MEDINFO 2007
HISA Ltd Australia
© 2007 The authors. All rights reserved.

Conducting Time Series Analyses on Large Data Sets: a Case Study With Lymphoma

Vojtech Huser[a,b], Roberto A. Rocha[b], Martin Huser[c]

[a] *Intermountain Healthcare, Salt Lake City, Utah,* [b]*University of Utah, Department of Biomedical Informatics*
[c]*Obstetrics and Gynecology Clinic, Masaryk University Hospital, Brno, Czech republic*

Abstract

We describe the application of our previously developed analytical infrastructure called RetroGuide to conduct an observational retrospective cohort study using a clinician friendly flowchart approach.

Keywords:

time series analysis, retrospective observational study, Hodgkin lymphoma, workflow technology

Introduction

Retrospective observational studies can provide important background knowledge for preparing expensive clinical trials. Our main purpose is to demonstrate the use of a suite of medical informatics tools called RetroGuide. The problem of investigating the pregnancy rate in women who underwent toxic therapy for Hodgkin lymphoma is used as a case study.

Methods

Intermountain Healthcare's Enterprise Data Warehouse (EDW) was used as the data source. It contains a lifetime Electronic Health Record (EHR) for each patient and it includes encounters from multiple healthcare facilities. RetroGuide was used to create graphical scenarios representing a temporal sequence of events of interest associated with analytical questions. RetroGuide uses a workflow engine to execute the modeled process on real retrospective data and produces a series of analytical reports about how well the model fits the data from EDW.

Women with a diagnosis of Hodgkin lymphoma between the age of 18 and 40, and with at least 5 years of follow up data were included in the study (140 subjects total).

Results

A simplified version of the executable flowchart addressing the main question about the pregnancy rate is shown in Figure 1. Arrows represent transitions and may contain a condition: e.g. "previous event found" (black) or "not found" (grey). The variables inside brackets show that the engine can remember the timestamp of a previously identified event. After the execution of this scenario, the resulting RetroGuide

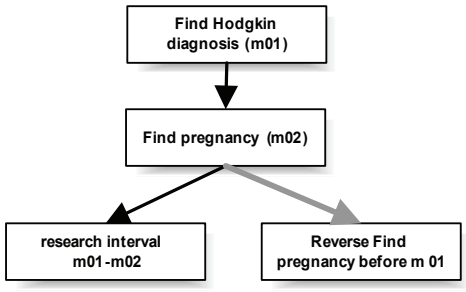

Figure 1 - RetroGuide Scenario (simplified version)

summary report showed that 22.9% of the patients treated for Hodgkin's lymphoma successfully completed a pregnancy within 5 years of the diagnosis (32/140; 95% CI 15.5-30.2).

Discussion

We found only one observational study investigating pregnancy after treatment for Hodgkin's lymphoma (prospective). Our results are comparable to this study, however the retrospective design is cheaper and faster to conduct, with results directly relevant to our local population. Additional advantages of RetroGuide when compared to traditional SQL-based database tools are: (a) user-friendly flowchart model as a shared logic formalism between the data analyst and clinicians; and (b) support for extensive "drill down" capability into available EHR data via a hierarchy of customizable reports. We have used RetroGuide to investigate 13 additional clinical questions relevant to lymphoma therapy and pregnancy.[1] We have also initiated a controlled study to evaluate the suitability of RetroGuide's methodology to empower champion clinicians to effectively utilize a large EDW. The value of our graphical tool will increase with a more widespread adoption of longitudinal and detailed EHRs.

References

[1] Huser V, Rocha RA, and James BC. Use of workflow technology to analyze medical data. In: 19th IEEE CBMS Symposium Proceedings, 2006; pp. 446-450

1 additional data available at http://workflow.minfor.net