incidence is unclear. Improved understanding of this association may hold significant potential for population health improvement strategies. The purpose of this investigation was to study the association of adherence to varying levels of an optimal lifestyle with the incidence of chronic disease among employed adults (N=35,684). **Methods:** All baseline data were collected as part of a health assessment process during the year 2007. Participants were followed for two years. Descriptive statistics were calculated and reported as the proportion of the sample reporting a newly diagnosed chronic disease (diabetes, heart disease, hyperlipidemia, hypertension, back pain, or cancer) by the number of lifestyle behaviors that they practiced. **Results:** Compared with individuals who practiced no or only 1 optimal lifestyle behavior, the risk of developing a condition for individuals practicing 3 or 4 behaviors was 66% lower for diabetes, 45% lower for heart disease, 17% lower for hyperlipidemia, 15% lower for hypertension, 43% lower for back pain, and 24% lower for cancer. **Conclusions:** Adherence to optimal lifestyle is associated with lower 2-year incidence rates of several chronic conditions. Attention to population-based strategies that improve adherence to these four behaviors appears to positively affect risk of chronic diseases and their risk factors. To the extent that social marketing strategies and systematic measurement and feedback to individual members, patients, and groups (e.g., employer-based) increases adherence to multiple health behaviors, risk of chronic disease may be reduced.

Keywords: Optimal lifestyle adherence, Lifestyle behaviors, Health improvement strategies

doi:10.3121/cmr.2010.943.ps2-29

IRB Issues

C-A2-01:

IRB Review in the Era of ARRA: Streamlining Without Sacrifice

Jeffrey Braff, DrPH, MBA¹; Sarah Greene, MPH²; Karin Johnson, PhD²; Ella Thompson, BS²; Kate Rardin-Leahy, PhD³

¹Kaiser Permanente; ²Group Health Cooperative; ³Health Partners

Background and Aims: In the era of ARRA, we are routinely confronted with compressed timelines and an accelerated pace of research. Concurrently, discussions have intensified over recent months about the HMORN's capacity to conduct comparative effectiveness research on topics of national interest. Building on these unique challenges and opportunities, this interactive session will serve as a forum for conference attendees to discuss recent HMORN efforts to make multi-site IRB review as agile as possible. In particular, we will showcase the HMORN facilitated review process and discuss how it might be further improved or expanded. **Methods:** In early 2008, the HMORN ratified its first Standard Operating Procedure (SOP), which outlines procedures for inter-institutional IRB review of data-only studies. This SOP helped clarify what steps to take and who to contact, both in order to pursue ceding, and when multiple IRBs need to review the same study. This process has now been used in a majority of HMORN member institutions. As its use has increased, researchers and IRB administrators have become more familiar with the process. In September 2009, the HMORN Governing Board identified continuing such work to ensure IRB agility as a key 2010 priority for the HMORN. **Results:** In this session, we will review how to use the facilitated IRB review process outlined in the SOP and learn from attendees how we can further facilitate multi-site IRB review, along with attendant concerns such as HIPAA. In addition, we will give a snapshot of recent and upcoming developments across the Network sites, including human research protection program accreditation efforts and migration to electronic IRB systems. **Conclusions:** This session will enable research investigators, IRB administrators and HMORN leadership to engage in a constructive dialogue about the residual barriers to using the facilitated IRB process for data-only studies, and identify pathways to institutionalize this process in our Network's routine research processes. This session will also introduce how the demands of IRB accreditation and the implementation of new electronic IRB systems may facilitate or impede our progress in this area.

Keywords: IRB review, HMORN, Multicenter Studies

doi:10.3121/cmr.2010.943.c-a2-01

Health and Bioinformatics

C-A4-01:

Providing Care for Cancer Survivors in Integrated Delivery Systems: An Assessment of Current Practices and Future Opportunities

Jessica Chubak, PhD¹; Clarissa Hsu, PhD²; Leah Tuzzio, MPH¹; Beth Kirlin, BA¹; Larissa Nekhlyudov, MD, MPH^{3,4}

¹Group Health; ²Group Health Research Institute; ³Harvard Medical School;

⁴Harvard Vanguard Medical Associates

Background and Aims: There are currently more than 11 million cancer survivors in the United States; for many, cancer is a chronic condition that requires effective management and coordination of care. Little research on cancer survivorship models has been done, and most has focused on care provided in cancer centers. To our knowledge, no research has been conducted within population-based integrated delivery systems. To this end, the Cancer Research Network (CRN) Survivorship Scientific Interest Group (SIG) undertook a project to identify and describe; 1) current practices in cancer survivorship care, 2) goals and/or intentions to create cancer survivorship care programs, and 3) areas for future research in survivorship care models within integrated delivery systems. **Methods:** Semi-structured interviews were conducted with expert informants from 10 of 14 CRN sites. Individuals were selected based on their leadership roles in primary care/internal medicine, oncology, nursing and/or other cancer-related specialties. Interviews were conducted by telephone and were audiotaped, transcribed and coded. Analyses are underway; final results will be available at the HMORN meeting. **Results:** Of the initial 48 persons contacted, interviews were completed with 40 expert informants. Preliminary analyses suggest that in general respondents were supportive of providing cancer survivors with medical services customized to the needs of cancer survivors. At most sites, follow-up care for cancer survivors was provided by an informal collaboration between primary care physicians and oncologists. Respondents reported that the electronic medical records available at their sites appear to facilitate communication between providers caring for cancer survivors. Challenges in providing care to cancer survivors included allocation of time and resources. Few sites have survivorship care-specific guidelines, care plans or other formal tools to assist physicians in creating follow-up care plans for cancer survivors. The interviews identified 2 formal survivorship programs underway and one being pilot tested. **Conclusions:** While there appears to be an awareness of the needs of cancer survivors by the leadership in the integrated health care delivery sites, survivorship programs are generally not formalized. This study enhances our understanding of community-based cancer survivorship care and helps build the foundation for research on the effectiveness of different approaches to improve care.

Keywords: Cancer survivorship, survivorship models, integrated health delivery

doi:10.3121/cmr.2010.943.c-a4-01

C-B5-02:

H1N1 Vaccine Safety Monitoring in the Vaccine Safety Datalink Project: New Challenges and Useful Lessons

Tracy Lieu, MD, MPH¹; Sharon Greene, PhD, MPH¹; Ruihua Yin, MS¹; Rong Li, MS¹; Martin Kulldorff, PhD¹; Eric Weintraub, MPH²; Bruce Fireman, MA³; Roger Baxter, MD³; James Nordin, MD, MPH⁴; Steven Jacobsen, MD, PhD⁵; Edwin Lewis, MPH³; James Baggs, PhD²; Melisa Rett, MPH¹; Richard Platt, MD, MS¹; Grace Lee, MD, MPH¹

¹Harvard Pilgrim Health Care Institute and Harvard Medical School; ²Centers for Disease Control and Prevention; ³Kaiser Permanente Vaccine Study Center, Northern California; ⁴HealthPartners Research Foundation; ⁵Kaiser Permanente Southern California

Background and Aims: The worldwide introduction of influenza A (H1N1) vaccines in fall 2009 has raised widespread public questions about the safety of these new vaccines. The Vaccine Safety Datalink Project, a collaboration among CDC and eight health care systems in the HMORN, began accelerated

monitoring of this vaccine's safety in early fall 2009. **Methods:** The VSD Project conducts near real-time monitoring of potential adverse events following vaccination for all new vaccines. Each week, each site updates dynamic data files, and the coordinating center analyzes all data cumulatively. Repeated analysis of the same data requires adjustment for multiple testing. To address this issue, VSD uses sequential statistical methods. Evaluation of H1N1 vaccine safety poses special challenges, including confounding by indication, seasonality, and new priority groups including pregnant women. We are using several analytic methods, including: (1) Self-controlled case-series, which compares exposed vs. unexposed time windows within the same individual; (2) Poisson, which compares the outcomes during the current season with expected rates based on historical data for seasonal influenza vaccine; and (3) Difference-in-difference, which compares the relative risk in exposed vs. unexposed time windows for H1N1 vaccine during the current season with the relative risk for seasonal influenza vaccine during cumulative prior seasons. **Results:** Pre-specified outcomes of interest include: Guillain-Barre syndrome, neuropathies, seizures, encephalitis, Bell's palsy, ataxia, anaphylaxis, spontaneous abortion, pre-eclampsia, stroke, myocarditis and wheezing. Historical rates of these outcomes were calculated for time windows after seasonal influenza vaccination in prior seasons (2000/01-2008/09). The sequential analysis methods above will be used, and relative and absolute risks will be estimated for each outcome. For outcomes found associated with vaccination in preliminary analyses, additional evaluation will be conducted using temporal scan analyses, logistic regression, and when appropriate, medical record review.

Keywords: H1N1 safety, Influenza A vaccine monitoring, Vaccine Safety Datalink Project

doi:10.3121/cmr.2010.943.c-b5-02

C-B5-04:

An Algorithm for Classifying Potential Cases of Amyotrophic Lateral Sclerosis from Electronic Health Records Using Decision Tree Analysis

Kathleen Albers, MPH¹; Stuart Hoffman, DO²; Judith Readon, MS²; Mark Hornbrook, PhD³; Alison Naleway, PhD³; Joanna Bulkley, PhD³; Susan Andrade, PhD⁴; Daniel Newman, MD⁵; David Nerenz, PhD⁵; Margaret Gunter, PhD⁶; James Burmester, PhD⁷; Priscilla Velentgas, PhD⁸; Daniel Strickland, PhD⁹; Stephen Van Den Eeden, PhD¹⁰

¹Kaiser Permanente; ²Geisinger Medical Center; ³Kaiser Permanente Northwest; ⁴Meyers Primary Care Institute; ⁵Henry Ford Health System; ⁶LCF Research; ⁷Marshfield Clinic Research Foundation; ⁸Harvard Pilgrim Health Care; ⁹Kaiser Permanente Southern California; ¹⁰Kaiser Permanente Northern California

Background/Aims: The CDC is developing a national amyotrophic lateral sclerosis (ALS) registry that would be based on electronically available medical records and claims. The HMORN contracted with the CDC to help examine the feasibility of doing this in locations (e.g., HMOs) where complete longitudinal records are more likely to be available. Any final system used by the CDC will require an algorithm to classify putative cases. Our aim was to develop an automated algorithm to classify putative cases of ALS and other motor neuron diseases using electronic health records.

Methods: We identified potential cases of ALS by review of inpatient, outpatient, other provider services, mortality and pharmacy data for any evidence of ALS or, more generally, motor neuron disease (MND) from 2001 to 2005 at Kaiser Permanente, Northern California. The records of each individual were reviewed to classify into five categories: probable ALS, not ALS, other MND (not ALS), insufficient information, or unclear. Classification And Regression Tree (CART) analyses were used to create an algorithm based on a comprehensive set of variables that may predict the likelihood of having true ALS (or, conversely, ruling them out). **Results:** Multiple algorithms were developed with key variables being number times an ALS code appeared in the record, physician-type making diagnosis, alternative diagnoses, mortality records, and length of follow-up. In our best model, we were able to correctly classify 99.6% of those originally determined to be probable ALS. However, there remain a substantial number of individuals classified by review or algorithm in the unclear, insufficient, or other MND categories. In addition, with this analytic approach, the "costs"

of misclassification can be modified, and we will present alternative algorithms with a discussion of the benefits and disadvantages of each. Validation analysis results conducted in a second HMO will be presented. **Conclusions:** An algorithm was developed that appropriately classifies the vast majority of probable ALS cases. However, given the lack of detail in some records and the difficulty in diagnosing some clinical cases, there will be some misclassification when based solely on electronic medical records.

Keywords: Amyotrophic Lateral Sclerosis, Electronic health records, ALS Classification

doi:10.3121/cmr.2010.943.c-b5-04

PS1-15:

A Method for Discovering Variant Spellings of Terms of Interest in Clinical Text

David Carrell, PhD¹; David Baldwin, BS¹

¹Group Health

Background and Aims: Mining information from unstructured clinical text requires knowledge of synonyms and variant spellings corresponding to the term of interest. Lexicons such as the Unified Medical Language System identify recognized synonyms but are less helpful in identifying variant spellings or locality-specific language. Natural language processing algorithms employing dictionary look-ups frequently miss terms because of variant spellings, accounting for as much as 25% of algorithm failures. This problem is not easily resolved deductively or through domain expertise. We describe an inductive, empirically-driven method that identifies synonyms/variant spellings of a concept of interest. **Methods:** The conceptual basis for this method is the principle that terms with similar meanings tend to be found in similar textual contexts. Accordingly, sets of words found with high frequency in the context of two synonyms will overlap with high probability. Call the word of interest the target term; call the high-frequency words around it context terms; call words most likely to be synonyms/variants synonym candidates. A four-step process can be used to identify variants. Step 1: Identify the target term's highest frequency context terms. Step 2: Identify terms found in high frequency near each of the context terms. Step 3: Eliminate Step 1 & 2 terms appearing more often outside the context than in it. Step 4: Manually review remaining synonym candidates to identify true synonyms/variants. Using SAS we illustrate the method for target term HER2/NEU in Group Health chart notes for 2008, representing 200,913,281 lines of clinical text. **Results:** Four of the top 20 context terms near HER2/NEU had ratios of in_context_occurrence/out_of_context_occurrence > 1: OVEREXPRESSION, ONCOGENE, OVEREXPRESSED, C-ERBB-2. These context terms occurred 1196, 617, 540, and 170 times, respectively, and produced lists of 205, 119, 143, and 54 of their own context terms, respectively (excluding words with frequency < 2). Manual review yielded a final set of 23 synonyms and spelling variants, including HER-2, HER/2, HER-2-NEU, HER2/NUE (note spelling: "NUE"), FISH, IHC, and C-ERBB-2 among others. **Conclusion:** The method described provides an easily implemented approach to identifying synonyms/variants of terms of interest that can be used to enhance strategies for identifying concepts in clinical text.

Keywords: Identifying variant spellings, clinical text, medical terminology
doi:10.3121/cmr.2010.943.ps1-15

PS1-16:

Disenrollment Rates Among HMORN Sites by Age, Gender, and Medical Condition

Jennifer Ellis, MBA, MSPH¹; Marsha Raebel, PharmD^{1,2}; Addie Clark, BA¹; Susan Andrade, ScD³; Robert Davis, MD, MPH⁴

¹Kaiser Permanente Colorado; ²University of Colorado; ³Fallon/Meyers; ⁴Kaiser Permanente Georgia

Background: It is often desirable to include enrollment and disenrollment (or, conversely, retention) information in grant applications. Overall disenrollment has been evaluated as part of Virtual Data Warehouse quality