small set of SAS programs and a T-SQL script. The results at Group Health show the methodology to be an effective means of using the VDW to quickly populate I2B2.

Keywords: Data loading, I2B2, VDW

doi:10.3121/cmr.2011.1020.ps1-12

C-A1-01:

Using Data Transformations, Derived Values, and Cryptographic Functions to Protect PHI in the VDW

Daniel Ng, BSE, MBA; Kaiser Permanente Northern California, Division of Research

Background and Aims: HIPAA policies define values derived from PHI as also being PHI. Consequently, applying algorithmic functions to PHI has been viewed as having little benefit to research data. However, short of full de-identification, the use of transformed PHI may reduce compliance risk and increase security of routine data handling. Our aim is to: 1) present a general framework for evaluating methods of de-identifying/protecting PHI, and 2) evaluate how well selected mathematical functions, including common cryptographic functions, can enhance protection of PHI in the HMORN Virtual Data Warehouse (VDW). **Methods:** The methods used include a review of technical literature/material, from both within and outside traditional research disciplines, followed by analysis and application of findings to the issues addressed here, including: 1) reviewing both the regulatory and practical context for protecting PHI in research data; 2) developing a set of criteria to evaluate the benefits and costs of methods for PHI protection/de-identification; 3) reviewing the basic uses of general cryptography; 4) comparing/contrasting the needs of general cryptography with those of PHI protection in research data; 5) evaluating selected methods of protecting PHI against the proposed criteria. **Results:** The proposed framework for evaluating PHI protection methods includes five criteria: 1) effect on usefulness of data; 2) effect on ease of use or analytical efficiency; 3) net impact on data security; 4) system implementation costs; 5) negative effect on data quality.There are several methods for protecting PHI that can be easily implemented in the VDW, including: 1) the creation of linking variables that eliminate the need to routinely query PHI variables, such as service dates; 2) the MD5 cryptographic hash function, which can be used to obscure any PHI data and is relatively easy to implement in SAS, the standard analysis software platform used in the HMORN. **Conclusions:** Stewards of research data, like the VDW, should adopt the use mathematical functions, including cryptographic hash functions, to transform PHI into derived values. Such methods do not replace the need for full de-identification, but can enhance security and reduce compliance risk during routine data handling.

Keywords: VDW, De-identification, PHI

doi:10.3121/cmr.2011.1020.c-a1-01

C-B1-01:

Psychotherapeutic Prescription Patterns Across Health Plans in the Mental Health Research Network

Christine Stewart, PhD[1]; Greg Simon, MD[1]; Irina Miroshnik, MS[2]; Brian Ahmedani, PhD[3]; Matthew Lakoma, BS[2]; Dana Larkin, BS[3]; Karen Niedenfuer, BS[4]; Robin Whitebird, PhD, MSW[4]; Carsie Nyirenda, BS[5]; Arne Beck PhD[5]; Marvin Adams, BS[6]; Robert Davis, MD, MPH[6]; Mark Schmidt, BS[7]; Ameena Ahmed, MD, MPH[7]; Jamila Gul, BS[8]; Phillip Crawford, BS[9]; Frances Lynch, PhD[9]; Julie Liu, BS[10]; Karen Coleman, PhD[10]

[1]Health Research Institute; [2]Harvard Pilgrim Health Care Institute; [3]Henry Ford Health Systems; [4]Health Partners Research Foundation; [5]Kaiser Permanente Colorado; [6]Kaiser Permanente Georgia; [7]Kaiser Permanente Hawaii; [8]Kaiser Permanente Northern California; [9]Kaiser Permanente Northwest; [10]Kaiser Permanente Southern California

Background: The need for effectiveness and dissemination research is at least as great in mental health as in other areas of health care. The Mental Health Research Network aims to use the VDW for multi-site studies of mental disorder epidemiology, treatment patterns, or treatment effectiveness. We expect that the existing infrastructure is satisfactory in some areas, while others will need improvement or expansion. We present early results from validation studies of enrollment and pharmacy data on patterns of psychotherapeutic drug use. **Methods:** We identified the relevant denominator populations using existing VDW variables indicating continuous enrollment with any drug coverage in 2009. After developing National Drug Code (NDC) lists for drugs used in the treatment of depression, bipolar disorder, attention-deficit disorder, and other mental health conditions, we calculated rates of use of these drugs across ten health plans in 2009. **Results:** Antidepressant exposure rates vary across sites, although our data include no diagnostic criteria and will include uses for other conditions. Initial results indicate overall antidepressant use rates vary more than two-fold between ten sites, from 62 to 162 exposures per thousand members. In addition, the preferred member of a class (for example SSRIs), and the rate of any pharmacy utilization also differs across sites, although to a lesser degree. **Conclusions:** Data validation is an iterative process where results are compared to what is known and expected and discrepancies are investigated. The considerable variation in overall antidepressant exposure rates at different sites is unexpected, based on reports of the incidence of depression (ref) and warrants further exploration. Possible explanations include use of antidepressants for other conditions, cultural attitudes toward mental health treatment in different populations, or delivery system practices. Alternatively, they may be artifacts of information systems or incorrect interpretations of the underlying data. These possibilities can be explored with a combination of local expertise from each delivery system and further analysis of VDW data.

Keywords: Mental health, Pharmacy

doi:10.3121/cmr.2011.1020.c-b1-01

PS1-38:

Benefits of Best Practice for Multi-Site National Drug Code Identification

Christine Stewart, PhD[1]; Gene Hart, MS[1]; Chester Pabiniak, MS[1]; Brian Ahmedani, PhD[2]

[1]Group Health; [2]Henry Ford Health Systems

Background: Prescription drugs are identified in the VDW pharmacy file by National Drug Code (NDC.) Although the NDC constitutes a standard nomenclature, there is a no single, comprehensive resource for codes, either in the HMORN or beyond. Instead, each site maintains an "everNDC" file based on its own internal and external sources. Therefore, to identify NDC codes for drugs of interest for a multi-site study, the recommended process (VDW EverNDC Data Structure, Version 3) involves the collection and compilation of codes from each site. This is a frequent and time-consuming practice. We will present an exploration of the benefits of this approach in the experience of the Mental Health Research Network (MHRN.) **Methods:** We will compare the results of the recommended practice to those from a shortcut limited to RXnorm (RXN), an open-source library from the National Library of Medicine and First DataBank (FDB), a proprietary database licensed by Group Health, for antidepressants, lithium, anticonvulsant mood stabilizers, first and second generation antipsychotics, benzodiazepines, and stimulants. **Results:** In our process of identifying antidepressant NDCs from a list of 32 active ingredient names, RXN and FDB generated a list of 15401 codes. At Group Health (GH), our historical NDC file had 1653 additional codes. However, these additional codes only match 35 fills in our pharmacy file, the most recent being in 2002. For comparison, the full list of antidepressant NDC codes yielded an average of 480,000 fills per year at GH over the period 2000-2009. From nine other sites, we received an additional 1329 codes, increasing the number of codes by about 8%. **Conclusions:** Our initial results indicate that for antidepressants, site-specific NDC codes at GH represent rare and older prescriptions. The RXN/FDB method yielded ~90% of the total NDC codes identified at our site, and >99.99 % of the prescriptions identified. We need to confirm this across other sites and other drug classes, but it is possible that a simplified process for defining a list of NDC codes for a drug class would be suitable for many multi-site studies.

Keywords: Pharmacy, National Drug Code, everNDC

doi:10.3121/cmr.2011.1020.ps1-38