow me on Twitter @emlitofnote