

POSTER ABSTRACTS

14th Annual HMO Research Network Conference
April 13-16, 2008 – Minneapolis, MN

12:45–1:30 pm • Tuesday, April 15th
Salon C

PS2-41

Tumor Registry Content Area of the Virtual Data Warehouse

Karen E. Wells, BA, Henry Ford Health System / Department of Biostatistics and Research Epidemiology;
Richard Krajenta, BS, Henry Ford Health System / Department of Biostatistics and Research Epidemiology;
Lois Lamerato, PhD Henry Ford Health System / Department of Biostatistics and Research Epidemiology;
Donald J. Bachman, MS, The Center for Health Research, Kaiser Permanente Northwest, Portland, OR

Background: The Virtual Data Warehouse (VDW) was created as a mechanism to produce comparable data across health systems for the purpose of facilitating multi-site research within the Cancer Research Network (CRN). Each site maintains its own standardized files according to mutually agreed upon dataset definitions. The common structure of the VDW files enables a site programmer to distribute a SAS program to all participating sites that, with minimal modifications, can run against their local VDW. The program produces de-identified summary results that can be transferred to the coordinating site programmer. The Tumor content area is just one of eight standardized files maintained at each of the local CRN sites with tumor registries or access to tumor registry data.

Methods: The Tumor content area was one of the first to be developed by the Scientific & Data Resources Core (SDRC). The common format for each element of the Tumor content area - variable name, label extended definitions, code values and value labels, was largely driven by the data standards as defined by the North American Association of Central Cancer Registries (NAACCR). The Tumor content area contains detailed information on patient demographics and incident primary tumors such as date of diagnosis, ICD-O site, morphology, stage at diagnosis and first course of treatment. Eleven CRN sites maintain a Tumor content area for the VDW.

Results: The Tumor content area has been used for case identification in several funded multi-site CRN studies. Additionally, the Cancer Counter on the CRN website uses aggregate data from the Tumor content areas of all CRN sites to provide counts of primary tumors for every combination of selected variables. The Cancer Counter has proven to be a valuable tool in facilitating proposal development.

Conclusions: The standardization of the Tumor content area across CRN sites enables the sharing of compatible data in multi-site studies by improving programming efficiency, accuracy and completeness of data.