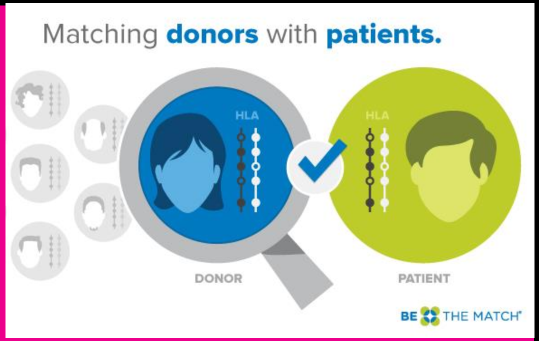
MEDINFO23

8 – 12 JULY 2023 | SYDNEY, AUSTRALIA



Ensemble Clustering to Generate Phenotypes of Kidney Transplant Donors and Recipients

Syed Sibte Raza Abidi, Kranthi Jalakam, Syed Hani Raza Abidi, Karthik Tennankore

Dalhousie University, Canada

Background

- Kidney transplantation is a treatment option for end-stage kidney disease
- Donor and recipient characteristics impact outcomes for kidney transplant
- Understanding the relationships between donor and recipient characteristics
- Understanding the relationships between donor and recipient characteristics and how they influence post transplant outcomes may have future implications for kidney allocation

Objectives and Approach

- To identify phenotypes of kidney transplant recipients using their clinical characteristics (inclusive of both donor and recipient features)
- To determine how these phenotypes influence graft survival.

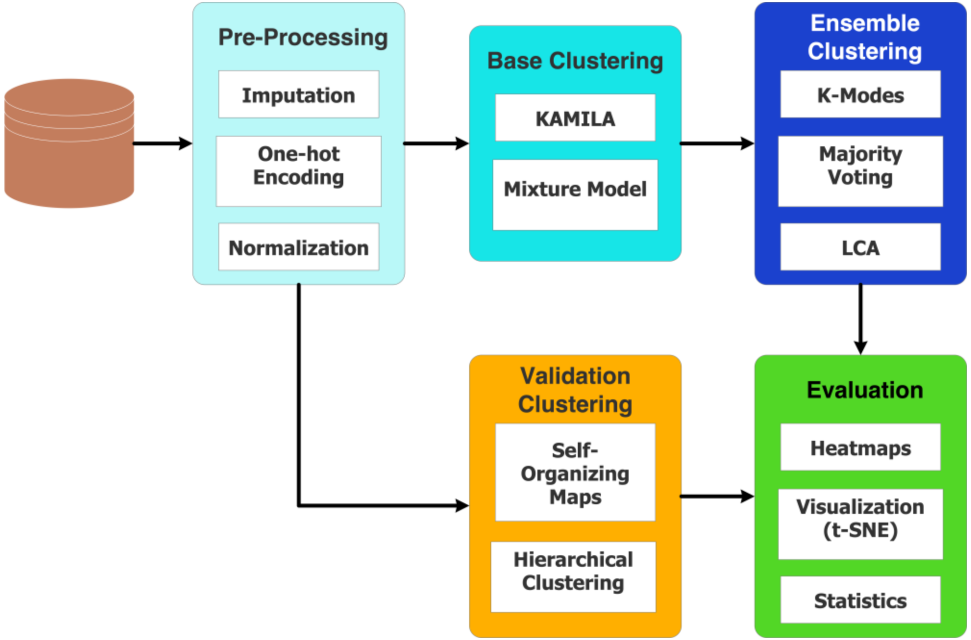
- Investigate a dataset of kidney transplant donor and recipient characteristics
- Generate phenotypes by applying unsupervised clustering to multi-type data

Data

- Scientific Registry of Transplant Recipients (SRTR)—2009 to 2011
- 25,824 recipients of a deceased donor kidney transplant
- 28 numerical and categorical variables
- Recipient outcome (experienced graft failure, died or survived) not used

Methods

- Multi-type data clustering
  - Model based clustering: Mixture models and Kay-means for Mixed Large data (KAMILA)
  - Neural network clustering: Self Organizing Maps
- Ensemble clustering
- Cluster evaluation metrics
  - Silhouette score, Dunn index, Calinski-Harabasz score

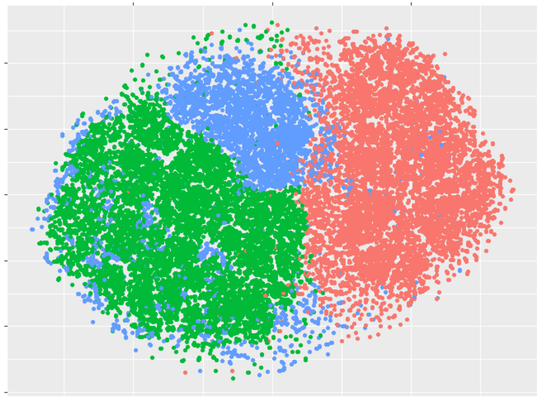


Results – Base vs Ensemble Clustering

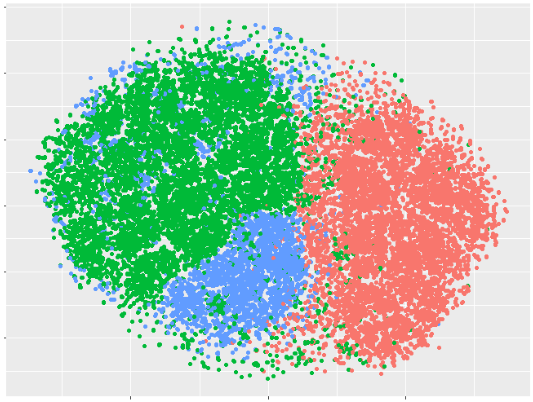
Method	Silhouette Score (↑)	Dunn Index (↑)	CH Index (↑)
Base Clustering			
KAMILA	0.108	0.0536	2760
Mixture model	0.091	0.0099	2514
Ensemble Clustering			
k-modes	0.108	0.0200	2704
Majority Voting	0.076	0.0282	2065
Latent Class Analysis (LCA)	0.113	0.0538	2843
Validation Clustering			
Self-Organizing Map (SOM)	0.100	0.0512	2595

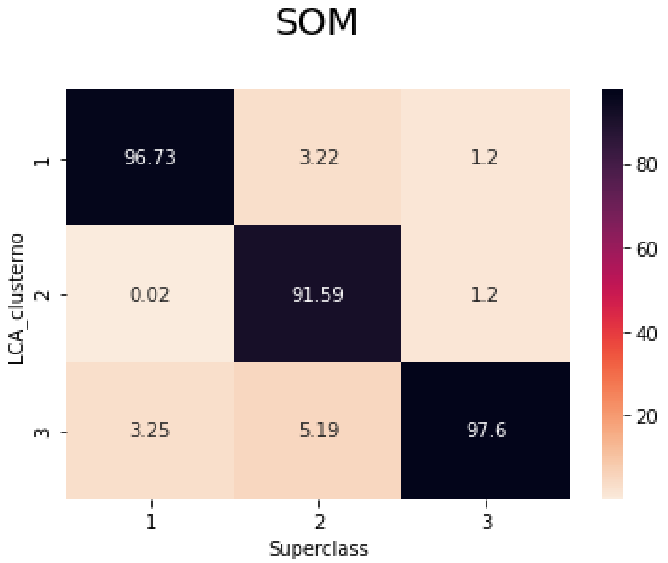
Clustering Results

- Latent Class Analysis
  - Cluster 1:11472
  - Cluster 2: 9583
  - Cluster 3: 4769

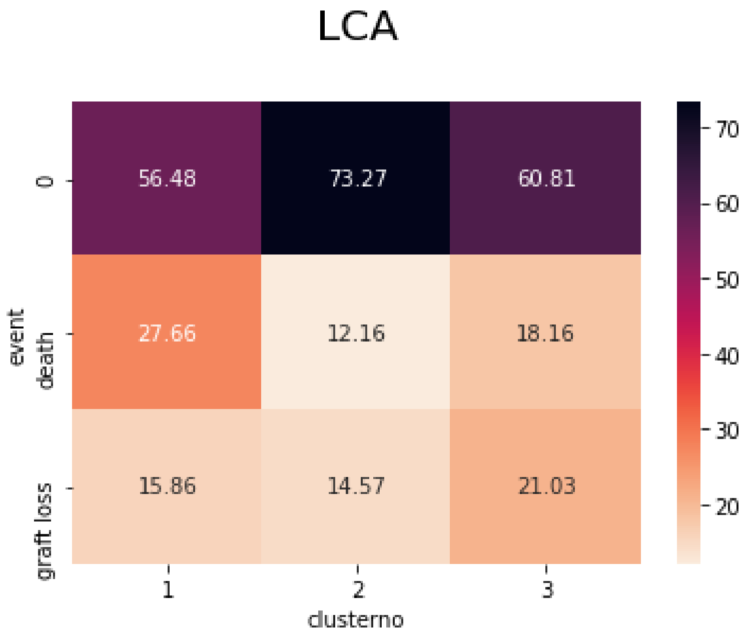


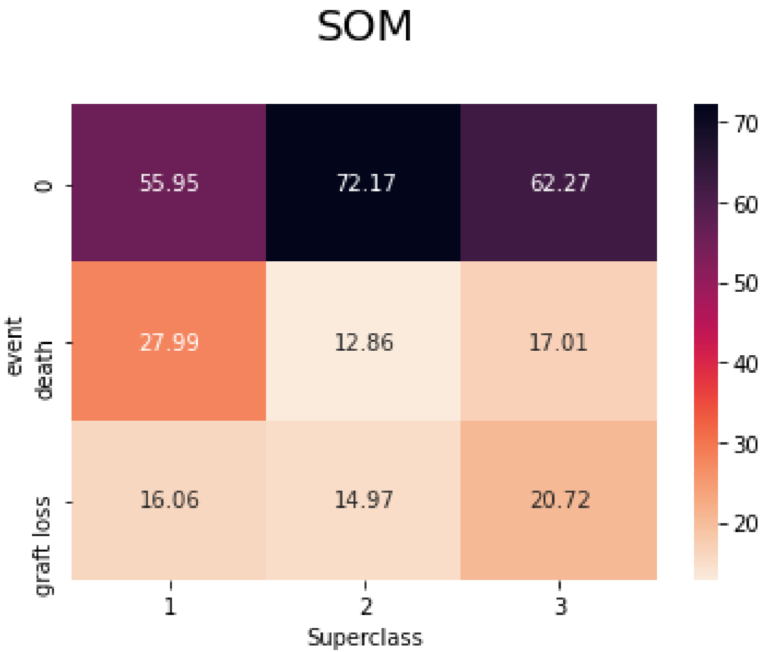
- Self-Organizing Map
  - Cluster 1:12472
  - Cluster 2: 9443
  - Cluster 3: 3909





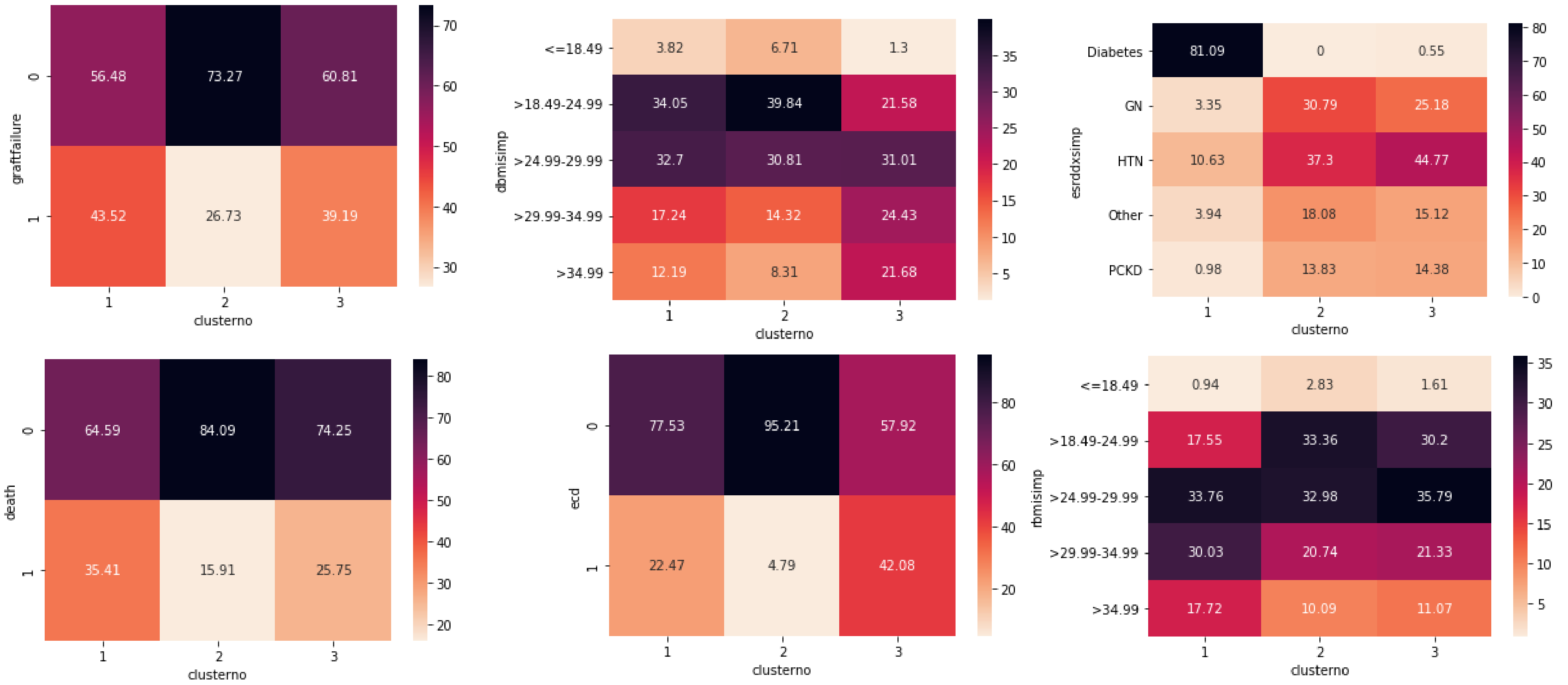
Cluster Validation





MEDINFO23

8 – 12 JULY 2023 | SYDNEY, AUSTRALIA



Phenotype Feature	Cluster A (11,472)	Cluster B (9,583)	Cluster C (4,769)
Donor & Recipient Age	Youngest Donor & Recipient	Oldest Recipient	Oldest D & Older R than cluster A
Survival Time	Longest	Similar	
Years on dialysis pre-transplant	Longer than cluster B	Shortest	Longest
Recipient Event	Highest censored	**Highest Death**	**Highest Graft Loss**
Recipient Death	Lowest	**Highest**	Higher than cluster A
ESRD Diagnosis	GN and HTN	Diabetes	GN and HTN
Recipient Diabetes	None	**Almost entirely Yes**	Mostly No
Recipient Sex	**Highest proportion of female sex**	Lowest	Higher than cluster B
Recipient BMI	Higher presence between 18.49 and 29.99 than cluster B	**Significantly higher >29.99 and lower >18.49 – 24.99**	Similar to cluster A
Donor Hypertension	Almost None	Majority No with significant Yes	**Almost entirely Yes**
Expanded Criteria Donor	Almost None	Majority No with significant Yes	**Highest Yes (~40%)**
Donor Race	Similar		**Highest proportion of black donors**
Donor Diabetes	Almost None	Slightly higher than cluster A	**Highest Yes**
Donor BMI	Highest >18.49 – 24.99	Higher presence >29.99 than cluster A	**Significantly higher >29.99 and lower >18.49 – 24.99**

Concluding Remarks

- Phenotypes of kidney donors and recipients
  - Characterize anticipated outcomes for each cluster assignment
- Outcomes not used for cluster derivation
  - Clusters identify high risk phenotypes at differential risk of graft failure
- Multi-type data ensemble clustering approach
- Observed concordance between the clustering method provides a new strategy to validate cluster membership
- Generalizable to other organ transplant studies

THANK YOU