Innovative methods to validate latent groups of high-risk primary care populations using Mixture-Item Response Theory with Clinical Diagnoses

Adam J Batten[1], Sandeep Vijan[2], Rachel Piegari MS[3], Erin Jaske[1], Karin Daniels[4], Kari Nelson[1,5], Stephan D Fihn[5], Ann-Marie Rosland[4]

[1]VHA Primary Care Analytics and Evaluation, VA Puget Sound Health Care System; [2]Ann Arbor VA Center for Clinical Management Research [3]VHA Office of Clinical Systems Development and Evaluation;[4]Center for Health Equity Research and Promotion, VA Pittsburgh Healthcare System; [5]Division of General Internal Medicine UW, Seattle

PACT Demonstration Laboratory Initiative

BACKGROUND

- Mixture-item response theory (M-IRT) or latent trait analysis has been used to categorize high risk patients into subgroups with unique diagnosis.
- Latent multimorbid subgroups offer the potential to tailor care practices to the specific needs of otherwise undifferentiated high-risk patients.
- Because latent trait analysis depends heavily on the underlying data structure, validation is necessary.

OBJECTIVES

- Test the primary assumption of the M-IRT method, that the sum of conditions is a sufficient statistic for a patient.
- Identify a parsimonious or best clustering.

POPULATION STUDIED

940,714 VHA primary care patients in 2014 with a predicted probability of hospitalization in the next year >0.25

METHODS

- **Cohort Design with latent traits**
- **Pre-Conditioning variables:** Care Assessment Needs (CAN) Score
- **Traits (independent variables)**: 30 ICD9 based comorbid conditions.
- **Analysis:** Relying on bootstrap resampling, bagging, and grid search methods we tested the primary assumptions of the M-IRT model across 3 dimensions, including:
  1. Model fit: An Information Criteria (AIC). Figure 1
  2. Patient level subgroup similarity: the Variance of Information (VI) statistic. Figure 2
  3. The rank order of condition burden (item difficulty) or condition prevalence in the subgroups: Kendall's τ. Figure 3
- **Prediction:** A Neural Net is trained to predict subgroup and patient complexity on new and unseen patients.

M-IRT RESULTS

Figure 1: 5-8 subgroup models are the most parsimonious.

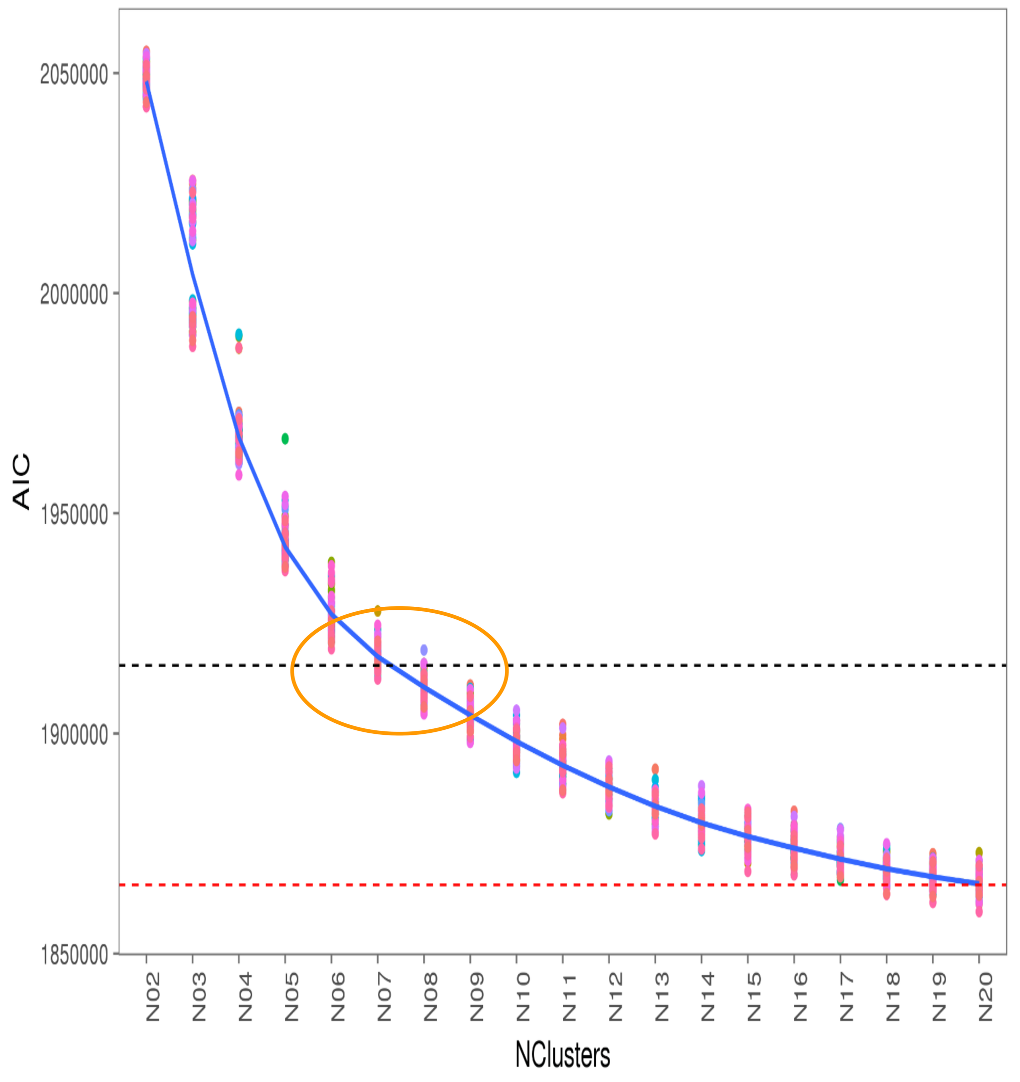


Figure 2: VI heatmap reveals a clear separation between the 2-5, 6-14, and 15-20 cluster specifications.

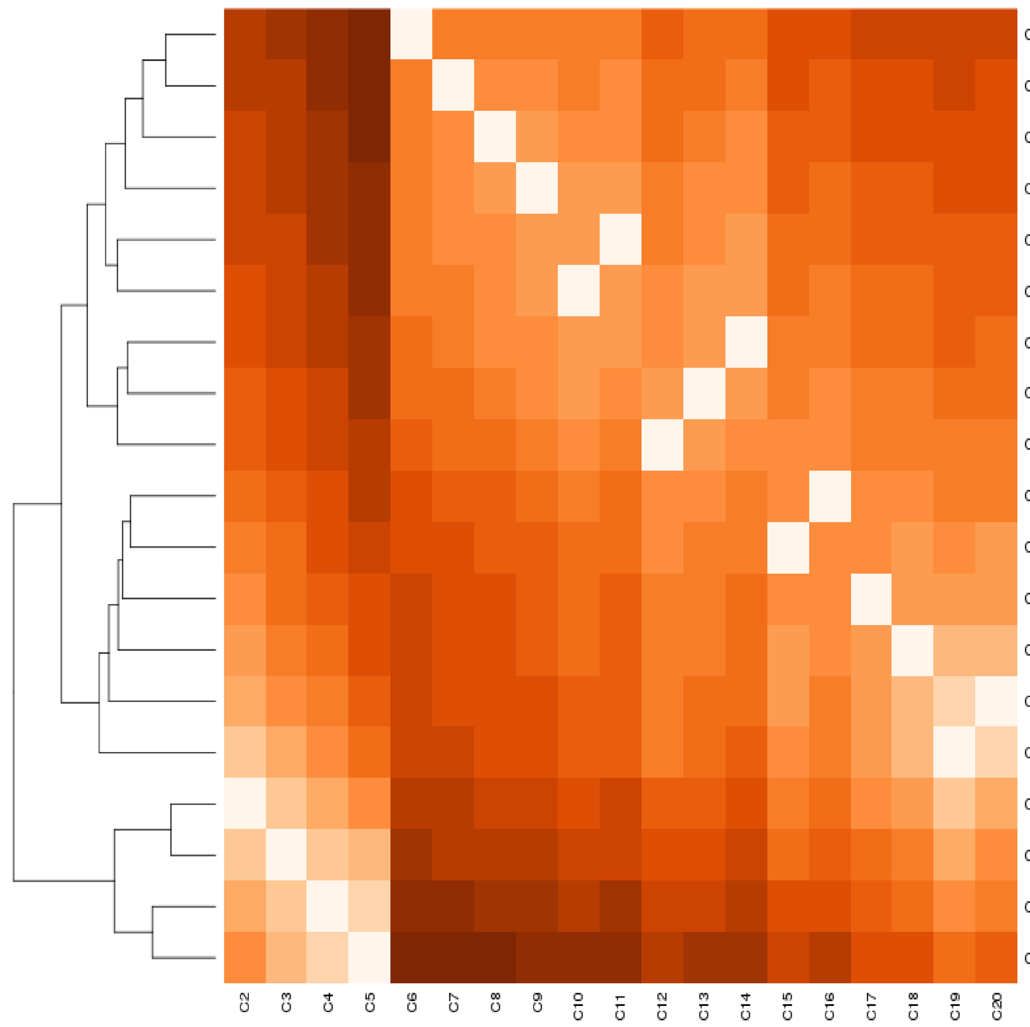
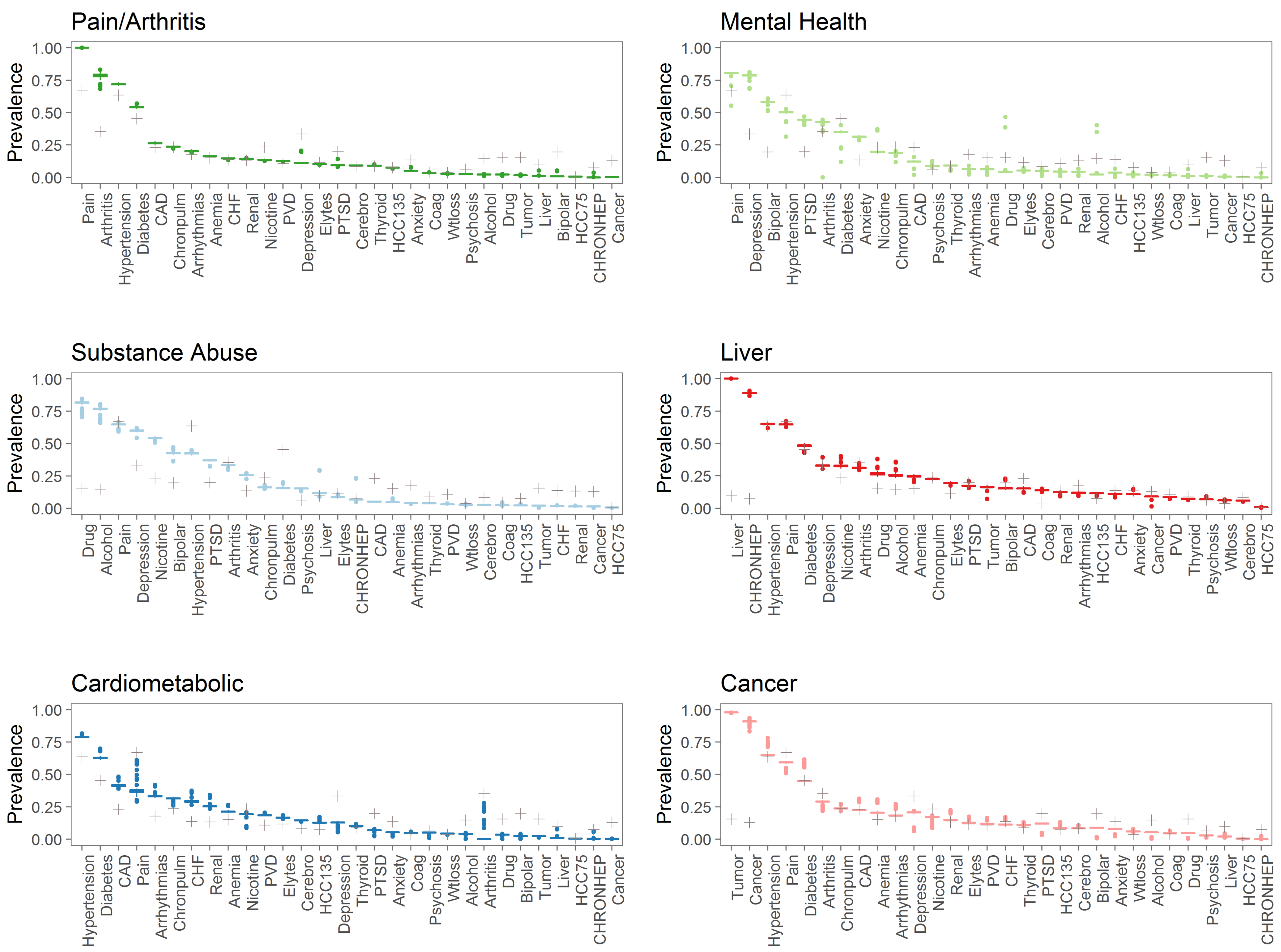


Table 1: Most common subgroups identified by the procedure.

..	Top 3 Conditions	Samples	HRSubgroup
1	Cancer, Depression, Tumor	13	
2	Chronic Arthritis, Depression, Diabetes	16	
3	Chronic Artery Disease, Chronic Arthritis, Diabetes	83	Pain/Arthritis
4	Bipolar, Depression, PTSD	85	Mental Health
5	Alcohol, Depression, Drug Abuse	90	Substance Abuse
6	Chronic Hepatitis, Diabetes, Liver Disease	98	Liver
7	Arrhythmias, Chronic Artery Disease, Diabetes	100	Cardiometabolic
8	Cancer, Diabetes, Tumor	100	Cancer

Figure 3: Significantly different diagnostic characteristics among the subgroups.



PERCEPTRON RESULTS

- Using cross validation with stratified partitioning we fit a neural net with SAS Enterprise Miner's Autoneural Node.
- Outcome variables: subgroup and patient complexity.
- Independent variables: 30 ICD9 comorbid conditions.
- Optimal model identified by balancing the average squared error of prediction (ASE) with the risk of overfitting.
- The best model consisted of a single hidden layer and had an overall ASE of 0.09.

CONCLUSION

- Approaches typically used in machine learning can provide a robust test of the assumptions underlying soft clustering methods like latent trait models.
- Neural nets can be trained to replace these models.

IMPLICATIONS

EMR-based, robust and validated groupings of high risk patients have the potential to assist health care teams in efficiently developing more tailored, and possibly more effective, care management plans for heterogeneous high risk patients.

ACKNOWLEDGEMENTS

This work was undertaken as part of the Veterans Administration's PACT Demonstration Laboratory initiative, supporting and evaluating VA's transition to a patient-centered medical home. Data for this report were developed by the national evaluation team at the PACT Demonstration Lab Coordinating Center and the VHA Clinical Systems Development and Evaluation. Funding for the PACT Demonstration Laboratory Initiative is provided by the VA Office of Primary Care. The views expressed in this poster are those of the authors and do not necessarily reflect the position or policy of the Veterans Administration.