

Adolescent Depression
Symptom Clustering:
Machine Learning Driven
Hypothesis Generation

Eric Lin, MD, Victor Avila, MD, Jose Flores, MD, PhD and Michael Bloch, MD, MS

BACKGROUND: Despite consensus about the heterogeneity of major depression, depression subtype models have not converged for adolescent depression. While machine learning approaches have been used to study adult depression subtypes, it has not yet been applied to the pediatric population.

METHODS:

- 1. Used the TADS trial (Treatment for Adolescents with Depression Study)
 - a. n = 439
 - b. Examined baseline symptom clustering
 - i. Primary measure: Children’s Depression Rating Scale – Revised (CDRS)
 - ii. Secondary measures: Reynold’s Adolescent Depression Scale (RADS) & Beck Depression Inventory (BDI)
- 2. Applied hierarchical clustering
- 3. Sensitivity analyses
 - a. Examined clusters with different times, parameters, clinical scales, and methods
 - i. K-means clustering
 - ii. HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise)
 - iii. t-distributed Stochastic Neighbor Embedding

DISCUSSION:

- 4. 3 cluster solution (with several unclustered items) appears to emerge
 - a. Silhouette score = 0.252, bootstrapped 95% CI (0.226, 0.270)
 - i. Scores bounded between -1 and +1 for incorrect to dense clustering
 - ii. Scores around 0 indicate overlapping clusters
- 5. Without clear consensus in cluster evaluation, sensitivity analyses are used:
 - a. Clusters are comparable across techniques (Tables 1 & 2), time, parameters (not shown)
 - b. Comparison across clinical scales was limited by differences in scale items
- 6. Such preliminary findings require external validation or replication with larger sample

We investigated symptom clusters of adolescent major depressive disorder using unsupervised machine learning techniques.

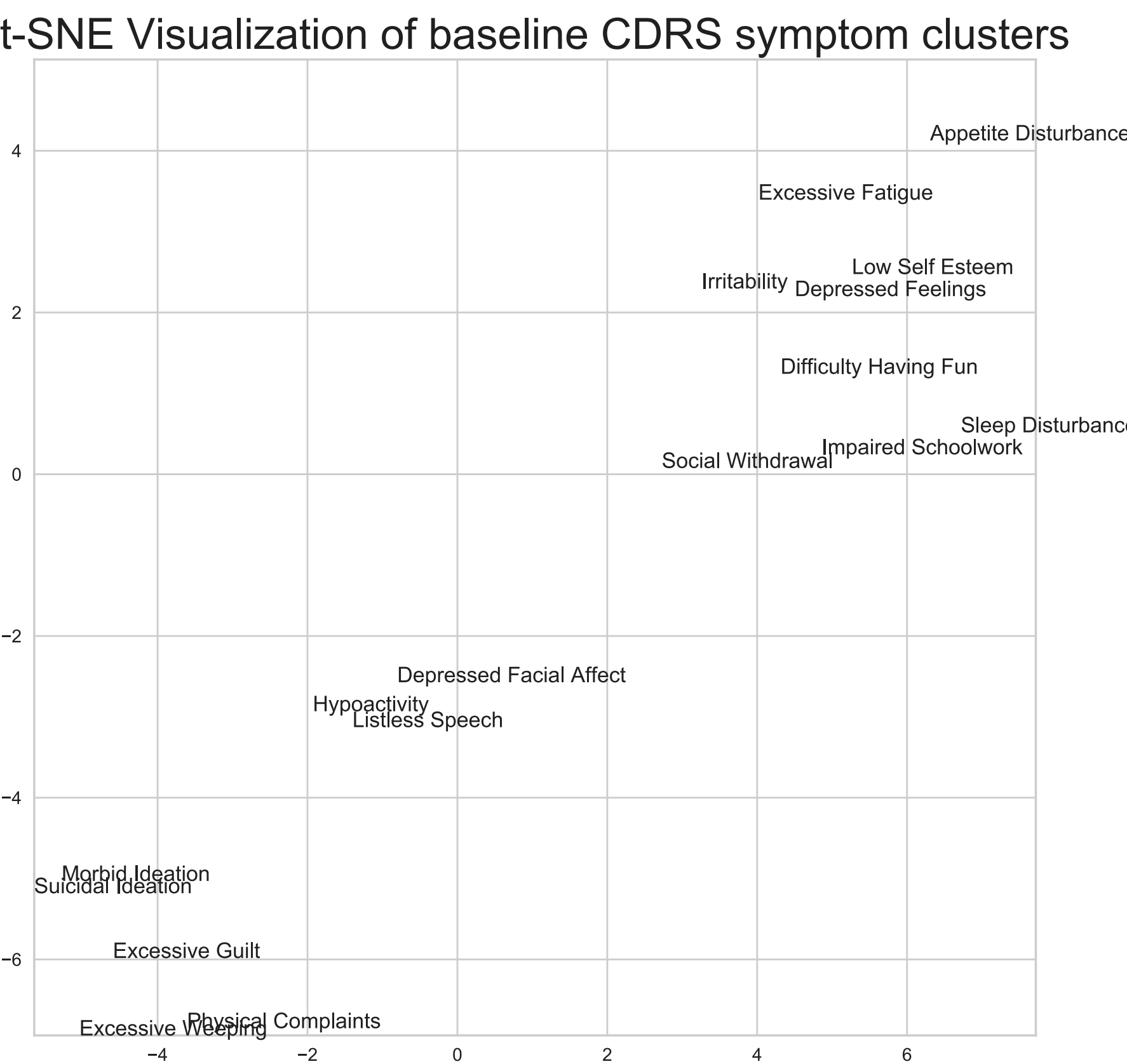
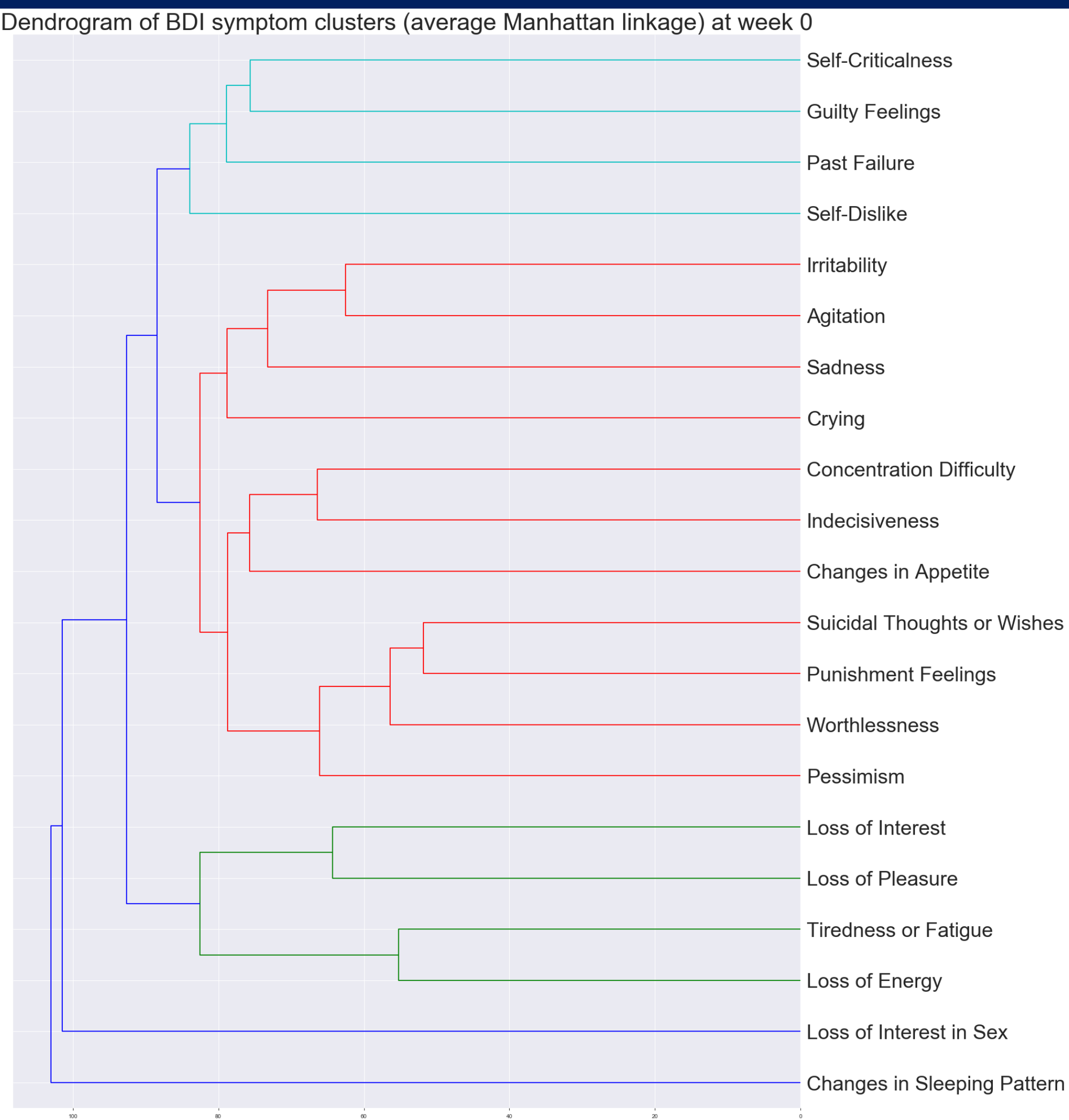
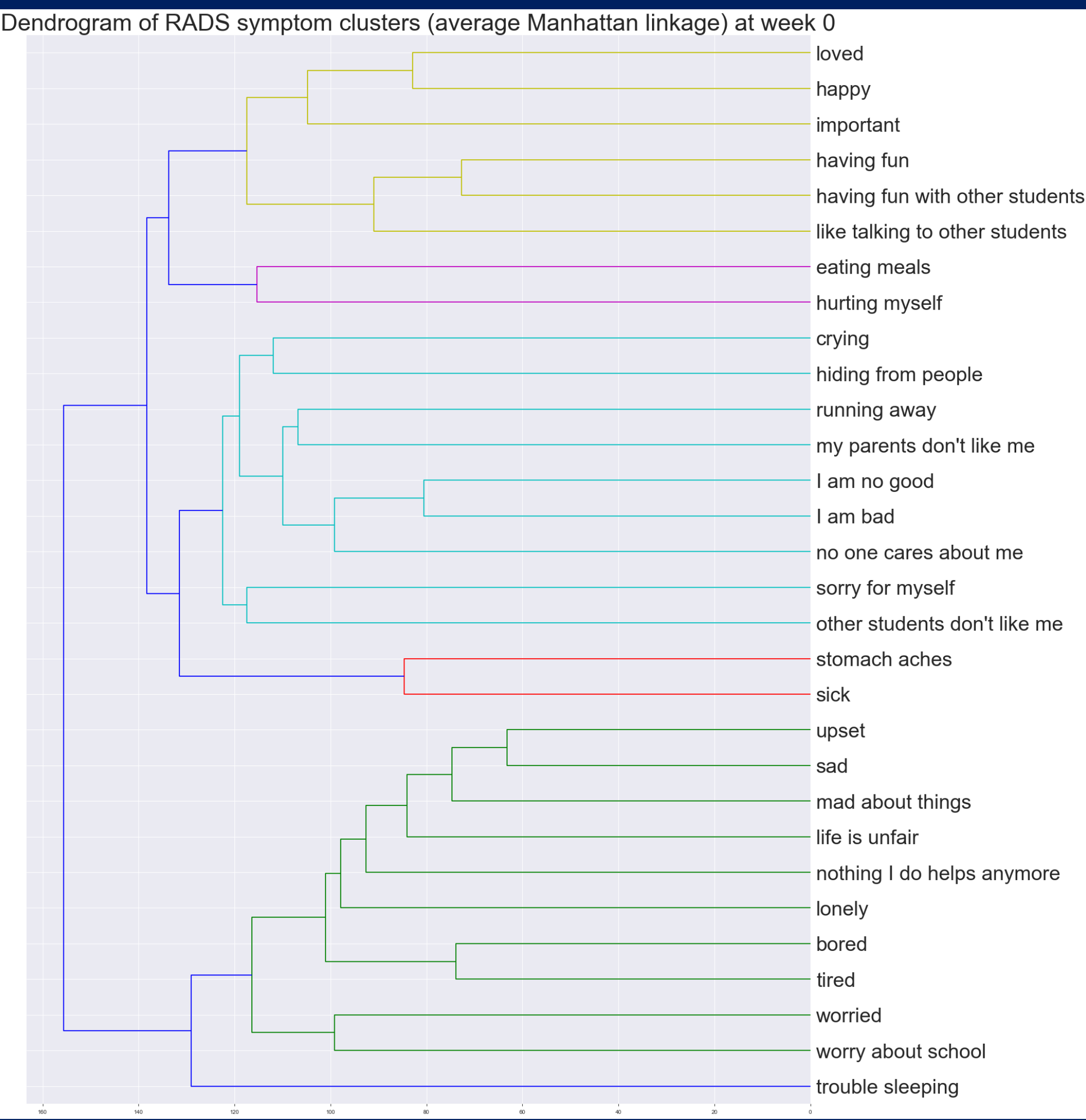
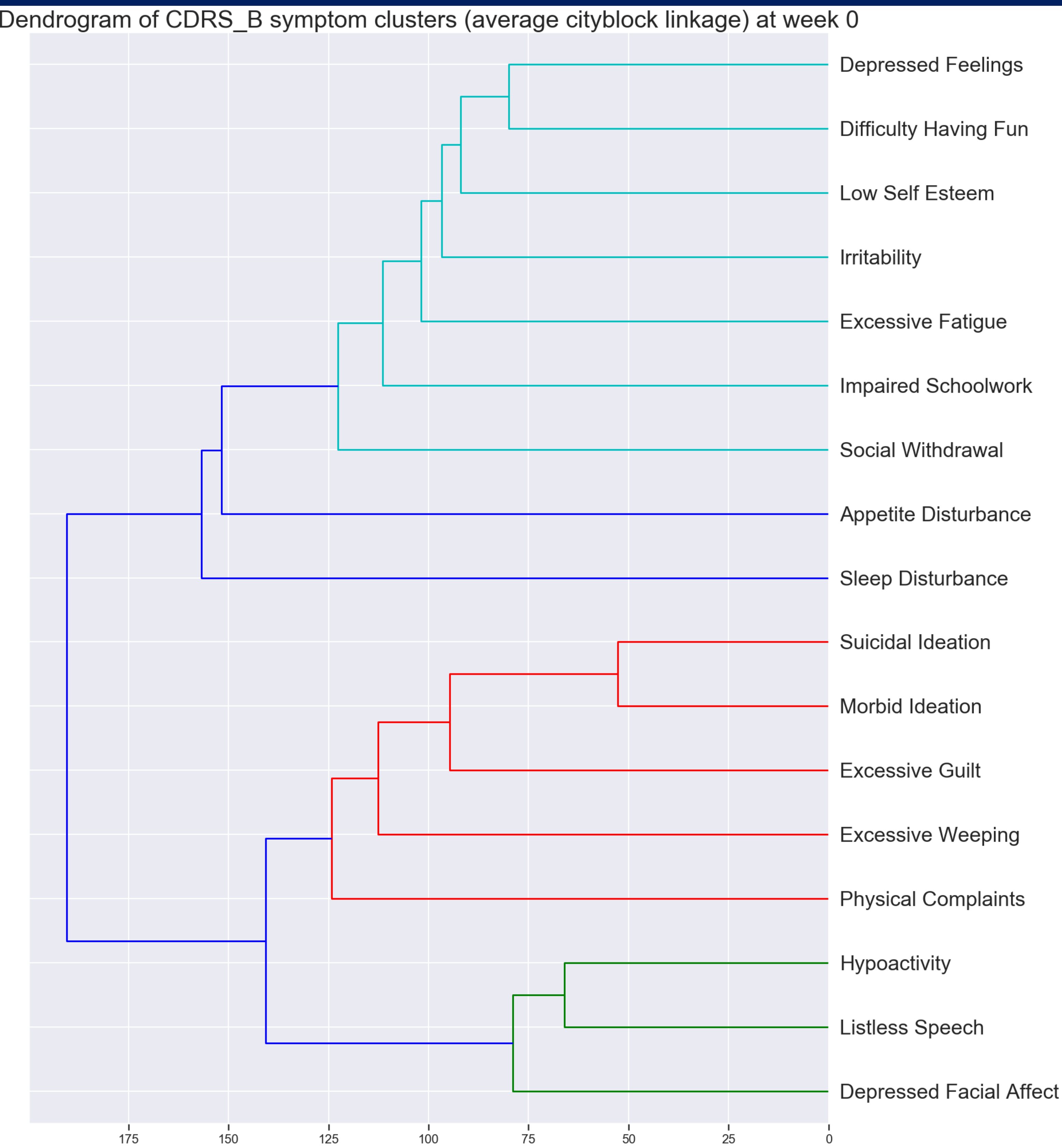


Table 1. K Means Clusters (k=5) for Baseline CDRS	
Cluster 1	Physical Complaints, Excessive Guilt, Morbid Ideation, Suicidal Ideation, Excessive Weeping
Cluster 2	Impaired Schoolwork, Difficulty Having Fun, Social Withdrawal, Excessive Fatigue, Irritability, Low Self Esteem, Depressed Feelings
Cluster 3	Depressed Facial Affect, Listless Speech, Hypoactivity
Cluster 4	Sleep Disturbance
Cluster 5	Appetite Disturbance

Table 2. HDB Scan Clusters for Baseline CDRS	
Cluster 1	Excessive Guilt, Morbid Ideation, Suicidal Ideation, Excessive Weeping
Cluster 2	Impaired Schoolwork, Difficulty Having Fun, Social Withdrawal, Excessive Fatigue, Irritability, Low Self Esteem, Depressed Feelings
Cluster 3	Depressed Facial Affect, Listless Speech, Hypoactivity
Unclassified	Sleep Disturbance, Appetite Disturbance, Physical Complaints



Contact for questions:
QR Code for LinkedIn
eric.lin@yale.edu

