USING RWD TO IDENTIFY SUBPOPULATIONS OF NAFLD AND ASSOCIATIONS WITH CARDIOVASCULAR AND LIVER-RELATED OUTCOMES

Laura Evans, Seth Kuranz, Richard Rast
TriNetX, Inc. Cambridge, MA, USA | seth.kuranz@trinetx.com

TriNetX

OBJECTIVES

The aims of this study were to:

(1) Identify NAFLD subgroups within a large RWD network using both traditional and novel risk factors described in the literature
(2) Measure the occurrence of cardiovascular and liver-related outcomes across subgroups.

METHODS

NAFLD patients from the TriNetX network were defined based on the selection criteria in Figure 1. To identify NAFLD subgroups, a k-means clustering model incorporated NAFLD risk factors described in Table 1 and Figure 2. The incidence of NASH, fibrosis/cirrhosis, HCC, cardiovascular disease, and all-cause mortality was measured across clusters. All definitions were based on ICD9/10 and LOINC codes.

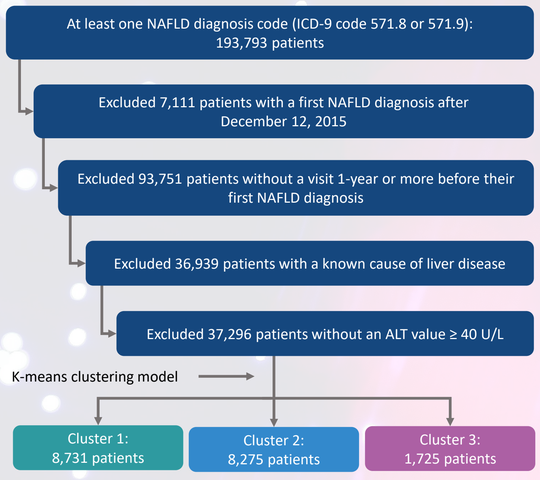
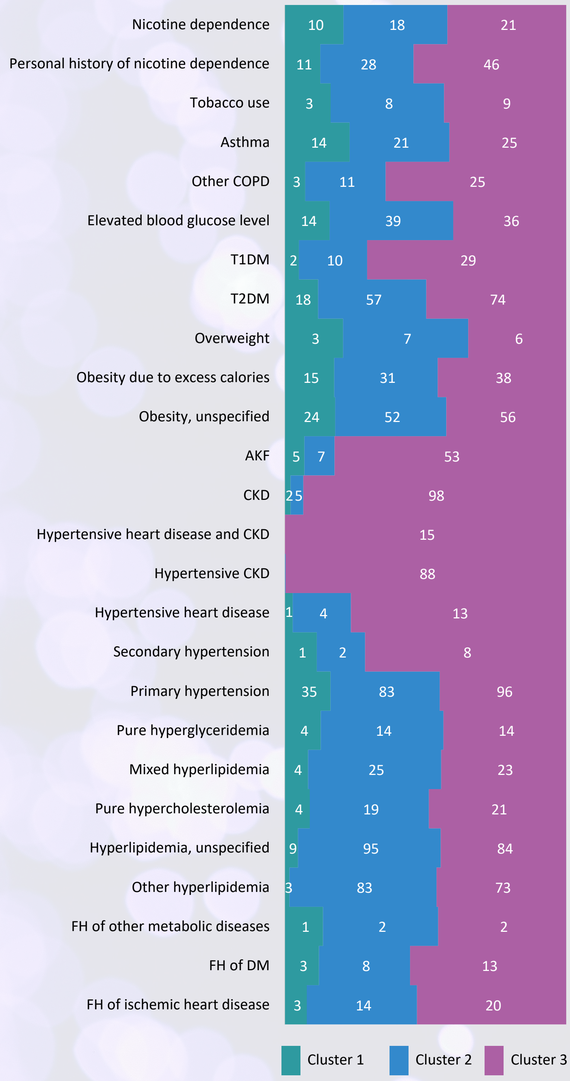


Figure 1. Patient flow diagram

	Cluster 1	Cluster 2	Cluster 3
N	8,731	8,275	1,725
Mean age	42	53	58
Female (%)	49	53	51
Hispanic (%)	10	5	4
White (%)	76	86	80
Black (%)	7	6	12

Table 1. Demographic characteristics in clustering model



Note: numbers reflect the percentage of patients in each cluster with the characteristic in each row.

Figure 2. Characteristics included in clustering model

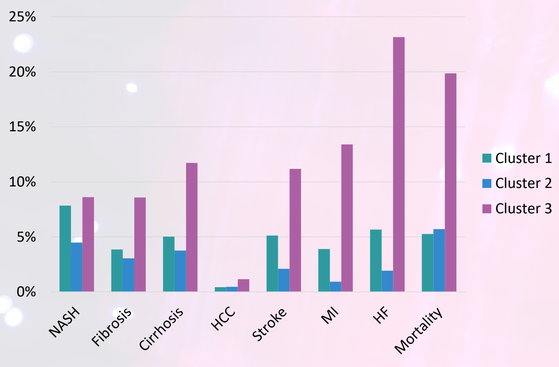


Figure 3. Incidence of outcomes by cluster

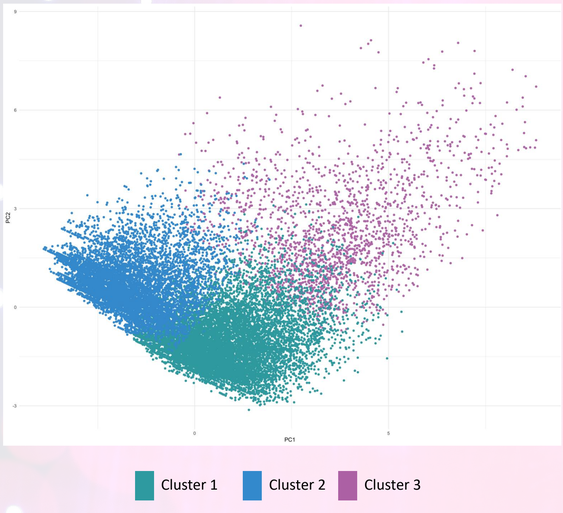


Figure 4. Visualization of NAFLD clusters in two dimensions

	Cluster 1	Cluster 2	Cluster 3
Mean creatinine (mg/dL)	0.9	1.3	3.3
Mean eGFR (MDRD; mL/min/1.73 m²)	81.2	64.4	31.0

Table 2. Kidney biomarkers by cluster

RESULTS

The k-means clustering model identified 3 clusters. Cluster 1 had the lowest prevalence of comorbidities. Cluster 2 had a high prevalence of hyperlipidemia and primary hypertension, but a low prevalence of secondary hypertension. Cluster 3 had the highest prevalence of hypertension, type 1 diabetes, kidney disease, and kidney-related hypertension. Creatinine and eGFR laboratory results confirm Cluster 3 is characterized by decreased kidney function (Table 2). Cluster 2 was similar to Cluster 1 with respect to fibrosis/cirrhosis and all-cause mortality, but not cardiovascular disease.

CONCLUSIONS

This analysis identified three distinct groups with varying levels of cardiovascular and liver-related risk. Comorbid kidney disease was a significant differentiator for Cluster 3. Lower risk patients found in Cluster 1 should be further explored to identify early NAFLD intervention opportunities.

Abbreviations: RWD = Real world data, NAFLD = Non-Alcoholic Fatty Liver Disease, NASH = Nonalcoholic steatohepatitis, FH = family history, AKF = acute kidney failure, CKD = chronic kidney disease, DM = diabetes mellitus, T1DM = type 1 diabetes mellitus, T2DM = type 2 diabetes mellitus, HCC = hepatocellular carcinoma, MI = myocardial infarction, HF = heart failure.