about whether screening and detection rates vary by race/ethnicity. **Methods:** Data from electronic health records (EHR) in a large ambulatory group practice were used to construct a cohort of persons aged 35 and older without prevalent diabetes ($N = 116,355$). They were followed for diabetes screening (average follow up 17 months in 2007-2011). We computed standardized screening and detection rates in each racial/ethnic group and examined predictors of screening. We compared positive detection rates for those who met the U.S. Preventive Services Task Force (USPSTF) or American Diabetes Association (ADA) screening criteria and those who did not. **Results:** Most patients (82%), including 75% of those meeting neither criteria, were screened during follow-up. Racial/ethnic minority groups had higher screening rates – Latino (82%), Asian subgroups (83% (Japanese) to 90% (Filipino)), non-Hispanic Black (92%) – than non-Hispanic White (80%) ($P < 0.001$). Detection rates were also significantly higher for some minority groups (Latinos (2.1%), Asian Indians (2.4%), and Filipinos (3.8%)) than non-Hispanic White (1.3%) ($P < 0.05$). Among screened, 1.6% had positive test results (number needed to screen (NNS) = 64). The NNS was 46 for those meeting ADA criteria, 36 for USPSTF criteria and 132 for those meeting neither. **Conclusions:** Nearly all patients in our study population received screening, and racial/ethnic minority groups were not less likely to be screened. Current clinical practice screens a high proportion of individuals meeting neither ADA nor USPSTF guidelines. Given low screening costs that may be appropriate because many new cases were detected among those without risk factors. Given the substantial benefit of early detection of diabetes our study suggests that universal, preventive screening for diabetes for asymptomatic individuals may have utility.

Keywords: Diabetes; Screening; Prevention

doi:10.3121/cmr.2013.1176.ps2-4

Genetics

B4-1:

An Overview of Kaiser Permanente's Research Program on Genes, Environment, and Health

Cathy Schaefer¹; Sarah Rowell¹; Mary Henderson¹; Lawrence Walter¹; Marianne Sadler¹; Sunita Miles¹; Donna Schaffer¹; Lisa Croen¹; Eric Jorgenson¹; Lawrence Kushi¹; Charles Quesenberry¹

¹Kaiser Permanente Northern California

Background/Aims: The Research Program on Genes, Environment, and Health (RPGEH) is a large, population-based resource for genetic epidemiology developed by scientists at the Kaiser Permanente (KP) Northern California Division of Research. The goal of this program is to link research data for 500,000 broadly consented health plan members, including longitudinal electronic medical records (EMR), genomic data from biospecimens, and environmental exposure data from surveys and geographic information system (GIS) databases in order to support research on many different health conditions. **Methods:** In 2011, the RPGEH established a state-of-the-art biorepository. The KP clinical lab infrastructure is used to collect and transfer blood samples to the biorepository where saliva, serum, plasma, buffy coat and extracted DNA samples are processed and archived. The biorepository includes a facility for separating blood into components for storage, as well as DNA extraction and normalization. Storage capabilities include -80 C and -20 C freezers, an ambient storage unit and LN2 freezers. The biorepository was designed to store blood components for up to 500,000 health plan members. Although initially participant recruitment was by postal mail, electronic methods of recruitment have now been developed. The goal of this effort was to decrease the per person cost of contact and to recruit younger members. By leveraging the KP EMR, blood draw orders are now automatically entered by RPGEH staff for consenting participants. These samples are then stored in the RPGEH biorepository and tracked in a Laboratory Information Management System. **Results:** A survey of adult health plan members has provided data on health-related behaviors and other risk factors for 430,000 RPGEH participants. To date, 190,000 biospecimens from consenting RPGEH participants have been collected. Genome-wide genotyping and telomere length analysis of 100,000 RPGEH participants have been performed. **Conclusions:** At >200,000 participants (with specimens for >100,000 genotyped), the RPGEH is already an extremely unique and valuable resource.

Keywords: Genetic Resource; Biorepository; Recruitment of a General Cohort

doi:10.3121/cmr.2013.1176.b4-1

B4-3:

Demographic and Behavioral Influences on Telomere Length and Relationship with All-Cause Mortality: Early Results from the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH)

Cathy Schaefer¹; Stan Sciortino¹; Mark Kvale²; Kyle Lapham²; Dilrini Ranatunga¹; Sarah Rowell¹; Marianne Sadler¹; Dana Ludwig¹; Lawrence Walter¹; Stephen Van den Eeden¹; Rachel Whitmer¹

¹Kaiser Permanente Northern California; ²University of California, San Francisco

Background/Aims: Telomere shortening is a biomarker of aging, but it is still unclear whether it directly influences aging-related health changes and mortality, or largely reflects the accumulating influences of environmental influences and declining health. Prior studies have reported associations between short telomeres and many health conditions. Mendelian syndromes and heritability studies demonstrate the biological underpinnings of telomere length (TL), yet other studies show that environmental and behavioral factors also influence telomere lengths. **Methods:** As part of the RPGEH, the Genetic Epidemiology Research Study on Adult Health and Aging (GERA), a multi-ethnic cohort (average age = 63 years) has measured TL from saliva samples on over 100,000 individuals with linked electronic medical records. A survey of demographic and behavioral factors was conducted 2 years prior to saliva collection. With these data, we examined demographic relationships with TL, behavioral influences, and relationship of TL with all-cause mortality following sample collection. **Results:** As expected, TL is inversely correlated with age, and women have longer telomeres than men except as young adults. All analyses controlled for age and gender. As seen in other studies, we find significantly longer TL among African Americans than other groups. TL is positively correlated with education and body mass index (BMI), and negatively correlated with cigarette smoking and alcohol consumption. We found no relationship with major depression or stress-related disorders, even when limited to recent episodes. We found that short TL was prospectively associated with mortality, although only those with the shortest TL were at increased risk; the association suggested a critical threshold of short TL determines effects on mortality. The association persisted even after adjustment of the demographic and behavioral factors such as age, sex, race, education, BMI, smoking, and alcohol consumption. **Conclusions:** While this could indicate a direct effect of TL on health, it will also be important to examine the extent of pre-existing morbidities in these individuals to understand their possible role in the pathway between TL and longevity.

Keywords: Genetics; Mortality; Epidemiology

doi:10.3121/cmr.2013.1176.b4-3

B4-4:

Genome-Wide Association Study of Macular Degeneration: Early Results from the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH)

Eric Jorgenson¹; Stan Sciortino¹; Ling Shen¹; Dilrini Ranatunga¹; Thomas Hoffmann²; Mark Kvale²; Yambazi Banda²; Pui-Yan Kwok²; Lawrence Walter¹; Neil Risch²; Cathy Schaefer¹

¹Kaiser Permanente Northern California; ²University of California, San Francisco

Background/Aims: Age-related macular degeneration (AMD) is the most common cause of vision loss in individuals over the age of 50 in the United States. Genetic factors explain a large portion of the risk of developing AMD, and genetic variants at the CFH, HTRA1/ARMS2, C2/CFB, and C3 gene loci have previously been associated with the disease. **Methods:** We conducted a genome-wide association study (GWAS) of AMD in the Kaiser Permanente Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort. The GERA cohort includes 110,266 subjects with extensive electronic medical record information on eye examinations, diagnoses and treatment of

vision disorders, and dense genome-wide genotype information on more than 675,000 genetic markers generated using Affymetrix Axiom arrays. The cohort is ethnically diverse, with 7.5% Asian, 7% Latino, 3.5% African American, and 81% non-Hispanic white subjects. We identified a total of 2,147 AMD cases (46 Asian, 125 Latino, 11 African American, and 1,965 non-Hispanic whites) and 37,521 controls (2,013 Asian, 3,201 Latino, 1,168 African American, and 31,139 non-Hispanic whites) for analysis. Analyses were conducted separately for each race/ethnicity group. **Results:** In the largest group, non-Hispanic whites, we identified highly significant associations with variants in the CFH and HTRA1/ARMS2 gene regions, and genome-wide significant associations in the C2/CFB and C3 gene regions. **Conclusions:** These results confirm those of previous studies and demonstrate the power of the GERA cohort for combining information from electronic medical records with extensive genotype data. This approach can be applied to additional vision disorder phenotypes, including response to treatment and disease progression.

Keywords: Genetics; Vision Disorders; Age-Related Macular Degeneration
doi:10.3121/cmr.2013.1176.b4-4

B4-5:

Neighborhood Deprivation and Telomere Length: Preliminary Findings from the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH)

Stephen Van Den Eeden¹; Mark Kvale²; Jun Shan¹; Kyle Lapham²; Dilrini Ranatunga¹; Sarah Rowell¹; Marianne Sadler¹; Dana Ludwig¹; Lawrence Walter¹; Rachel Whitmer¹; Charles Quesenberry¹

¹Kaiser Permanente Northern California; ²University of California, San Francisco

Background/Aims: Shortened telomeres have been associated with numerous adverse health outcomes. In addition, a number of environmental or external exposures, including smoking, air pollution and stress, have been reported to be associated with short telomeres. We sought to examine how neighborhood quality of participants in the RPGEH Genetic Epidemiology Research Study on Adult Health and Aging (GERA) cohort affected telomere length. **Methods:** The GERA cohort is a multi-ethnic cohort (average age = 63 years) of over 100,000 individuals with linked electronic medical records and questionnaire data. Telomere length was determined from a saliva sample in the Blackburn Laboratory using the novel Automated Telomere Length Analysis System (ATLAS) to handle the required high throughput processing of samples. Each sample was assayed six times using qPCR. Relative telomere length (T/S) was obtained from the initial concentrations of the sample telomere (T) with the corresponding sample reference gene (S). The distribution of (T/S) was found to be positively skewed and a log transformation was used to normalize the distribution. The final telomere length end point was the difference in adjusted means of telomere length per standard deviation unit by accounting for age and gender. The NDI is a standardized composite score of neighborhood quality derived from eight 2000 US Census data variables related to poverty/income, occupation, family structure, education and unemployment and normalized to a 100-point scale at the block-group level. **Results:** A higher NDI indicates greater neighborhood deprivation. The NDI was linked with residential address at time of sample collection. We observed a pattern of shorter telomere length with increasing level of neighborhood deprivation. The pattern persisted even after accounting for age, gender, race/ethnicity, smoking, BMI, and the presence of cardiovascular disease, diabetes and other comorbidity. **Conclusions:** Our results suggest that neighborhood can adversely impact telomere length. Future plans will be discussed.

Keywords: Telomere; Neighborhood Deprivation; Environment
doi:10.3121/cmr.2013.1176.b4-5

PS3-1:

Marshalling Site-Specific Data in Multi-Site Studies: Lessons from the eMERGE Network

Jane Grafton¹; Eric Baldwin¹; David Carrell¹

¹Group Health

Background/Aims: The Electronic Medical Records and Genomics (eMERGE) Network is a national consortium of nine institutions supported by the National Human Genome Research Institute (NHGRI) to study genetic correlates of disease by pooling data from local biorepositories and electronic data ecosystems. Three HMORN sites participate. Twenty-one of >40 planned genome-wide association studies (GWAS) have been completed, including GWAS for chronic, cognitive, cardiovascular, gastro-intestinal, hemotologic, infectious and other phenotypes, without patient contact. Transportable algorithms rely entirely on structured data from the EMR and, optionally, clinical text using natural language processing (NLP). Salient themes include bioinformatics, genomic medicine, privacy, and community engagement. **Methods:** Algorithms defining phenotype cases and controls are developed iteratively at a primary site in conjunction with one or two secondary sites. SAS, KNIME, Python, and random manual review at multiple sites establish an algorithm's positive predictive value (PPV) and portability. Validated algorithms, published as site-agnostic pseudo code documents on a secure Web site are implemented with local tailoring at remaining sites; data are pooled for analysis. **Results:** Pseudo code documents are an efficient way for communicating the logic and content of phenotype algorithms across sites when data not available in multi-site standardized formats (such as the Virtual Data Warehouse) are required or data must be obtained using NLP. Iterative, random-sample chart validation is an important method for developing robust transportable algorithms. Business intelligence rules systems such as KNIME simplify implementation of complex algorithms and NLP. A Web site for sharing pseudo code and validation results aids communication. **Conclusions:** Experiences from the eMERGE network offer valuable lessons for conducting multi-site studies in the HMORN when non-VDW and/or NLP-derived data are required.

Keywords: Genomic; Phenotype; Bioinformatics
doi:10.3121/cmr.2013.1176.ps3-1

PS3-2:

Informatics Challenges to Implement Pharmacogenetics to Clinical Practice

Zhiyi Zhou¹; Max He¹; Murray Brilliant¹; Ariel Brautbar¹; Aaron Miller¹; Bryan Weichelt¹; Simon Lin¹

¹Marshfield Clinic / Security Health Plan of Wisconsin

Background/Aims: A widely-held vision arising from genomics research is to be able to use information on genomic variation to guide clinical prescriptions. This translational project is designed to test the concept that sequence information from a patient can be coupled with electronic medical records for use in personalized medicine. **Methods:** Marshfield Clinic is a health care system in Wisconsin with 2 hospitals and 52 community care centers, a member of the HMORN, and partner of the Institute for Clinical and Translational Research, University of Wisconsin – Madison. Cattails, Marshfield Clinic's internally developed outpatient EHR, has been in existence since the late 1980s. The full-feature EHR is deployed on laptop, tablet and PDA technology and supports a variety of clinical decision support (CDS) applications surrounding care management, prevention, radiology orders and medication prescribing. Much of the underlying infrastructure that supports these CDS applications will be used when integrating genomic information into clinical practice. The underlying development principles for this initiative emphasize the importance of clinical workflow, genetic result significance, and the ability to reconfigure the system as knowledge improves. **Results:** A strategy of implementing an informatics system to satisfy the needs of pharmacogenetics from both research and clinical practice was tested. One in three of the 411,851 Marshfield Clinic patients in 2011 could benefit from a pharmacogenomic test. Implementing pharmacogenetics tests of three drugs (Warfarin, Clopidogrel, and Tamoxifen) at Marshfield Clinic could prevent potential adverse events that cost \$11.3 million annually. **Conclusions:** Such an informatics system will further