

## POSTER ABSTRACTS

10<sup>th</sup> Annual HMO Research Network Conference

May 3-5, 2004 Dearborn, MI

### Information Technology 27

#### **Distributed Data Model for the Vaccine Safety Datalink (VSD) Project**

Baggs, J; Lewis, E; Amundson, JH; Berger, N; Blake, M; Butani, AL; Canton, F; Drew, L; Fox, R; Lee, M; Malais, D; Norman, R; Riedlinger K; Shay, D

**Background:** To monitor and evaluate vaccine safety, the Centers for Disease Control and Prevention (CDC) established the Vaccine Safety Datalink (VSD) project in 1990. As part of this research, the VSD collects automated medical care data on a large population currently or previously enrolled at eight health maintenance organizations (HMOs) located across the United States. Previous to the year 2001, annual automated data formatted to CDC specifications as SAS datasets were sent as a cycle tape to the CDC. Because of recent heightened confidentiality concerns, this centralized data model has been replaced by a distributed data model.

**Methods:** In the distributed data model, the HMOs no longer provide annual automated data sets to CDC. Rather, the data are formatted as SAS datasets and reside at the HMO sites. Two different methods of data access have been developed to compile data sets for specific studies. In the first method, the HMO computer at specified intervals retrieves SAS program code for a specific study from a secure, pre-specified location. Four SAS macros have been developed to facilitate access of the data and retrieval of SAS logs and output. In the second method, CDC researchers access the data files required for a specific study through a secure SAS remote session using SAS Connect. Both methods will use encryption technology to insure security of data transmissions.

**Results:** Distributed data models have been developed and successfully tested at all eight HMO sites. Six sites developed distributed data models using the first method described above while two chose the latter method.

**Conclusions:** The application of the distributed data model with some site-specific components will permit VSD research to continue in a timely manner while providing state-of-the-art protection of the confidentiality of individual medical information used in VSD studies.