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

(a RetroGuide Case Study)

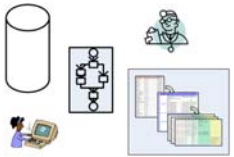
Vojtech Huser, MD Roberto A. Rocha, MD PhD Martin Huser, MD

Intermountain Healthcare & Biomedical Informatics, University of Utah, USA

O&B Clinic, Masaryk University Hospital, Brno, Czech republic

Intermountain℠ Healthcare

The University of Utah Biomedical Informatics



RetroGuide *clinician-friendly flowchart alternative to SQL !*

Problem

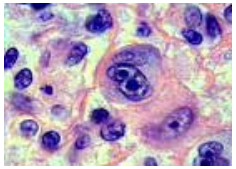
```
SELECT rownum, VAL_NUM2, T2TXT, times
FROM
(
SELECT  empi, VAL_NUM2, T2TXT, count(*) as times
FROM sandbox.ev9
WHERE VAL_NUM2 is not null
group by empi, val_num2, T2TXT
)
order by VAL_NUM2
```

- Amount of EHR data stored in databases grows exponentially
- SQL is currently predominant way of analyzing it
- Clinicians need extra analyst person in order to run advanced queries
- Clinicians can not review SQL code and effectively collaborate with the analyst
- SQL has limitations (e.g. temporal queries, GUI, learning curve)

RetroGuide

- internally developed analytical suite based on workflow technology
- uses standard process modeling language (XPDL)
- re-uses workflow editor as the modeling application
- provides set of external libraries called by the flowchart nodes

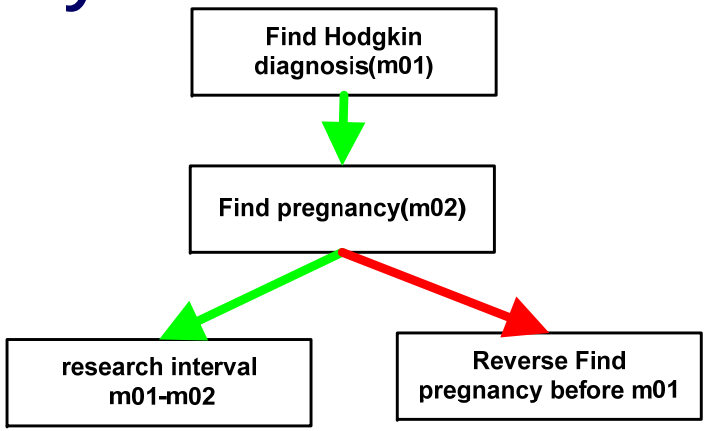
Lymphoma case study



- problem in cancer domain chosen as a case study to demonstrate features of RetroGuide
- retrospective observation study with 140 cancer subjects followed for 5 years
- primary outcome: rate of completed pregnancy after toxic cancer therapy

RetroGuide methodology

- 4 phases (data extraction, scenario creation, scenario execution and reports review)
- RetroGuide scenario has 2 layers
  - flowchart layer
    - (user-friendly to clinicians)
  - code layer
    - (makes the flowchart executable)



Analogy to manual chart review

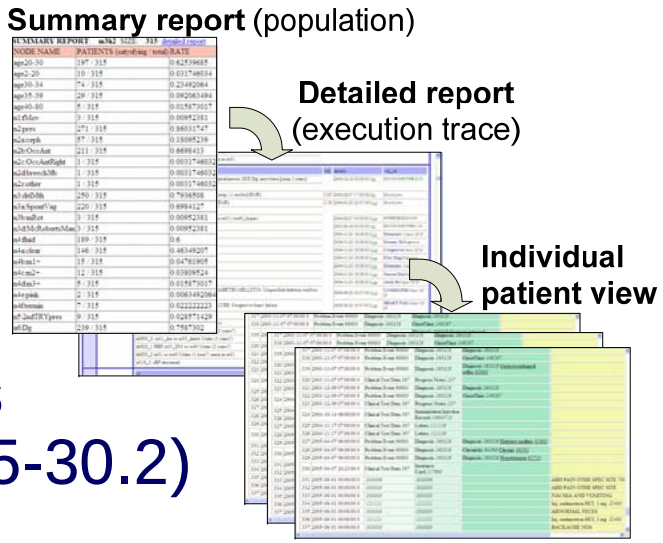
- RetroGuide scenario resembles instructions for manual chart review
- RetroGuide uses procedural approach as opposed to a declarative SQL code
- Analytical problems can be formulated having single patient in mind
- Ability to use constructs like “Time Jump” and remember values or timestamps by using variables

Intermountain's EHR data

- 3.2 TB large EDW covering 3+ million patients
- integrated lifetime EHR
- clinical data (lab, medication, clinical) combined with administrative data
- only coded data was utilized

Intermountain℠ Healthcare

Results



- 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)
- Other clinical factors were investigated with RetroGuide scenario but necessary EHR data elements were present only in limited subset of patients
- Comparable prospective study with 85 subjects with 8.5 years follow-up time found 36.9% rate of pregnancy [1]

[1] Franchi-Rezgui (2003) "Fertility in young women after chemotherapy with alkylating agents for Hodgkin and non-Hodgkin lymphomas

Conclusion

- RetroGuide is a user friendly modeling and analytical approach based on established technology (Bussines Process Management)

- We have demonstrated the use of RetroGuide for conducting a retrospective observational study in cancer domain

RetroGuide advantages

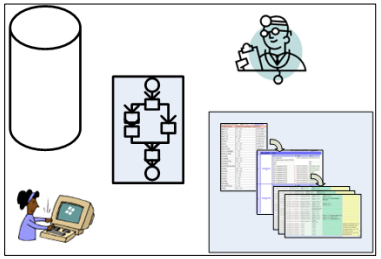
- graphical element of the analytical model (middle layer)
- drill-down capabilities via linked reports
- ability to run code not covered by the query language
- combine multiple queries into one model
- single patient execution model resembling decision support mode of operation

Future

- RetroGuide use and potential will grow with more detailed EHR data and longer time-span available
- RetroGuide functionality is intended to be an exploration and feasibility study of alternative analytical approach to current code-only based methods

References

- Huser V, Rocha RA, James BC (2006) Use of workflow technology to analyze medical data. Proc of IEEE CBMS06 Symposium pp. 446-450
- Huser V (2007) Running Decision Support Logic Retrospectively to Determine Guideline Adherence: a Case Study With Diabetes, Spring AMIA2007
- Huser V, Rocha, RA (2007) Analyzing medical data from multi-hospital healthcare information system using graphical flowchart models," BMIC Symposium (accepted).
- Huser V, Rocha RA (2007) Retrospective Analysis of the Electronic Health Record of Patients Enrolled in a Computerized Glucose Management Protocol," IEEE CBMS07 Symposium (accepted)



- RetroGuide project website: workflow.minfor.net
- Contact email: vojtech.huser@intermountainmail.org

  Salt Lake City, Utah, USA