Tools and Data Management For Multi-Center Cancer Genetics
Jamie Malloy, M.S.; Rebecca Sutphen, M.D.; Lynn Anderson, M.S.,
H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida

Introduction. A typical cancer genetics counseling session consists of a genetics counselor gathering personal and familial risk data from a patient and analyzing this data to determine the risk a patient has of developing cancer, or of being predisposed to a greater than average risk of developing cancer. This process has, in the past, been done manually, with the counselor sketching out the pedigree by hand, often after the patient leaves. An added requirement to this scenario is a research need to not only gather this data, but also to save it in some format that can later be evaluated by a biostatistician for cancer research purposes. The need for a database application arises.

A web-based computer system for Cancer Genetics practice and research has been developed to provide a toolkit for the genetics counselor, as well as an environment for data and study management. The system is used not only within Moffitt but also accommodates the multi-institutional Florida Cancer Genetics Network (FCGN), a statewide network designed to empower the Florida Cancer Genetics community and to provide a cancer genetics research repository for analysis of a variety of study hypotheses. This system will be demonstrated during the theatre presentation. System features will be illustrated from Moffitt Cancer Center in Tampa, Florida, via the Internet. System architecture will also be described.

System Features. The system 1.) Provides a communications and tools support environment for a community of Cancer Geneticists and Genetics Counselors distributed around Florida, 2.) Facilitates the collection, storage, and management of data for cancer genetics research projects, 3.) Provides a seamless interface for pedigree drawing and BRCA risk assessment, 4.) Provides tools for study/subject management, such as reports, labels, and schedule helpers, 5.) Protects the confidentiality of the data, and 6.) Delivers the data for reporting and analysis.

System Design. The need to support cancer genetics professionals distributed geographically across the state motivated design of a web-based system delivered over the Internet. This decision led directly to particular security and data collection strategies. Cancer Genetics data is particularly fragile. Use of the Internet for data exchange (even de-identified data) motivated an elaborate security strategy. In addition to application level security, requiring username and password, the system uses 128-bit encryption and a 2nd level of authentication using a key fob mechanism. The key fob mechanism requires the user to enter in a pin# plus the dynamic password provided by the key fob.

Workflow requirements of the cancer genetics counselors motivated design decisions regarding data collection mechanisms and provision of tools. It was determined that counselors would download/print the personal medical history and family history (tele)forms, fill them in, and fax them directly into the database. Within moments the data can be used as input for delivering a family pedigree (pdf) and a risk assessment.

System Architecture. The application uses an HTML web interface together with Microsoft COM components on a Windows 2000 IIS application server to allow add, edit, and delete of data stored in an Oracle backend DBMS. The interface is forms-based. Online forms include a family history form, containing risk assessment and relationship information for members of a family, and a personal and medical history form, which gathers detailed risk and contact facts about the proband of that family. The forms are filled in by the patient or genetic counselor and faxed or scanned into the data system using automated forms processing software. The data is automatically exported to the database. The system then automatically, based on the form data, draws a pedigree and calculates a BRCApro risk assessment for the proband and presents it via the web for the genetics counselor to view. Other tools to the genetic counselor include automated printing of identifiers on forms and add/edit of data to the database via the web interface.

The data model is also forms-based. Most sections of the personal and medical history form correlate to a table in the database system. The interesting part of the data model has more to do with the family history. Each individual entered into the system is given a unique identifier, which is actually a combination of a unique family identifier followed by a sequence number unique to that individual within the family. This unique identifier is used to not only identify an individual but also, when grouped by family identifiers, identify single families. The only relationship information required to identify an individual’s place within the family is her mother and father’s unique identifiers. Parental ID’s are the only information (along with family grouping ID) needed to reconstruct all other family relationships.

Conclusions. The system is successful with users at Moffitt and in the FCGN. An important strength is the pedigree drawing tool embedded in the application. Use of automated forms processing software to enter the data into the database without manual intervention is another strong feature, although vulnerability to snags in faxing or scanning the forms into the system causes some frustration on the part of the users.