RHEA: Real-world observational Health data Exploration Application

Soobeen Seol[1], Jimyung Park[1], Chungsoo Kim[1], Byungjin Choi[2] , Dong Yun Lee[2], Rae Woong Park[1,2]

[1]Department of Biomedical Sciences, Ajou University Graduate School of Medicine [2]Department of Biomedical Informatics, Ajou University School of Medicine

Background

- Patient data is fragmented in electronic health records (EHRs).
- We developed a framework named RHEA (Real-world observational Health data Exploration Application) that can profile longitudinal patient data.
- Using RHEA, we visualized the profiled data in patients with cancer.

Methods

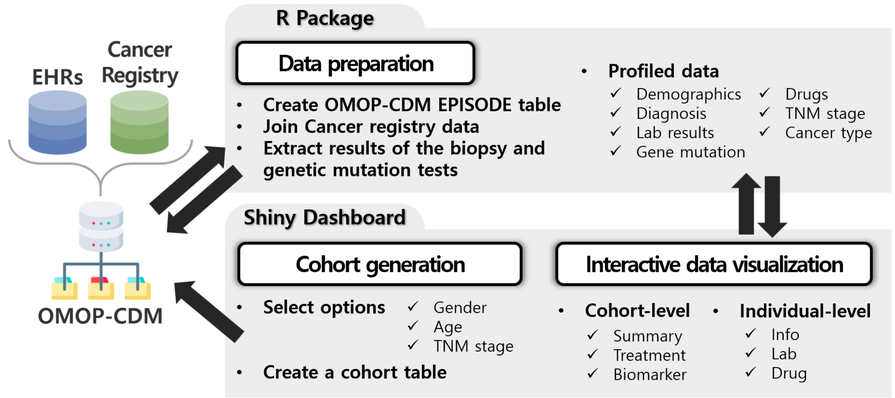

Figure 1. The schematic workflow of the framework. EHRs: electronic health records; OMOP-CDM: Observational Medical Outcome Partnership-Common Data Model.

1. Data source

- Ajou University School of Medicine (AUSOM), from January 1994 to July 2021.
- As a proof-of-concept study, we defined a target cohort as the patients with colorectal cancer (CRC), which showed a rapidly increasing health problem.

2. Data preparation

- We updated the TRACER (Tool for Regimen-level Abstraction of Chemotherapy Episode Records) with 66 newly introduced chemotherapy.
- We generated EPISODE table using the TRACER package.
- We extracted the TNM stage from cancer registry data and insert into the MEASUREMENT table.
- We extracted biopsy and genetic mutation test results (microsatellite instability, KRAS mutation, NRAS mutation, and BRAF mutation) from the pathology reports using regular expressions.
- Profiling was performed to reflect on the visualization dashboard based on the previously extracted data, and the profiled data included demographics, diagnoses, laboratory tests, medications, procedures, and TNM stages.

3. RHEA function

- The cohort-level visualization provides overall information about patients with cancer.
- The individual-level visualization provides detailed information about each patient.
- The function 'Cohort Generation' that generate subset of original cohort by gender, age and TNM stage.

Results

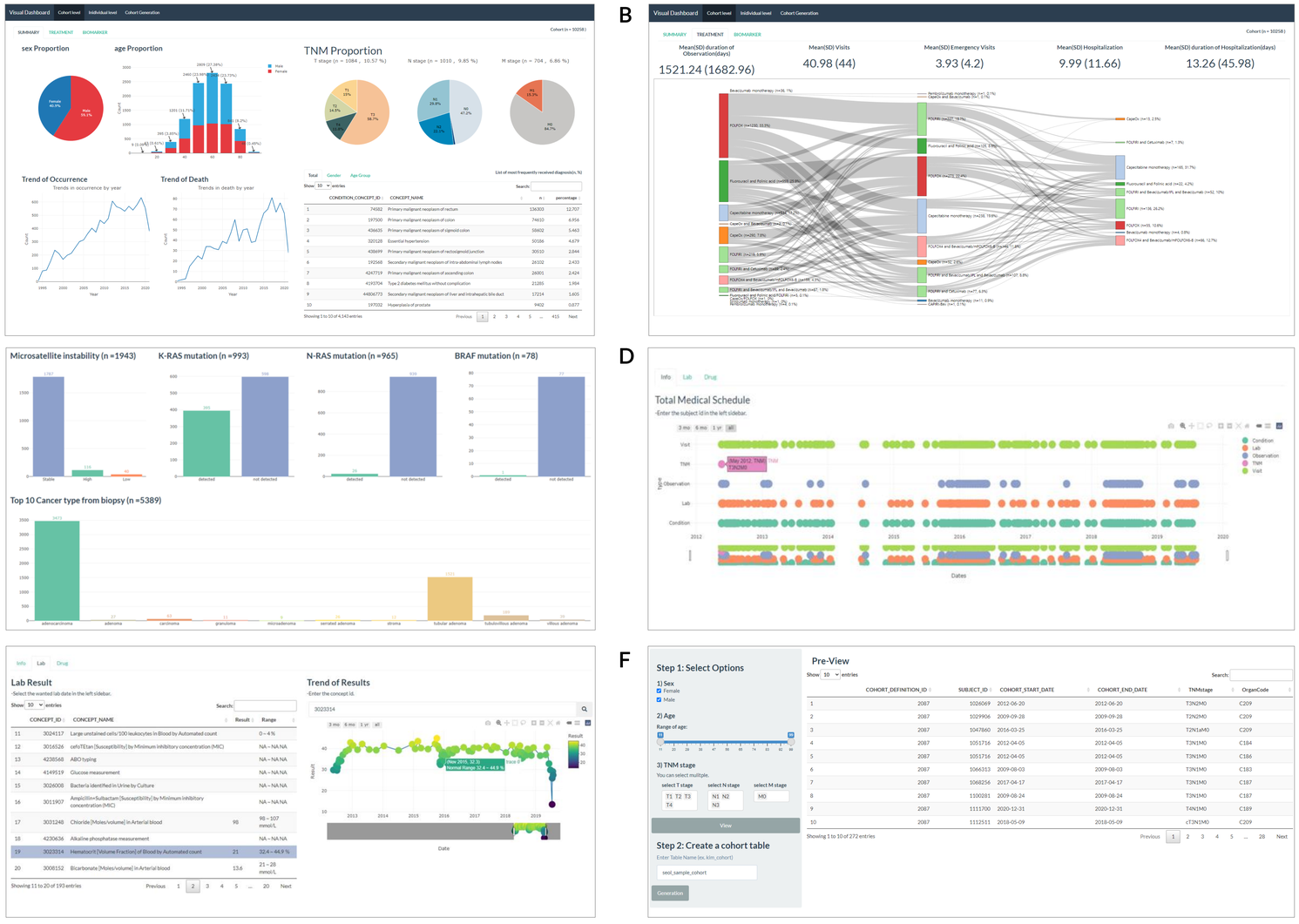

Figure 2. Snapshots of the dashboard that visualize the profiled patients' data. A, summary tab on cohort-level page. B, treatment tab on cohort-level page. C, biomarker tab on cohort-level page. D, info tab on the individual-level page. E, lab tab on the individual-level page. F, cohort generation page.

Results

- 10,258 patients with colorectal cancer were identified in the AUSOM database.
- We derived the regimen-level information of chemotherapy and inserted 212,327 episodes into the OMOP-CDM EPISODE table.
- Pathologic clinical information like cancer type and the result of gene mutation tests was extracted from 16,072 biopsy reports.
- The framework creates a dashboard that visualizes the profiled patients' data and consists of three main parts: 1) cohort-level visualization (Fig 2A, 2B, 2C), 2) individual-level visualization(Fig 2D, 2E) and 3) cohort generation (Fig 2F).

Conclusions

- We developed a standardized framework named RHEA, representing the longitudinal status of patients with cancer.
- This visualized information can help clinicians to monitor or analyze patients with cancer more in-depth.
- Further, we will insert unstructured data in RHEA.

This research was supported by a grant of the project for Infectious Disease Medical Safety, funded by the Ministry of Health, Republic of Korea (grant number: HG22C0024). This research was also funded by the Bio Industrial Strategic Technology Development Program (20003883, 20005021) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea) and a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health &Welfare, Republic of Korea (grant number: HR16C0001).