MEDINFO 2023 — The Future Is Accessible
J. Bichel-Findlay et al. (Eds.)

doi:10.3233/SHTI230968

Temporal Phenotyping for End-Stage Renal Disease Using Longitudinal Electronic Health Records

Shengqiang CHI[a], Feng WANG[a], Xueyao LI[a], Minghong XU[a], and Jingsong LI[a,b,1]
[a] *Research Center for Healthcare Data Science, Zhejiang Lab, Hangzhou, China*
[b] *Engineering Research Center of EMR and Intelligent Expert Systems, Ministry of Education, College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China*

Abstract. End Stage Renal Disease (ESRD) is a highly heterogeneous disease with significant differences in prevalence, mortality, complications, and treatment modalities across age, sex, race, and ethnicity. An improved knowledge of disease characteristics results from the use of a data-driven phenotypic classification strategy to identify patients of different subtypes and expose the clinical traits of different subtypes. This study used topic models and process mining techniques to perform subtyping of ESRD patients on hemodialysis based on real-world longitudinal electronic health record data. The mined subtypes are interpretable and clinically significant, and they can reflect differences in the progression of the disease state and clinical outcomes.

Keywords. phenotype mining; end-stage renal disease; disease subtype; process mining; topic model

1. Introduction

According to the 2021 Annual Data Report of the United States Renal Data System, the number of patients with end-stage renal disease (ESRD) is increasing year by year. The total number of patients increased by 41.0% over the past decade to 809,103[1]. ESRD is a highly heterogeneous disease. The prevalence, treatments, complications, and mortality of ESRD vary depending on demographics, etiology, clinical manifestations and disease progression[1-3]. In this situation, the benefit of using a single patient characteristic to make treatment decisions and prognostic analysis is restricted.

In recent studies on the subtyping of ESRD, Wang et al[4] used the blood pressure fluctuation patterns during dialysis to classify patients into 5 subtypes. Patients in the subtype with a decreasing systolic blood pressure (SBP) during dialysis had a better prognosis than those in the subtype with an increasing SBP. Lioulios et al[5] identified different immune subgroups in ESRD patients and defined the immune patterns of each subgroup by analyzing immune expression data. These studies were based on a single feature of multiple visits or the initial static data, thus could not fully utilize the long-term longitudinal electronic health record (EHR) data of ESRD patients.

[1] Corresponding Author: Jingsong Li, e-mail: ljs@zju.edu.cn.

This study combined the topic model (TM) and process mining techniques to mine the treatment processes, so as to study the subtypes of ESRD. The main contributions are as follows: The TM was used to transform the low-level, fine-grained data into high-level, coarse-grained data, so as to summarize the clinical events into a single visit label; Visit labels were used in clinical process mining to enable process mining methods to handle the coexistence of multi-granularity information such as events within a single visit and events between multiple visits in the longitudinal EHR data; The temporal phenotypes and subtypes of ESRD were extracted from the longitudinal HER data. Differences in clinical outcomes of different subtypes were evaluated.

2. Methods

2.1. Data sources

Patients with ESRD who started receiving hemodialysis at the First Affiliated Hospital of Zhejiang University between 2000 and 2016 and whose follow-up time was at least 90 days were selected. In this study, the initial hemodialysis was used as the starting point of observation, and the cutoff time of follow-up is August 20, 2017. Events of interest included death and kidney transplantation. The clinical events of interest included diagnoses, laboratory tests, medical examinations, procedures and medications. Among them, laboratory tests focused only on abnormal test results.

2.2. Clinical process mining, subtype classification and evaluation

First, we used TMs to preprocess the clinical process data by treating the clinical events in a single visit as documents and labeling visits. We used the latent Dirichlet allocation algorithm to provide topic probability distributions for documents. The topic with the highest probability was selected as the visit's topic. To determine the optimal number of topics K, we used four metrics of consistency, redundancy, importance, and perplexity[6].

Secondly, different treatment processes of ESRD patients were found by mining the extracted visit labels. The visit labels obtained from the TM were used to construct event logs. Each visit was represented as an event log, whose event name was the label of that visit and timestamp reflected the duration of that visit. Consecutive visits with the same event names were merged. The duration of the new event was equal to the time period covered by two visits. The clinical process mining algorithm is as follows:

Step 1: event logs were converted into an event matrix with a dimension of $N \times M$, where N is the number of patients and M is the maximum number of patient visits. Each row in the event matrix represents a treatment sequence of a patient's visits. Each column represents the label of a visit, and patients with less than M visits are filled with 0.

Step 2: the support of all starting events (the first column in the event matrix) was calculated and only those events with the support greater than a threshold were selected as starting events for the clinical process. The support is defined as the number of patients who experienced a specific sequence of events divided by the total number of patients.

The algorithm adds subsequent events to the clinical process by iterating the support computation on the events that follow the initial set, until no more frequent events can be extracted or the maximum length is reached. The quality of process mining results is

evaluated using four metrics: fitness, accuracy, generalization, and simplicity[7]. The quality score (QS) of the model is defined as the average of these four metrics.

Finally, patients were classified by clinical process and subtypes were characterized using patients' clinical manifestation and prognosis. To determine whether there are clinically interpretable differences between subtypes, we first determined whether statistical differences exist in demographics. The Mann-Whitney test was used to examine the differences between the medians of continuous variables and the χ^2 test to examine the differences between the distributions of categorical variables. Then, the differences in important medications and complications among different subtypes were determined. In addition, Kaplan-Meier survival curves were plotted, and Log-Rank test was used to assess the differences in survival curves of patients with different subtypes.

3. Results

3.1. Data pre-processing and cohort data statistics

Patients with less than 3 visits in the follow-up as well as visits with only one clinical event were excluded. The following analysis was then performed on data of 73,090 visits from 2,010 patients. The median number of visits per patient was 8 (interquartile range (IQR): 4-39). The median number of clinical events per visit was 3 (IQR: 2-5).

3.2. Subtype mining and evaluation

We conducted a grid search on hyperparameters of the TM in a range of K from 2 to 10. When K is equal to 4, the model achieved a good trade-off between high consistency and importance and low redundancy and perplexity. The K was set to 4. The 4 generated topics as well as topic keywords and their probabilities were shown in Table 1.

Table 1 Topic keywords and their probability.

Topic	Topic keywords and their probabilities
1	***Multiple complications and their treatments (17 keywords)*** ***Keywords:*** Hypertension (0.095), anti-peptic ulcer (0.086), antimicrobials (0.077), anticoagulants (0.053), anemia (0.050), amino acids (0.045), antacidosis (0.038), sodium channel blockers (0.036), anti-platelet drugs (0.036), hyperparathyroidism (0.035), calcium supplements (0.033), mucopolysaccharide polysulfate cream (0.033), sedative-hypnotics (0.030), laxatives (0.030), anti-anginal (0.025), nitrates (0.024), diuretics (0.024)
2	***Hemodialysis treatment, abnormal bone metabolism, anemia and their treatments (7 keywords)*** ***Keywords:*** Abnormal bone mineral metabolism (0.180), calcium supplements (0.172), anti-anemia drugs (0.113), levocarnitine (0.109), anti-platelet drugs (0.064), diabetes (0.06), amino acids (0.052)
3	***Hemodialysis treatment, anemia, cardiovascular diseases and their treatments (8 keywords)*** ***Keywords:*** anti-anaemic agents (0.148), anti-anginal (0.147), levocarnitine (0.107), calcium channel blockers (0.100), anti-abnormal bone mineral metabolism (0.086), vitamins (0.063), beta-blockers (0.061), angiotensin II receptor blockers (0.049)
4	***Abnormal laboratory test results (11 keywords)*** ***Keywords:*** routine blood test (0.120), routine kidney function test (0.110), routine urine test (0.102), routine potassium, sodium, chloride, calcium, magnesium and phosphorus measurement (0.098), routine blood lipid test (0.072), routine liver function test (0.066), bone metabolism marker measurement (0.053), glucose measurement (0.048), routine myocardial enzyme profile test (0.038), routine coagulation test (0.035), ultrasensitive C-reactive protein measurement (0.026)

We performed a grid search on hyperparameters of the process mining model in the range of minimum support from 0.01 to 0.05 and maximum length from 2 to 10. Due of

the model's best QS, we decided to set the minimum support at 0.03 and the maximum path length at 4. The excavated treatment processes were shown in Figure 1.

We reduced the number of subgroups by grouping similar event sequences. Six subtypes were obtained in the end, as shown in Table 2. There were fewer differences between subtypes 2, 3 and 6, while more differences between the other groups. Thus, we grouped subtypes 2, 3 and 6 into one group, and grouped the remaining subtypes into the other group. We found that the latter group of patients had abnormal laboratory test results (ALTRs) earlier in the course of the treatment (named as abnormal group) compared to the former group (named as normal group). The median time to the first occurrence of ALTRs is 1.74 years (IQR: 0.50-4.01 years) in the normal group and 0.96 years in the abnormal group (IQR: 0.13-3.89 years, $p<0.0001$). Besides, there were statistical differences in demographics, lifestyles, disease characteristics and comorbidities in both subtyping methods. The survival curves were shown in Figure 2. The mortality in the abnormal group was higher than that in the normal group, and the probability of kidney transplantation was lower than that in the normal group.

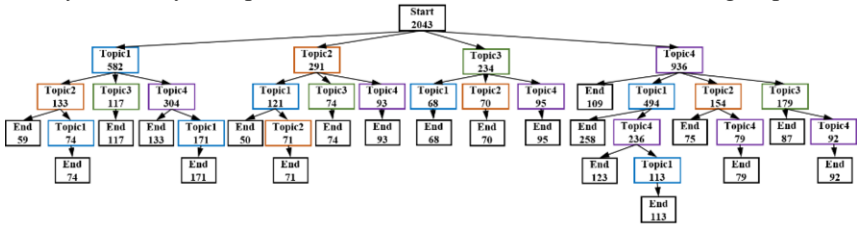

Figure 1 Treatment processes.

Table 2 End-stage renal disease subtypes and their characteristics based on the process mining model.

Subtype	Treatment process	Subtype description
Subtype 1	Topic 2 alternates with Topic 4	Abnormal bone metabolism and its treatments, ALTRs
Subtype 2	Topic 2 alternates with Topic 3	Abnormal bone metabolism and its treatments, cardiovascular disease and its treatments
Subtype 3	Topic 1 alternates with Topic 2	Abnormal bone metabolism and its treatments, diagnosis and treatment of multiple complications
Subtype 4	Topic 3 alternates with Topic 4	Cardiovascular diseases and their treatments, ALTRs
Subtype 5	Topic 1 alternates with Topic 4, Topic 4	ALTRs, and/or diagnosis and treatments of multiple complications
Subtype 6	Topic 1 alternates with Topic 3	Cardiovascular diseases and their treatments, diagnosis and treatment of multiple complications

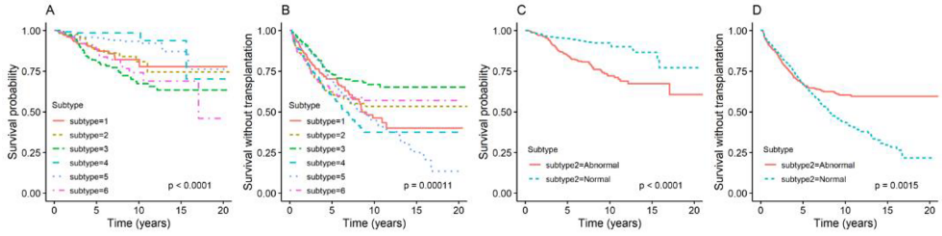

Figure 2. Kaplan-Meier survival curves.

4. Discussion

The results of the phenotyping of ESRD patients demonstrated the effectiveness of the method, and the findings are interpretable and clinically significant. In particular, the

differences in the time to the first occurrence of ALTRs between two subtypes obtained by coarse-grained subtyping, may account for the differences in clinical outcomes. In the abnormal group, patients' complete blood count, urine routine, renal function, liver function, myocardial enzyme and other indicators show abnormalities earlier, indicating that this group had a poor ability to control the disease. As a result, this group of patients had higher mortality, worse prognosis, and a lower probability of kidney transplantation.

These subtypes may improve clinical decision-making and disease management after they have undergone clinically validation. In particular, more refined clinical assessments and interventions can be used earlier to improve a patient's prognosis. However, this study is still in its early stage, and further researches are needed to improve and produce more interpretable, clinically experienced, and fine-grained characteristics of visits and subtypes.

5. Conclusions

This study aimed to identify subtypes of ESRD patients on hemodialysis using longitudinal EHR data. We utilized TMs and process mining techniques to classify patients into meaningful subgroups based on their multiple visit data, and presented interpretable connections between the disease status and clinical outcomes.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 82001925), the Zhejiang Provincial Natural Science Foundation of China (No. LQ21H180002 and No. LQ21F030006), the Key Research Project of Zhejiang Lab (No. 2022ND0AC01).

References

[1] U.S.R.D. System, 2021 USRDS Annual Data Report: Epidemiology of kidney disease in the United States, in, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD

[2] B. O and L. SL, End-Stage Renal Disease. [Updated 2021 Sep 16]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK499861/,

[3] M.M. O'Shaughnessy, M.E. Montez-Rath, R.A. Lafayette, and W.C. Winkelmayer, Patient characteristics and outcomes by GN subtype in ESRD, Clinical Journal of the American Society of Nephrology 10 (2015), 1170-1178,

[4] F. Wang, J.-y. Zhou, Y. Tian, Y. Wang, P. Zhang, J.-h. Chen, and J.-s. Li, Intradialytic blood pressure pattern recognition based on density peak clustering, Journal of biomedical informatics 83 (2018), 33-39,

[5] G. Lioulios, A. Fylaktou, A. Xochelli, E. Sampani, I. Tsouchnikas, P. Giamalis, D.-V. Daikidou, V. Nikolaidou, A. Papagianni, I. Theodorou, and M. Stangou, Clustering of End Stage Renal Disease Patients by Dimensionality Reduction Algorithms According to Lymphocyte Senescence Markers, 13 (2022), doi: 10.3389/fimmu.2022.841031.

[6] L. Chiudinelli, A. Dagliati, V. Tibollo, S. Albasini, N. Geifman, N. Peek, J.H. Holmes, F. Corsi, R. Bellazzi, and L. Sacchi, Mining post-surgical care processes in breast cancer patients, Artificial Intelligence in Medicine 105 (2020), 101855, doi: https://doi.org/10.1016/j.artmed.2020.101855.

[7] A. Berti, S.J. Van Zelst, and W. van der Aalst, Process mining for python (PM4Py): bridging the gap between process-and data science, arXiv preprint arXiv:.06169 (2019),