Name:	Sungjae Jung
Affiliation:	Dept. of Biomedical Informatics, Ajou Univ. School of Medicine, Suwon, Korea
Email:	sungjae.2425@gmail.com
Presentation type (select one):	Poster

A web based integrated code generating system for cohort analysis

Sungjae Jung, B.E.[1], Dukyong Yoon, M.D. Ph.D.[1], Rae Woong Park, M.D. Ph.D.[1]
[1]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Korea

Abstract

Cohort study is a commonly used method for clinical data analysis. Observational Health Data Sciences and Informatics (OHDSI) consortium released various method libraries for observational study including cohort analysis and web based analytic tools. Existing CIRCE and CohortMethod are powerful tools to define and generate SQL code for a cohort, and to perfume a cohort analysis easily. However to run the CohortMethod with CIRCE, we need to define separate codes for outcome, exposure group, comparator group, and finally R codes for cohort method. As we need four codes to run a cohort study, sometimes it is confusing without prior information on the study. Thus we developed a web application conveniently generating codes for data extraction and analysis. The Code generator (a tentative name) integrates all the necessary definitions (outcome, exposure and comparator). Applying the CohortMethod on the three coupled definitions, a user can easily perform a cohort study in a seamless manner.

Introduction

The cohort study is a commonly used method for clinical data analysis. To conduct a cohort analysis with CohortMethod in the Methods library of Observational Health Data Science and Informatics (OHDSI), defining cohorts of exposure and comparator groups are necessary. Although a cohort can be easily defined by using CIRCE tool, however each of exposure group and comparator group should be defined separately in the list. As the list of definitions are independently registered in the CIRCE, sometimes it is hard to match the exposure group and corresponding comparator group in the lists without prior informed knowledge.

In this study, we developed a web application integrating all the necessary definitions (outcome, exposure and comparator) and R codes for CohortMethod to enable easy execution of a cohort study in a seamless manner.

Method

In this study, we developed a web interface (hereafter 'Code generator' – a tentative name), which enable cohort extraction to cohort analysis in a seamless manner. The Code generator is an expanded cohort definition tool based on CIRCE. The Code generator provides Graphical User Interface (GUI), data access and execution of substantive functions based on WebAPI.

Result

A user can access to Code generator through web browsers (e.g., Google chrome, Microsoft Internet Explorer, etc.), and can define outcome and set specific conditions, then define cohorts (e.g., exposure and comparator groups), and generate codes for data extraction and cohort analysis. Finally, a user can conduct cohort analysis manually using the generated codes. Figure 1 illustrate the structure of the application and figure 2 shows how it works. The Code generator has four tabs as of outcome, exposure, comparator and Cohort method (figure 3). Each tab has SQL codes for outcome definition, exposure group extraction, comparator group extraction, and R codes for CohortMethod for

running against the cohorts. The Code generator depends on and heavily utilize the functions of CIRCE and WebAPI.

Figure 1. Conceptual system diagram of the Code generator	**Figure 2.** Example of the usage of the Code generator
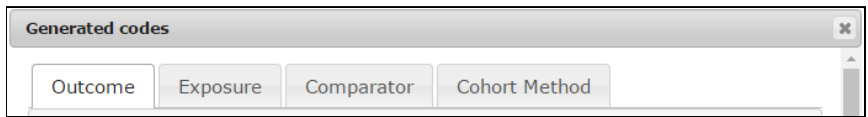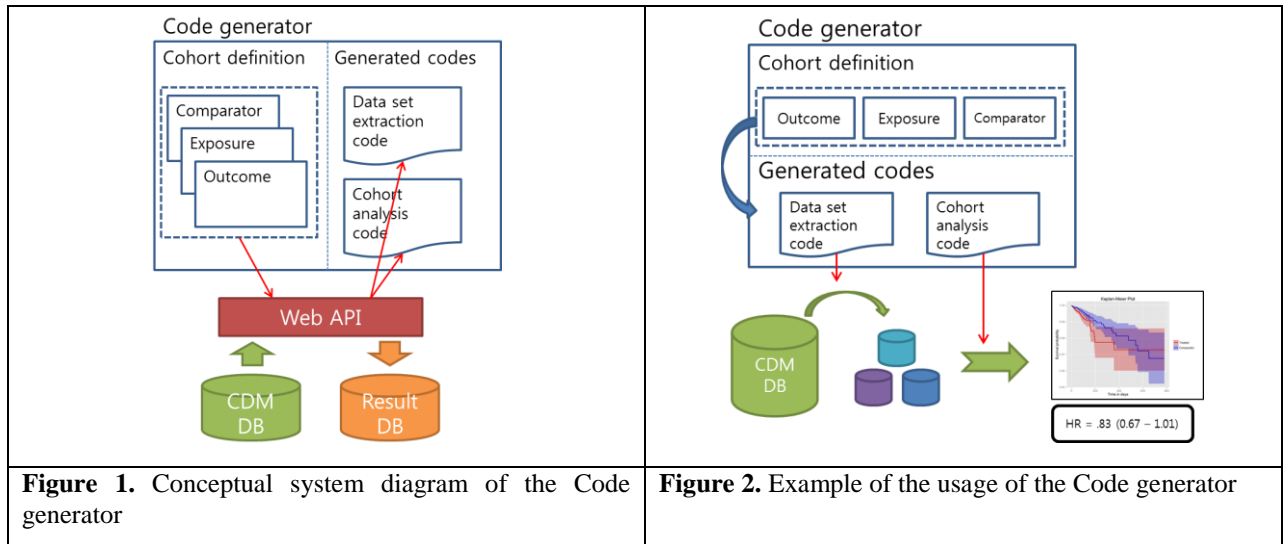
Figure 3. Generated codes menu of the Code generator.

Conclusion

In this study, we developed a web application conveniently generating codes for data extraction and analysis. The Code generator integrates all the necessary definitions (outcome, exposure and comparator). Applying the CohortMethod on the three coupled definitions, a user can easily perform a cohort study in a seamless manner.