

LiverMatch

Modeling Acceptance Behavior to Optimize Liver Allocation

[Proposal](#) [Midterm](#) [Final](#)

Project Proposal

Background

Literature Review

Center variability in liver offer acceptance for the highest priority patients varies from 15.7% to 58.1% [2]. This reluctance significantly impacts patient survival, as a 5% absolute decrease in a center's adjusted acceptance rate correlates with a 27% increased odds of waitlist mortality [1]. Institutional rejection patterns are so varied that candidates are up to three times as likely to receive frequently declined grafts from older donors by listing elsewhere [3][4]. Acceptance practices and the donor-to-candidate ratio independently cause geographic disparities in the median allocated MELD score. To address utilization issues, the OPTN implemented the risk-adjusted offer acceptance rate ratio performance metric, which relies on predictive modeling using donor characteristics.

Dataset

The primary datasets for LiverMatch are from the [Scientific Registry of Transplant Recipients \(SRTR\) Database](#) and the [Organ Procurement Transplantation Network \(OPTN\)](#).

The SRTR dataset is the in process of being requested and the OPTN data can be found in the repository.

Problem Definition

Problem

The process of liver transplantation involves a sequence of offers for each donor organ, but transplant centers vary in their acceptance thresholds [2]. Some organs are accepted at the top of the match run, while others cascade through multiple centers before being selected. This creates variability in allocation efficiency, and any delay in the process can result in viable organs expiring or reaching patients later than optimal, increasing patient mortality.

Motivation

Eliminating this inefficiency in the liver allocation process reduces waitlist times, maximizes utility of donor organs, and increases allograft survival rates. By introducing a standardized prediction model, the project shortens waiting times for scarce organs and can reduce the disparity in waiting times and geographical variability between different transplant centers [2][3]. Streamlining the selection process also reduces operational costs and logistical burden on hospitals and organ procurement organizations.

Methods

Data Preprocessing

- **Dimensionality Reduction:** Principal Component Analysis (PCA) will be applied to reduce the high-dimensional medical dataset ($\approx 3,400$ features), retaining components that explain most variance while minimizing noise and redundancy.
- **Categorical Encoding:** One-hot, ordinal, and target encoding will be used to transform categorical variables into machine-learning-ready representations.
 - **Donor Snapshot:** Encodes organ quality and offer context (age, BMI, steatosis, DRI, ischemia times, sequence number).

- **Recipient Snapshot:** Captures clinical urgency and compatibility (MELD score, blood type match, size match, distance, match run position).
- **Medical Center Snapshot:** Reflects structural and behavioral features (location, historical acceptance rates, median sequence numbers).
- **Feature Scaling:** Continuous variables will be standardized or log-transformed as appropriate to normalize distributions and stabilize model training.

ML Algorithms

We have several features we want to look at:

- **Sequence Number:** the position at which an organ is first accepted (e.g. if a liver is offered to 27 candidates before being accepted, its sequence number is 27).
- **Early Acceptance:** acceptance within the first decile of the national/regional distribution (e.g., sequence numbers 1–10). Represents organs that are highly desirable or broadly acceptable.
- **Late Acceptance:** acceptance beyond the median of the distribution (e.g., sequence numbers >50). Represents organs that are repeatedly declined before finally being matched.
- **Snapshot:** a structured summary of donor liver characteristics, including continuous (age, BMI) and categorical variables (biopsy steatosis <10%, 11–30%, >30%).

We want to categorize snapshots by looking into features such as steatosis, donor age, liver quality measures, etc.. There are several models we will use:

- **Clustering:** K-means, Hierarchical, DBSCAN, and Gaussian Mixture Models for grouping snapshots into distinct profiles.
- **Classification models:** logistic regression, gradient boosting for predicting the probability that a donor organ is accepted early or late.
- **Fixed effects for centers/regions:** indicator variables to capture differences in acceptance behavior attributable to local practice patterns rather than donor attributes alone.

Results and Discussion

Quantitative Metrics

- AUC - Classification Performance
- F1 Score - Classification Performance
- Early Acceptance - Sequence Number ≤ 10
- Moderate Acceptance - $10 < \text{Sequence Number} < 50$
- Late Acceptance - Sequence Number ≥ 50

Project Goals

The project aims to streamline the liver selection process by:

- Identifying patterns in liver offer sequences and acceptance positions.
- Quantifying regional and center-level differences in acceptance behavior.
- Building donor organ profiles to characterize which organs are accepted early vs. late.

Expected Results

The model should produce a set of characteristics that maximize the success of the transplant, called the **ideal donor**. We will map the real-world donors and find the donor(s) whose characteristics most closely correspond to the ideal donor's characteristics through proximity analysis.

Sustainability and Ethical Considerations

We will ensure that the decisions made by the model are transparent and provide equitable access to organs, minimizing biases that can cause unfairness. We will also protect sensitive data of donors and recipients and adhere to the **Health Insurance Portability and Accountability Act of 1996** by de-identifying data where appropriate.

Gantt Chart

See our Gantt Chart [here](#).

Contributions

Name	Contribution
Thinh	Problem Definition, Formatting, Evaluation Methods
Arjun	Data Preprocessing, Expected Results
Alex	Problem Definition, Project Goals, GitHub Website
Jordan	Data Analysis, Data Preprocessing, and Feature Engineering
Connor	Literature Review, References, Metrics

References

- [1] C. E. Haugen et al., "Offer Acceptance Patterns for Liver Donors Aged 70 and Older," Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society, vol. 28, no. 4, pp. 571–580, Apr. 2022, doi: <https://doi.org/10.1002/lt.26309>.
- [2] D. S. Goldberg et al., "Liver transplant center variability in accepting organ offers and its impact on patient survival," Journal of Hepatology, vol. 64, no. 4, pp. 843–851, Apr. 2016, doi: <https://doi.org/10.1016/j.jhep.2015.11.015>.
- [3] A. Wey et al., "Offer acceptance practices and geographic variability in allocation model for end-stage liver disease at transplant," Liver Transplantation, vol. 24, no. 4, pp. 478–487, Mar. 2018, doi: <https://doi.org/10.1002/lt.25010>.
- [4] "New pre-transplant performance metric now in effect, offer acceptance rate ratio - OPTN," Hrsa.gov, 2025. <https://optn.transplant.hrsa.gov/news/new-pre-transplant-performance-metric-now-in-effect-offer-acceptance-rate-ratio/>.

Special Reward Consideration:

Yes, we opt-in for Outstanding Project.

GitHub Pages Link

The link to our project page can be found [here](#).