Topological Visualization Uncovers Novel Clinically Relevant Clusters

Ryan P. Radecki, MD[1]; Eithon Cadag, MS[2];Tanya Petrossian, PhD[2]

[1]The University of Texas Health Science Center at Houston; [2]Ayasdi, Inc.

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) thrombolysis 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.

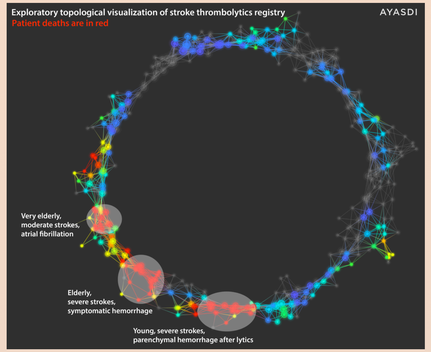


Figure 1. Exploratory visualization of SPOTRIAS.

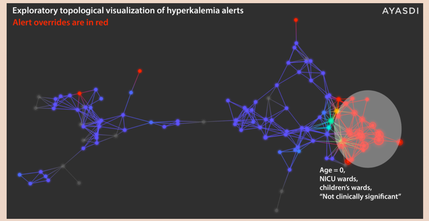


Figure 2. Exploratory visualization of alert database.

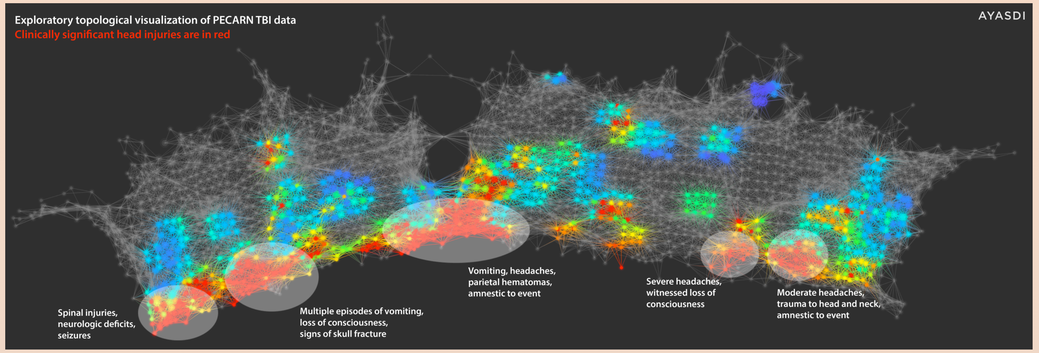
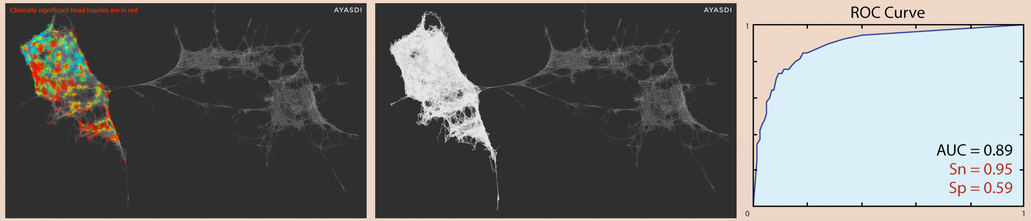


Figure 3. Exploratory visualization of PECARN TBI dataset, annotated with novel clinical clusters of cTBI.



Figures 4-6. Using an alternative PECARN TBI topology for predictive model, with ROC performance as above.

Methods Cont.

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.

Results/Discussion

Figure 1 above displays network clusters annotated with clinical features predictive of death after thrombolytic administration in acute ischemic stroke.

Figure 2 identifies overridden medication safety alerts for hyperkalemia. These alerts were later identified as erroneous, and based on neonatal heel-stick blood.

Results/Discussion Cont.

Figure 3 displays several novel clusters of clinical features associated with pediatric TBI in minor trauma.

Figures 4-6 show a network used to define a group containing all instances of TBI. Group membership was used to predict TBI in the hold-out set with an AUC of 0.89.

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 toplogical 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