ADDRESSING THE DUAL RISK OF LIVER DYSFUNCTION AND TUMOR PROGRESSION IN HEPATOCELLULAR CARCINOMA THROUGH SUPERVISED CLUSTERING

Author

Abdelghani Halimi

Linked in

SCAN ME

Supervisors

Nesma Houmani

Sonia Garcia-Salicetti

s@movar

Collaborators

Ilias Kounis

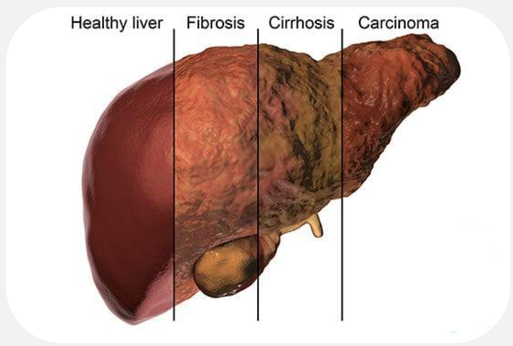
Audrey Coilly

(Chaire BOPA)

1) CONTEXT & OBJECTIVES

Introduction

- **Liver cancer** → 3rd leading cause of **cancer-related deaths** (2020)
- **Hepatocellular carcinoma (HCC)** → most common type
- **Liver transplantation (LT)** = definitive therapy for early-stage HCC
- **Problem**: Organ scarcity → need for efficient **waitlist prioritization**



Current Best Practices

MELD (Model for End stage Liver Disease) scores to objectively prioritize patients with highest risk of death within 3 months.

Designed for **cirrhosis**, not **tumor-related risks**

Existing Risk Scores in the litterature

- **Liver function–based**: Child-Pugh, ALBI, MELD-based scores
- **Tumor-based**: AFP score
- **Both types of risks**: HALT-HCC, Mehta Model

Use of **linear models** with a **limited** number of risk factors
→ **miss complex interactions**

Goal

- Combine predictive **accuracy**, **interpretability**, and **personalization** to improve liver transplant prioritization for HCC patients using ML

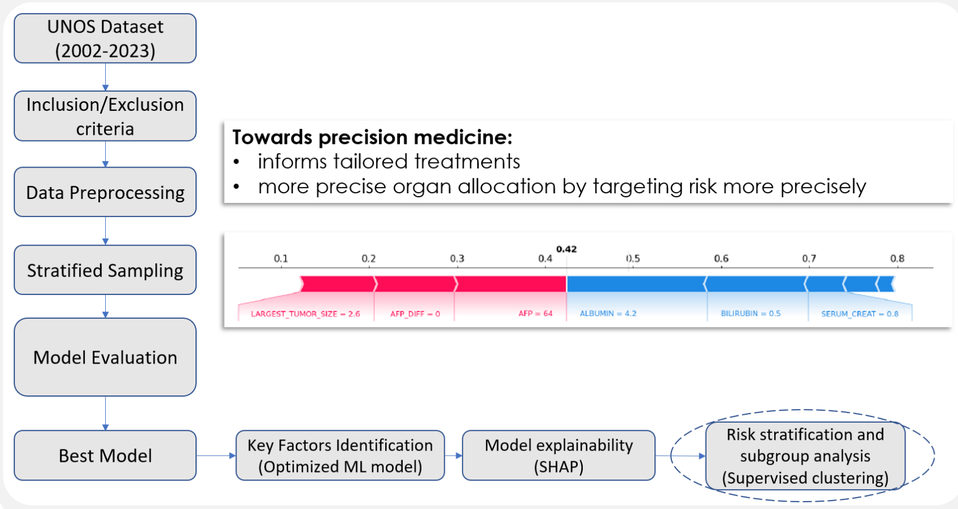
2) DATA & METHODS

UNOS* Dataset	Classes	
	Died	Survived
Patients	448	11,199
Observations (visits)	1,053	31,300

- **A tree-based ensemble Learning model** for mortality prediction
- **Classify** patients into **two classes:** "survived" vs. "died"
- **Mortality score**: the **ML** model's **output probability** of dying

→ Ensemble Learning Mortality score for HCC (ELM-HCC)

- **SHAP analysis** → interpretable, patient-specific risk factors
- **Supervised** clustering in an **embedded SHAP values** space
- Identify **distinct patient subgroups** with unique risk profiles



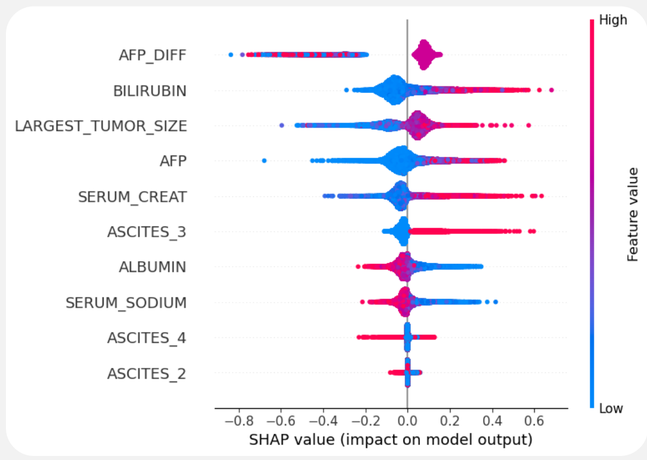
3) RESULTS

COMPARATIVE PERFORMANCE ANALYSIS

	AUROC	SENSITIVITY	SPECIFICITY
AFP score	0.592	35.61	81.99
Child Pugh score	0.705	56.65	76.93
ALBI score	0.707	59.31	74.64
MELD	0.734	62.24	76.94
MELD-Na	0.743	63.88	78.39
MELD 3.0	0.750	64.90	77.30
HALT-HCC	0.763	72.17	69.30
Mehta model	0.782	71.78	74.21
ELM-HCC	0.835	77.14	75.64

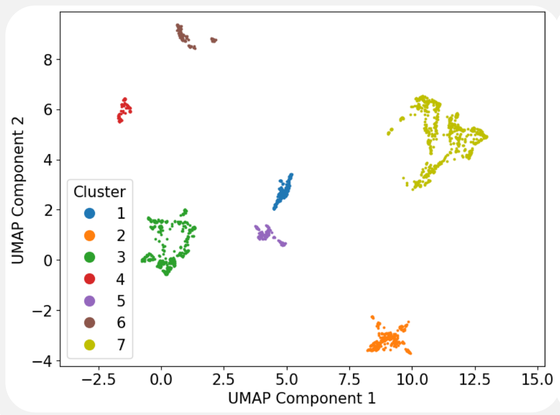
MODEL EXPLAINABILITY

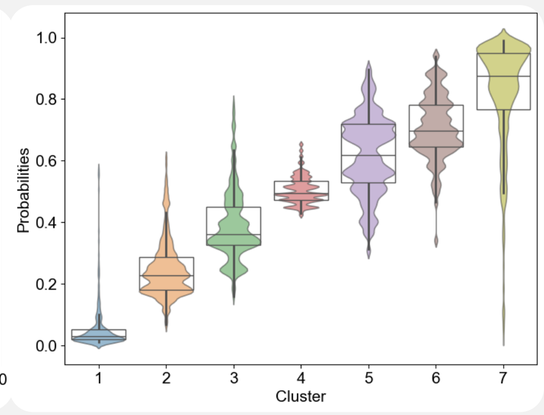
Varying impact of features on model predictions **aligning with well-established clinical knowledge**



RISK STRATIFICATION AND SUBGROUP ANALYSIS

Each cluster shows a **distinct risk profile**
(Kruskal–Wallis & Dunn's test, $p < 0.05$)





Cluster Characterization

- **Cluster 7 (highest risk)**: High bilirubin & creatinine, moderate ascites → severe liver dysfunction
- **Clusters 1–2 (lowest risk)**: Small tumors, low AFP, preserved liver function (low bilirubin/creatinine, no ascites)
- **Clusters 3–4**: Similar to 1–2 but with larger tumors & higher AFP
- **Cluster 5**: Low albumin, sodium, AFP → transitional liver failure
- **Cluster 6**: Highest AFP → aggressive tumor-driven profile

Different pathways to mortality (liver failure vs. aggressive tumor)

UNOS*: Supplied by the United Network for Organ Sharing (UNOS) as the contractor for the Organ Procurement and Transplantation Network (OPTN) in the USA.

Reference

Halimi, A., et al. "Explainable Mortality Prediction for Liver Transplant Candidates with Hepatocellular Carcinoma: A Supervised Clustering Approach." Health Data Science, 6 (2026): 0295.

PhD Students' Day
February 2026

Contact abdelghani.halimi@telecom-sudparis.eu www.telecom-sudparis.eu