University of Florida Division of Transplantation and Hepatobiliary Surgery

Identifying Clinical and Molecular Determinants of Health Disparities in Hepatitis C

Mariah Jordan, Protiva Rahman, Ali Zarrinpar

BACKGROUND

Despite the curability of hepatitis C (HepC), racial and ethnic minority populations are disproportionately unaware of their HepC status and encounter significant barriers to treatment when diagnosed, leading to a higher burden of liver disease[1,2]. We sought to enhance and validate a binary classification model for assessing HepC risk using various patient selection criteria assumed to correlate with HepC infection. Our hypothesis is that incorporating race, ethnicity, and related clinical covariates will improve the model's predictive accuracy for identifying HepC patients within the UF Health cohort.

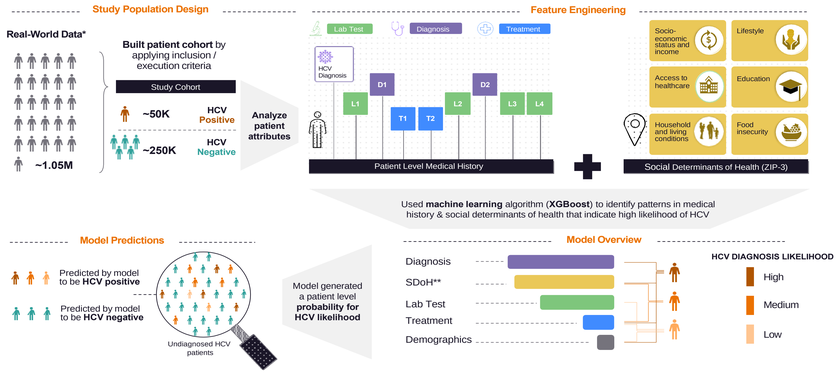
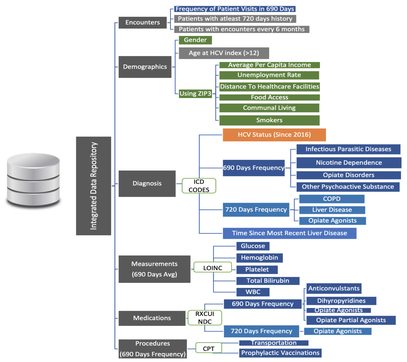


Figure 1: Overview of Preliminary Study Design: A binary classification model was developed using a large-scale US-based real-world dataset, to predict the likelihood of HCV diagnosis; *Total HCV positive & negative patients from 2014-2020; **SDoH - Social Determinants of Health

METHODS



We utilized the UF Integrated Data Repository (IDR) to gather data on diagnoses, procedures, medications, labs, and zip codes for patients over 11 years of age. The data were filtered to include only patients with records of HCV antibody tests, HCV RNA tests, HCV diagnosis codes, and HCV treatment results. Patients were then categorized into HCV positive and negative groups and excluded based on criteria including an index date before January 2016, age under 12, less than 720 days of medical history, and non-continuous semesterly enrollment. We generated 24 time-based features related to diagnoses, labs, treatments, and procedures, and 8 features related to gender, age, and social determinants of health.

Figure 2. Data Protocol.

RESULTS

Data Assessment

Cross-Data Cohort Comparison

Dataset	Cohort Size	HCV Positive	HCV negative	Male	Female
Original Training Data	295,244	50,081 (17%)	245,163 (83%)	115,848 (39%)	179,396 (61%)
UF Data	24,598	2,695 (11%)	21,903 (89%)	11,180 (45%)	13,418 (55%)

Figure 3. Cohort Comparison: While fine-tuning the model, we under-sampled the HCV negative patient population to address class imbalance and to maintain the similar ratio of HCV positive to HCV negative populations as training data.

Cross-Data Feature Comparison

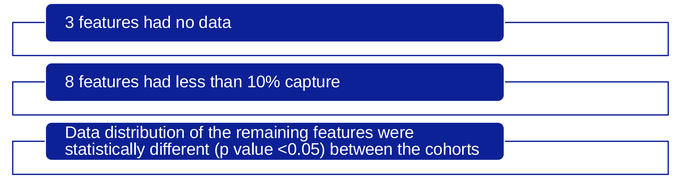


Figure 4. Feature Comparison.

Model Performance

Metric	Description	Performance on Training Data[3]	Performance on UF Data
AUROC	Overall predictive ability of model	95%	84%
Precision	When model predicts HCV positive, how often is it correct?	93.5%	80%
Sensitivity *(Recall)*	How many of the true HCV positive patients are identified by the model?	50%	35%

Figure 5. Model Performance: On UF data the model was able to score ~80% precision while maintaining the sensitivity of ~35% at an AUROC of 84%.

KEY TAKEAWAYS

- Difference in **model performance** can be attributed to inherent differences in data distributions between training & UF data
- Such **performance differences** are expected as seen in similar applications in multiple **industry use cases**[3].
- 84% **AUROC implies** that model **accurately distinguishes** the HCV positive vs negative patients with high confidence
- Precision and sensitivity numbers can be further adjusted based on business requirements.

CONCLUSIONS

We hypothesized that incorporating race, ethnicity, and related clinical covariates will improve the model's predictive accuracy for identifying HepC patients within the UF Health cohort. The ability of the final model to capture undiagnosed HCV patients while producing limited false alarms is directionally comparable to the model trained on EHR data, amidst the inherent differences in datasets. Immediate next steps include feature readjustments for bilirubin and model re-validation. Future directions include collaborating with other healthcare systems to expand the model's application and to further validate and refine the model. We also plan to implement the HepC prediction model in healthcare systems and testing its effectiveness in improving HepC screening rates and subsequent diagnosis and treatment, ultimately reducing the incidence of HCC and the disparities.

UF College of Medicine UNIVERSITY of FLORIDA

REFERENCES

1. Falade-Nwulia, Oluwaseun et al. "Hepatitis C in Black Individuals in the US: A Review." *JAMA* vol. 330,22 (2023): 2200-2208. doi:10.1001/jama.2023.21981
2. Centers for Disease Control and Prevention. Viral Hepatitis Surveillance Report – United States, 2020. https://www.cdc.gov/hepatitis/statistics/2020surveillance/index.htm. Published September 2022. Accessed 9/3/24
3. (*ATTR-CM* , *NASH* , *https://pubs.rsna.org/doi/10.1148/ryai.210064* , *https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0226255&type=printable*)

ACKNOWLEDGEMENT

Funding provided by The American Society of Transplant Surgeons Socio Economic and Racial Disparity Grant