Poster MSR89

Machine Learning-Based Approach to Discover Subtypes of Light-Chain Amyloidosis (AL) Using US Claims Data

Stuti Dhebar[1], Jeffrey Thompson[2], Pedro Laires[3], Julia Catini[2], Genevieve Lyons[2]

[1]Northeastern University, Boston, MA, USA; [2]Alexion, AstraZeneca Rare Disease Inc., Boston, MA, USA; [3]Alexion, AstraZeneca Rare Disease Inc., Barcelona, Spain

INTRODUCTION

- Light-Chain (AL) Amyloidosis is a rare disease caused by plasma cell dyscrasia, where there is an overproduction of light chain produced from a plasma cell clone, which misfold and aggregate to form amyloid fibrils.[1,2]
- Amyloid fibrils are deposited throughout the body, leading to organ dysfunction and failure, and may ultimately result in death if not treated. [1,2]
- As AL is a systemic disease within a heterogeneous presentation, patients may experience various organ manifestations, but the heart and kidneys are the most impacted in 60-75% of patients.[1,3]
- Machine learning techniques, including clustering, present a novel approach to identify patient subtypes from real-world data by segmenting patients into groups based on similarity metrics, such as the symptoms they exhibit.[4]

OBJECTIVES

- To utilize Machine Learning (ML) techniques to discover patient clusters and subtypes based on the clinical manifestations they experience before receiving a diagnosis for AL.
- To gain insight into frequently co-occurring comorbidities within the heterogeneous patient population.
- To evaluate the association between discovered subtypes and clinical outcomes following diagnosis, as well as healthcare resource utilization (HCRU).

CONCLUSIONS

- Machine Learning can be useful in effectively identifying patient subtypes for AL, which can inform treatment regimens to control symptoms.
- Using an unsupervised clustering method, six patient subtype groupings were identified based on the commonly co-occurring symptoms within each group.
- The subtypes reveal that AL patients within the Severe Cardiac and Cardiac + Renal subtypes manifest severe symptoms, leading to more severe outcomes compared to other groups.
- Unsupervised Machine Learning clustering algorithms are a promising avenue to facilitate a better understanding of the disease manifestations and treatment approaches and could potentially be extended beyond AL to characterize other diseases with heterogeneous patient populations.

REFERENCES

1. Blancas-Mejia LM, et al. Chem Commun (Camb). 2018;54:10664.
2. Al Hamed R, et al. Blood Cancer J. 2021;11:97.
3. Dittrich T, et al. Acta Haematol. 2020;143:388.
4. Carreira-Perpiñán, M. Á., Wang, W. arXiv 2013:1304.6478.

METHODS

STUDY DESIGN

- **Design:** Retrospective, non-interventional cohort study of adult patients with a confirmed diagnosis of AL amyloidosis entered in the IQVIA PharMetrics® Plus US Claims database.
- **Study Period:** January 1, 2016 to September 30, 2022
- **Identification Period:** January 1, 2018 to March 31, 2022
- **Baseline Period:** 24 months (730 days) pre-index
- **Follow-up Period:** 6 months (180 days) post-index
- **Index Date:** Second diagnosis of AL Amyloidosis (ICD-10: E8581)

INCLUSION/EXCLUSION CRITERIA

- Adult patients with ≥2 diagnostic codes of AL Amyloidosis (ICD-10: E8581) during the identification period, aged ≥18 years at index, with ≥24 months continuous coverage pre-index, and ≥6 months follow-up post-index were included.
- Patients with missing age or sex details were excluded.

MANIFESTATIONS OF AL AMYLOIDOSIS

- Presence of manifestations were defined using ICD-10 codes for each patient in the baseline period (24 months pre-index), including the following:
  - **Cardiac Manifestations:** Heart Failure, Dyspnea, Syncope and collapse, Cardiomegaly, Atrial Fibrillation and flutter, Other cardiac arrhythmias, Cardiomyopathy, Pleural effusion, Fatigue, Hypotension
  - **Renal Manifestations:** Nephrotic Syndrome, Proteinuria, Renal disease, Edema
  - **GI/Hepatic Manifestations:** Nausea/Vomit, Diarrhea, Constipation, Abdominal Pain, Hepatomegaly and splenomegaly, Bloating, Early satiety, Weight loss
  - **Neurologic Manifestations:** Peripheral Neuropathy, Autonomic Neuropathy, Numbness/Pain, Erectile dysfunction, Dizziness
  - **Other/Multi-organ Manifestations:** Dysphagia, Carpal Tunnel Syndrome, Purpura, Macroglossia, Submandibular swelling, Jaw claudication, Nail dystrophy, Lack of appetite, Change in taste (and smell)

OUTCOMES ANALYSIS

- Outcomes were defined as Emergency Room (ER) visits and inpatient (IP) stays for each patient during the follow-up period (6 months post-index).
- Cardiac-related ER visits and IP stays were defined as those with a Cardiac manifestation as the primary diagnostic code on the ER/IP claim.

MACHINE LEARNING (ML) ANALYSIS

Feature Engineering: Collected manifestations experienced by patients prior to AL diagnosis, including cardiovascular, renal, hepatic, neurologic, and others.

Train Algorithm: Partitioned patients into clusters based on K-Modes, an unsupervised ML algorithm that accounts for correlation between manifestations and partitions patients into distinct clusters based on similarity of co-occurring manifestations. Patients with manifestations that commonly co-occur were categorized into clusters (or subtypes) based on their full set of manifestations.

Performance Evaluation: Because this is an unsupervised ML problem, there is no concept of true accuracy or prediction. Clusters (subtypes) were evaluated using the silhouette score, which measures intra-cluster similarity and inter-cluster dissimilarity. The optimal number of clusters was selected to maximize the silhouette score.

Patient Subtype Discovery: We described and labelled each subtype based on the prevalent contributing manifestations that patients experience within that cluster.

Outcomes Analysis: For each subtype, we measured the probability of an ER visit or IP stay during the follow-up period and calculated the odds ratio (OR) using logistic regression, adjusting for age, sex, and region.

RESULTS AND INTERPRETATION

- A total of 1276 patients met the inclusion/exclusion criteria
- Over half of the cohort was male (N=736, 58%), and the mean (SD) age was 62.0 (9.8) years
- The highest silhouette score was 0.14 at 6 clusters/subtypes

Figure 1 displays each subtype and the prevalence of each manifestation within that cluster

We discovered the following distinct sub-types of AL Amyloidosis:

Cardiac (N=200, 16%)
Characterized by cardiomyopathy and heart failure without edema

Cardiac + Renal (N=96, 8%)
Characterized by heart failure with CKD, frequently comorbid with neuropathy

Severe Cardiac (N=116, 9%)
Characterized by heart failure with edema and pleural effusion

Non-Specific Manifestations (N=135, 11%)
Characterized by gastrointestinal and hepatic manifestations

Renal (N=253, 20%)
Characterized by proteinuria, chronic kidney disease

Early Mayo Stage (N=476, 37%)
Characterized by a lack of manifestations

FIGURE 1 - CLUSTER DISCOVERY

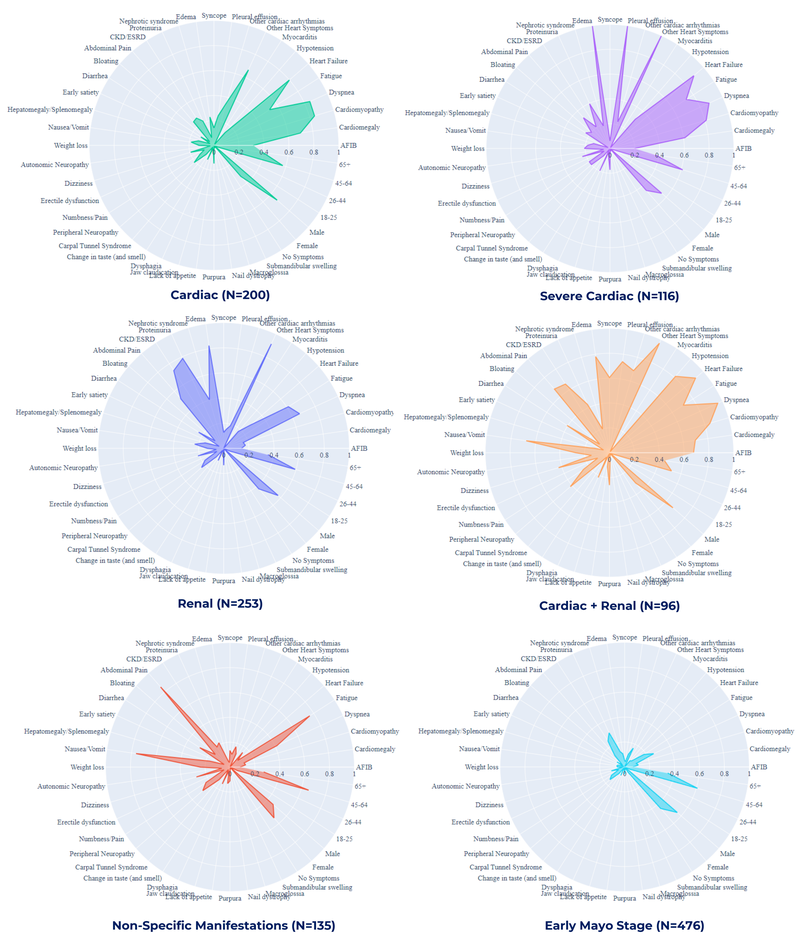

FIGURE 2 - OUTCOMES ANALYSIS

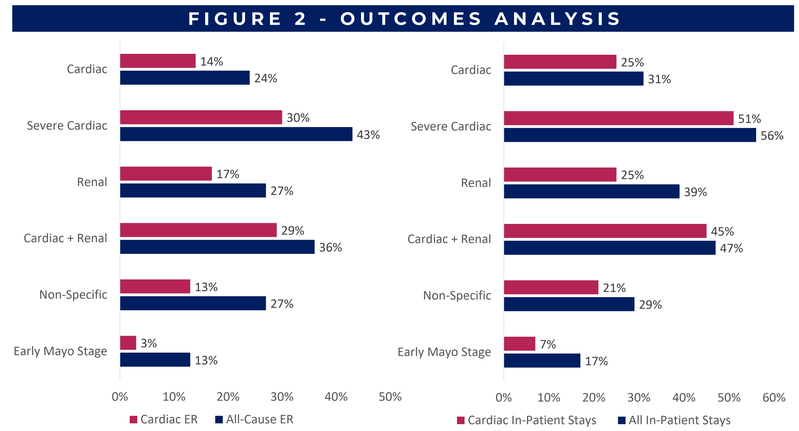

Figure 2 shows the proportion of patients within each discovered AL subtype experiencing emergency room visits (left) and in-patient hospital stays (right) during the 6-month follow-up period.

- In both the **Severe Cardiac** and **Cardiac + Renal** subtypes, ER visits and IP stays are common, with half of the patients having a Cardiac related IP stay (51% and 45%, respectively); and about a third of the patients experience a Cardiac related ER visits (30% and 29%, respectively).
- In the **Renal** and **Cardiac** subtypes, ER visits are still common, but relatively less frequent, with about a quarter of the patients experiencing an ER visit (27% and 24%), with over half of the ER visits related to Cardiac manifestations.
- Cardiac related ER visits and IP stays are somewhat common in the **Renal** and the **Non-Specific** subtypes, with about a quarter of the patients experiencing a Cardiac IP stay, despite the patients not exhibiting Cardiac manifestations at baseline.
- The lowest HCRU was observed for the **Early Mayo Stage** subtype.

FIGURE 3 - STATISTICAL ANALYSIS

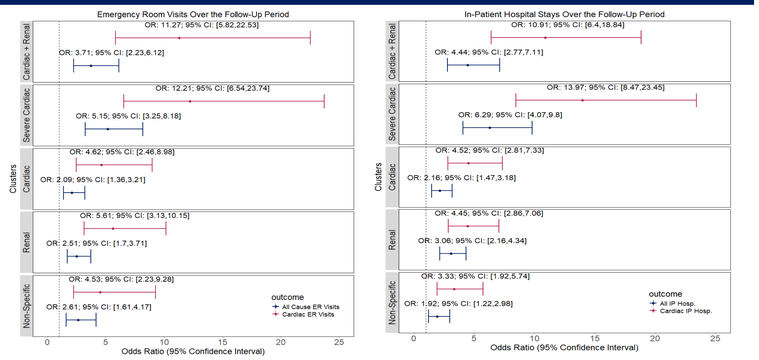

Figure 3 displays the adjusted odds ratios for emergency room visits (left) and in-patient hospital stays (right), adjusted for age, sex, and region, with the Early Mayo Stage subtype as the referent group.

- The probability of Cardiac ER visits were 11x - 12x higher in the **Severe Cardiac** and **Cardiac + Renal** subtypes as compared to the **Early Mayo Stage** subtype.
- Similarly, patients in these subtypes exhibit 10x - 13x increased probability of experiencing Cardiac IP hospital stays during the follow-up period.
- The probability of a Cardiac ER visit were 4x - 6x higher in the **Cardiac**, **Renal**, and **Non-Specific** subtypes as compared to the Early Mayo Stage referent; with similar results of IP stays.