Process mining for exploring treatment patterns in Chronic Lymphocytic Leukemia (CLL) in a real world oncology database

James Black[1]; Felipe Castro[1]; Linda Lundberg[2]; Michael Marino[3], Rudi Niks[4], Anne Rozinat[4]

[1]PHC Data Science, Hoffman-La Roche, Basel, Switzerland [2]Clinical Science, Hoffman-La Roche, Basel, Switzerland

[3]Marketing Science, Genentech, South San Francisco, United States [4]Fluxicon, Eindhoven, The Netherlands

Background

- Treatment patterns in oncology constantly evolve, while guidelines offer diverse options for increasingly personalised patient subsets. Defining normative treatment patterns is difficult. Process mining is the application of statistical and graphical methods to extract pathways from event logs and provides a novel approach in pharmacoepidemiology.

Evaluation plan

- Using line of therapy information from the Flatiron Health, electronic health record (EHR) derived oncology database, we evaluated the use of process mining methodology to aid in the understanding of CLL treatment patterns. Three defined use cases were explored using process mining methods by three different stakeholders (Marketing Science, Clinical Science and a Pharmaco-epi):
  1) using interactive analysis to identify data collection issues
  2) understanding complex patterns with high variance in treatment pathways within specific populations and
  3) understanding observed treatment flows leading to specific treatments or outcomes.

Results

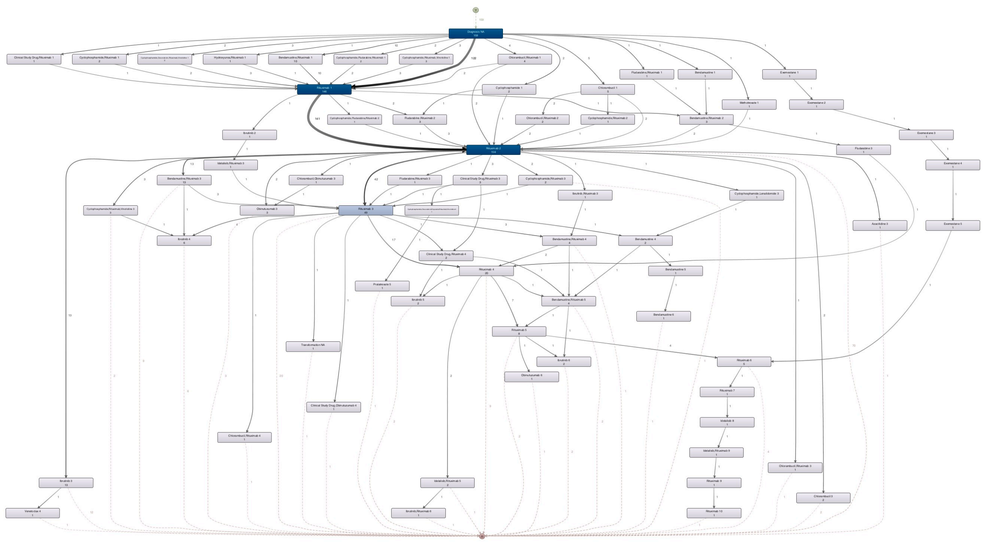
- Patients with CLL and SLL (N= 2,597; 81% CLL/19% SLL) ≥18 years old were selected from the Flatiron Health EHR derived database. Patients were included if they entered the Flatiron Health network without prior lines and first-line treatment was between 2011 and 2016.
- Using process mining methods to explore the data, we were able to quickly identify an abnormal treatment pattern (Figure 1), enabling us to account for it in later analyses. We also demonstrated the use of these methods to quickly gain an overview of treatment flows that are non-conformant with NCCN guidelines for 1L (Figure 2).

Discussion – Stakeholder experiences

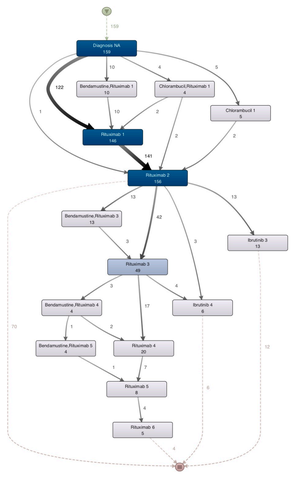
- In Pharmacoepidemiology, treatment pattern analyses are focused on out of sample estimation of what the complete population of interest experiences. Treatment pattern analyses routinely present only frequencies by line, or with 1st order transitions. Process mining provides a toolset to understand the full sequence of treatments within an observed sample.
- Process mining methods excel at identifying patterns within an observed set of data, which makes them susceptible to non-normative aspects of the input data. Without being to provide measure of variance, the ability to gain inference beyond the observed sample is limited.
- Epidemiology, clinical and commercial stakeholders found the most benefit came from being able to visualise and explore complex treatment patterns. Process mining methods and tools were seen as a tool to better understand complex process data, before testing hypotheses using traditional pharmacoepidemiologic methods.
- A potential limitation for commercial stakeholders is that while process mining tools excel at surfacing common, but unexpected treatment patterns, with the high heterogeneity present in real world settings it is difficult to identify uncommon patterns that may be of interest without a prior assumptions on where within the mapped process to filter to.
- Figure 1 shows how the visual methods used in process mining identified a pattern of interest in the data that was highly prevalent and did not conform to expectations.
- In Figure 2 a simple set of filters on the observed processes helped bring an understanding about general trends present in the data. This example highlights how plots produced via process mining workflows can help to quickly provide insights on new patterns in treatment that are emerging, and may help identify new treatment paradigms taking hold in the real world or be used as a reference in discussions with oncologists about what is the standard of care.

Figure 1: Steps taken to gain understanding about R-mono use that led to the unexpected discovery of R-mono re-treatment.

Step 1: Filter to R-mono exposed patients and view full set of treatment patterns experienced by the population



Step 2: Iteratively explore pathways filtering on factors like frequency



Step 3: Filter to identified data pattern of abnormal retreatment rates, highlighting potential data quality issue

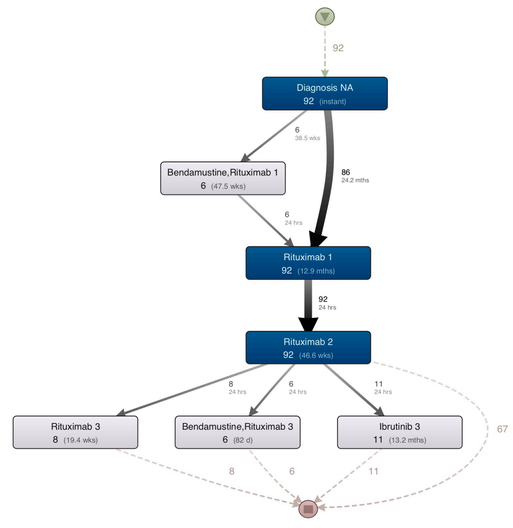
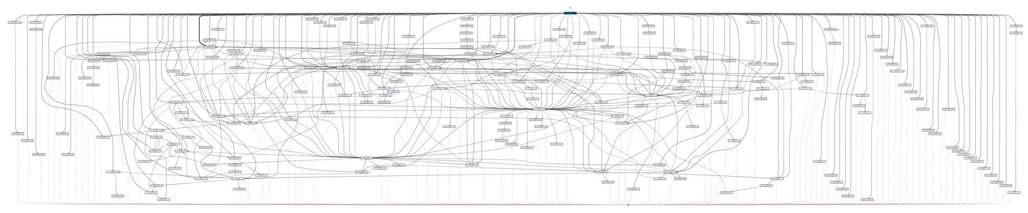
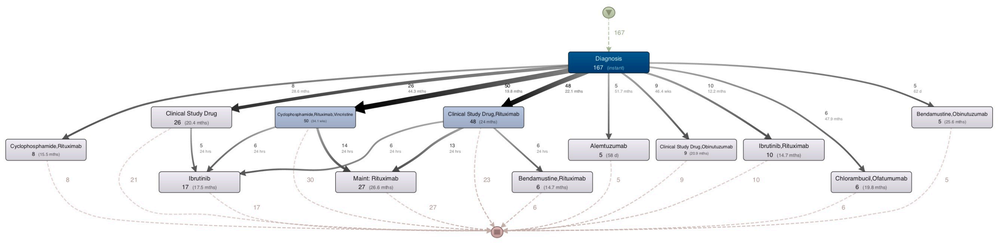


Figure 2: Filtering provides quick insights on dominant patterns present

Observed treatment patterns where patient had a 1L not specified in NCCN[1] guidelines.



Dominant patterns visualized by restricting to pathways experienced by 5 patients.



1L, First line of treatment;

Conclusions

- Application of process mining provided insights on complicated disease flows.
- Process mining techniques provide insight into the patterns present in observed data.
- The greatest benefit came from introducing process mining as an intermediate step to explore and validate a-priori assumptions about potential patterns present in the data, rather than to test hypotheses or estimate treatment prevalences.
- In addition to the statistical methodology developed within the process mining literature, process mining software provides mature and easy to use tools that can enable interactive exploration of complex longitudinal event data.
- Process mining is a diverse field, and a landscaping of methods to understand highly heterogenous data would be beneficial.

Acknowledgements

Dr Adrian Cassidy for reviewing drafts.

Conflicts of interest

JB, FC, LL, and MM are employees of the Roche Group. RN and AR are employed by Fluxicon. Flatiron Health is an independent subsidiary of the Roche Group.

References

1. NCCN Clinical Practice Guidelines in Oncology. (2016) https://doi.org/10.1007/978-3-540-85772-3

34th International Conference on Pharmacoepidemiology and Therapeutic Risk Management
Prague, Czech Republic | August 22 – August 26, 2018