Introducing ARTEMIS: Advanced Regimen Detection Using an Adapted Smith-Waterman Algorithm.

Asieh Golozar[1,2], Laurence Lawrence-Archer[1], Travis Zack[3], Jeremy L. Warner[4], Christian Reich[2,5]

[1]Odysseus Data Services, MA, USA [2]OHDSI Center at the Roux Institute, Northeastern University, Boston, MA, USA [3]University of California San Francisco, San Francisco, CA, USA [4]Brown University, Providence, RI, USA, [5]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands

Introduction

Chemotherapy is usually administered in complex cocktails called regimens, with varying antineoplastic agents, their dosing, cycle lengths, cycle number, and scheduling. Assessing the safety and effectiveness of chemotherapy regimens in the real-world requires the precise knowledge of their administration. However, this is often not consistently or sufficiently captured in observational databases, while the individual components, their timing, and their dosing is often known as part of drug exposure or administration data. The inference of regimens from the basic drug data is a very non-trivial task for several reasons: a large number (1500+) of established regimens, overlapping regimens, lack of standardization of regimen information, imprecise following of the protocol due to routine circumstances at the point of care, and the occurrence of side effects as well as non-standard treatment choices.

In 2019, the OHDSI Oncology WG developed a rule-based algorithm called THEANO to address this gap. THEANO identifies fixed combinations of agents administered within a fixed time interval (30 days) and matches them to HemOnc, a reference knowledge base of regimens. While Theano performs satisfactorily in detecting common treatment regimens, it has limitations: (i) handling different schedules of the same drug combinations, different dosing, different timing, and cycle lengths, (ii) deviations from the protocols, (iii) extracting dosing information and (iv) missing procedures such as radiotherapy or stem cell transplantation. Here, we introduce ARTEMIS, as the next generation regimen detection algorithm that addresses the limitations of THEANO.

Method

ARTEMIS incorporates an adapted Smith-Waterman (SW) algorithm, a standard bioinformatics approach for determining homology between nucleic acid or protein sequences. SW guarantees optimal alignment between two sequences of events by using recursive dynamic programming and scoring matrices. The standard SW, however, does not account for temporal variations such as missing events, treatment delays, protocol deviation, and relative time between events. ARTEMIS incorporates relative timing information in the alignment process to address this limitation. It identifies optimal matches between patients' longitudinal drug records and HemOnc regimens for a given scoring scheme, gap penalty, and temporal penalty. Cycles are

called when the algorithm detects an alignment score of > 0.6. The identified cycles are then used to assemble regimens and, out of these, lines of therapy (Figure 1).

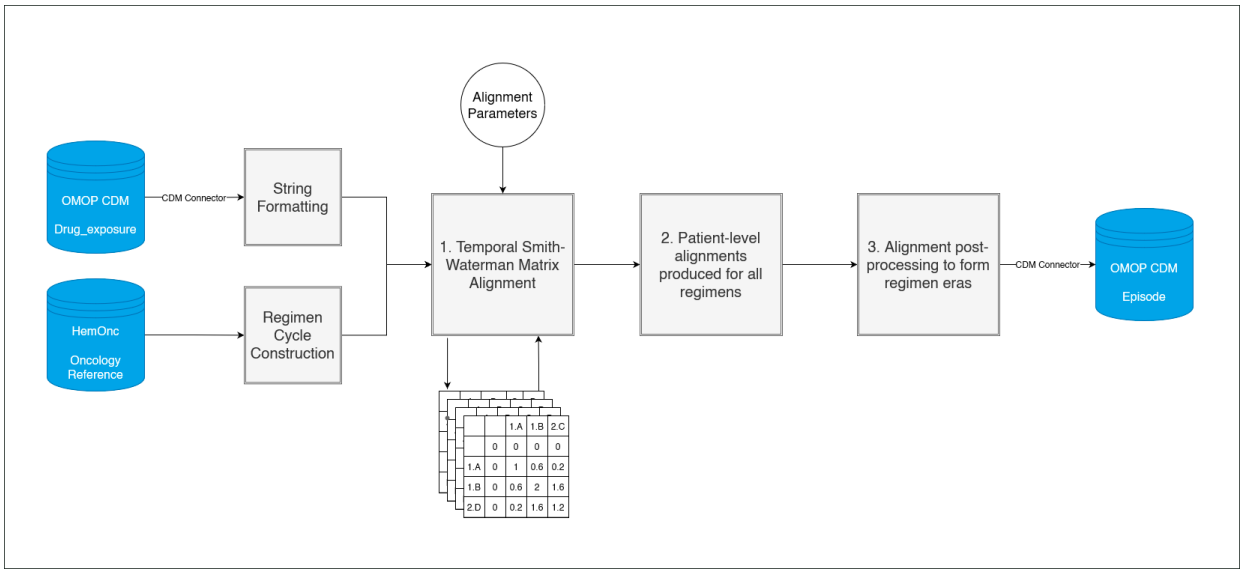


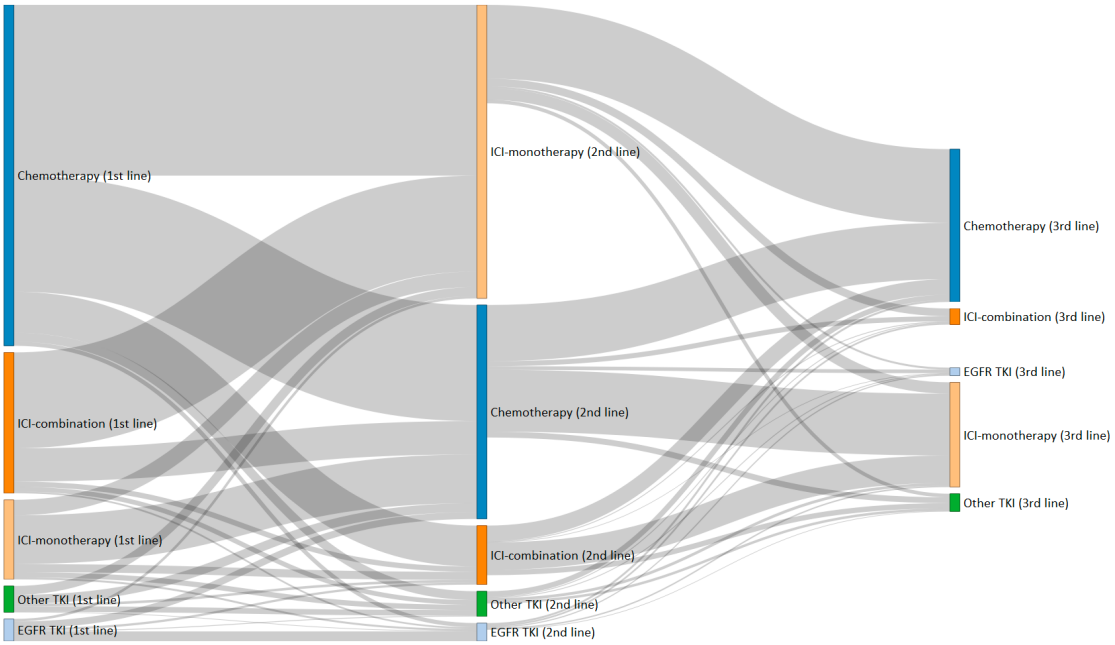
Figure 1. ARTEMIS schematic workflow

In its current version, ARTEMIS does not check dosing information or regimens that include non-drug treatments.

We are planning to test the validity of these matches in databases which have records of both the regimen and individual drug administration information.

Results

We applied ARTEMIS to a cohort of 3,297 metastatic lung cancer patients from the IQVIA PharMetrics Plus database to assess its ability to detect treatment regimens and lines of therapy. Here, we present the initial results. Regimens were identified in 3081 (81%) of patients with alignment score above 0.6. Distribution of treatment regimens in the first and second line of treatment and the treatment pattern if metastatic lung cancer patients are depicted in Figure 2.



1L: First line; 2L: Second line; 3L: Third line; ICI: immune checkpoint inhibitors; TKI: Tyrosine kinase inhibitor.

Figure 2. Treatment patterns in metastatic lung cancer patients

Conclusion

ARTEMIS offers a promising solution to the challenges of accurately identifying chemotherapy regimens from observational databases. Further evaluation is underway to assess the performance of this algorithm. Final results will be presented at the OHDSI symposium.