Development of A Medical Informatics Data Warehouse

Cai Wu\textsuperscript{a,b}, MS
\textsuperscript{a}Physician Relations, University of Texas MD Anderson Cancer Center, Houston, TX
\textsuperscript{b}Department of Family and Community Medicine, Baylor College of Medicine, Houston, TX

Abstract
This project built a medical informatics data warehouse (MedInfo DDW) in an Oracle database to analyze medical information which has been collected through Baylor Family Medicine Clinic (FCM) Logician application. The MedInfo DDW used Star Schema with dimensional model, FCM database as operational data store (ODS); the data from on-line transaction processing (OLTP) were extracted and transferred to a knowledge based data warehouse through SQLLoad, and the patient information was analyzed by using on-line analytic processing (OLAP) in Crystal Report.

MedInfo DDW Description
1. Basic Goal of MedInfo DDW
To analyze medical information by
- Season of patient coming to clinic
- Age group of patient, gender of patient, race of patient and marital status of patient
- Disease diagnosis
- Medication treatment
- Lab test observation
- Physician group
2. Dimensional Model of MedInfo DDW
There are 4 fact tables in MedInfo DDW:
- MedInfo: this is a factless fact table, which records patient general information after first in clinic.
- LabTest: this is a numerical fact table, which records patient general information and lab test item and test values. We can get the test values for each patient since the patient start test, and find the changes of the values during certain time period after medication.
- GenInfo: this is a transaction fact table, which records patient general information and problem start and stop date, medication start and stop date. We can find how long for the problem continuing, how long for the medication continuing, and the reason of the problem stopping and the reason of the medication stopping.
- ClinicReason: this is a textural fact table for textural items in observation, which records patient general information and check items in the clinic. We can find the reason for patient in clinic and the item to be checked, and the serial items during certain time period after medication treatment.

There are 12 dimensional tables in MedInfo DDW:
- Patient appointment time and lab test time (Time dimension)
- Patient information (Patient dimension, Age dimension, Gender dimension, Race dimension, Marital dimension)
- Patient physician information (physician dimension)
- Disease information (Disease dimension)
- Medication treatment information (Medication dimension)
- Disease stop information (DiseaseStop Reason dimension)
- Medication stop information (MedStop Reason dimension)
- Lab test values (Observation dimension)
3. ODS
Source tables in FCM Logician Oracle database
- Person (basic patient information)
- Problem (medical problem information)
- Medicate (medical treatment information)
- Provider (physician information)
- Appointment (appointment information)
- ObsHead (lab test item information)
- Observation (observation information for lab test)
4. Extract, Load and Transfer (ETL)
- Store Procedure
- Spool the source data from database
- SQLLoad
5. Building MedInfo DDW
- The data in the ODS will be transfer into MedInfo DDW period based on timestamp (weekly or monthly depending on the frequency of the data updating).
- For some index dimension tables (Age, Gender, Race, Marital Status, ProbStop Reason, MedStop Reason), the data will be inserted directly by hardcode; for Time dimension, the PL/SQL Procedure will be used to generate the date, month, year, season; for rest dimension and fact tables, the data will be selected from ODS and inserted into MedInfo DDW.