enable comparative effectiveness research in pharmacogenetics and pharmacogenomics.

Keywords: Pharmacogenetics; Clinical Implementation; Clinical Decision Support

doi:10.3121/cmr.2013.1176.ps3-2

PS3-3:

Biobanking for Research: A Survey of Member Attitudes and Understanding

Alanna Rahm¹; Michelle Wrenn¹; Heather Feigelson¹

¹Kaiser Permanente Colorado

Background/Aims: Population-based biobanks are an important resource for genetic research. To optimize recruitment of healthy individuals into a research biobank, it is important to know what potential participants understand about this type of research, and the use, risks, and benefits of providing samples to a biobank. **Methods:** We drafted informational recruitment and consent documents and then surveyed approximately 200 adult members in Kaiser Permanente Colorado clinic waiting rooms to evaluate: their understanding of biobanking and the materials we provided; their willingness to provide a blood sample to a biobank; and facilitators and barriers to their participation. Our materials stated that our proposed method of collecting biospecimens would be to collect an extra tube of blood at their next routine or clinical blood draw. Participants were also instructed that we were only gathering information, not actively recruiting for a biobank. **Results:** Our survey population was 65% female, mean age was 55, and 78% non-Hispanic white. Of 190 respondents, 79% understood that they would not need a separate blood draw to contribute to the biobank. Nearly all participants understood that they would not be paid for any products resulting from the use of their blood and would not receive results from their samples (91% and 84%). 81% understood that results from their sample would not be entered into their medical record. 67% of participants indicated they would be willing to contribute to a biobank, and 59% would still participate if providing information to US Government databases was required. 68% would donate a sample because "it is important to contribute to research," and over half the participants (55%) said they had no concerns about contributing to the biobank. Of those with concerns, 36% said information security was a reason. **Conclusions:** The data from this survey will help identify possible barriers to member participation in a research biobank and help identify ways to improve participation.

Keywords: Biobank; Genetic Research

doi:10.3121/cmr.2013.1176.ps3-3

PS3-4:

Genetic Risk Factors for Major Bleeding in Warfarin Patients in a Community Setting

Joshua Roth¹; Denise Boudreau¹; Monica Fujii¹; Federico Farin²; Allan Rettie²; Kenneth Thummel²; David Veenstra²

¹Group Health; ²University of Washington

Background/Aims: Studies evaluating newly initiated warfarin patients in anticoagulation clinics show that polymorphisms in CYP2C9, VKORC1, and CYP4F2 influence dose requirements, and CYP2C9 variants have increased risk of major bleeding side effects. However, these findings may not be applicable to long-term warfarin users or non-specialist care settings. Within this context, our objective was to evaluate the association between CYP2C9*2 and *3, VKORC1 1173, and CYP4F2*3 variants and major bleeding in long-term warfarin users in a community setting. **Methods:** We used a case-control design and recruited patients from Group Health (GH). Cases experienced a major bleeding event while receiving warfarin. Controls received warfarin on a randomly assigned index date, and had no major bleeding in the prior year. We identified major bleeding with an ICD-9 algorithm, and validated events with chart review. We obtained covariates from GH automated databases and a self-report survey. Our primary analysis used logistic regression to estimate the major bleeding odds ratio (OR) for variants vs. wild type patients. We also conducted an exploratory analysis to estimate the major bleeding OR for patients with variable vs. constant dietary vitamin K intake, stratified by genetic status. **Results:** We enrolled 265 cases

and 305 controls with an average of 3.4 and 3.7 years of warfarin use at the index date, respectively. In our primary analysis, the CYP4F2 variant was independently associated with decreased major bleeding risk (OR: 0.62, 0.43-0.91), and CYP2C9 and VKORC1 had null associations. In our exploratory analysis, CYP4F2 wild type patients demonstrated a trend toward increased major bleeding risk with variable vs. constant vitamin K intake (OR = 1.18, 0.64-2.16), while there was a null association in CYP4F2 variants. **Conclusions:** In the largest study of warfarin pharmacogenomics and major bleeding to date, we found that a common CYP4F2 variant is associated with a 38% reduction in risk of major bleeding. In contrast, CYP2C9 and VKORC1 variants demonstrated null associations. Our findings may reflect a gene-drug-environment interaction between CYP4F2*3, warfarin, and dietary vitamin K intake. Collectively, our findings expand understanding of genetic risk factors for major bleeding in warfarin therapy, and can potentially inform dosing and monitoring practices pending validation in independent cohorts.

Keywords: Warfarin; CYP4F2; Major Bleeding

doi:10.3121/cmr.2013.1176.ps3-4

PS3-13:

Re-Identification Risk Associated with Sharing Linked Genomic and Phenotypic Data from the Kaiser Permanente Research Program on Genes, Environment and Health (RPGEH)

Lawrence Walter¹; Stan Sciortino¹; Dirlini Ranatunga¹; Julia Kay¹; Dana Ludwig¹; Carol Somkin¹; Mary Henderson¹; Cathy Schaefer¹

¹Kaiser Permanente Northern California

Background/Aims: It is now understood that conventional de-identification methods such as the HIPAA Safe Harbor standard do not guarantee anonymity of patient records, which may be vulnerable to a variety of attacks aimed at re-identifying confidential information. We present an analytic framework for evaluating these risks quantitatively in order to be able to explicitly balance privacy and scientific utility. As a concrete example, we examine implications for patient privacy of plans to deposit over 70,000 full-genome genotypes and associated clinical data in the dbGaP federally-managed data repository, as a component of a NIH-funded study conducted by the Research Program on Genes, Environment, and Health (RPGEH) at the Kaiser Permanente Northern California Division of Research (KPNC DOR). Risks are examined from multiple perspectives and risk reduction strategies discussed. **Methods:** Two analytic approaches are described: (1) "k-anonymization", which computes risk based only on the distribution of cell sizes in the disclosed dataset; and (2) "k-map" which takes account of the characteristics of potential reference datasets – e.g., voter rolls, disease registries - which may be available to the attacker. Probabilities of re-identification were computed using a random sample of records from actual study participants, and assumed disclosure of the following phenotypic attributes: 5-year age group, sex, race (5 categories) and a set of 22 ICD9-defined common diseases. For method 2, the KPNC EMR was used as a proxy for a highly informative reference dataset. **Results:** The first method tended to yield very conservative estimates of risk: 9.5% of subjects in the disclosed dataset had unique phenotypic attributes, while 18% were in cells of size <5 and 24% were in cells of size <10. Factoring in characteristics of potential reference datasets, method 2, yielded substantially lower levels of risk: 2% of subjects were distinct, 4% in cells size <5, and 6% in cells of size <10. **Conclusions:** Assessment of re-identification risk of disclosed genomic-phenotypic data is complex, involving differing stakeholders' perspectives, attack types, and characteristics of both the disclosed data and the surrounding information environment. However, reasonable assumptions can be made which allow quantitative estimates of risk, and suggest strategies for risk reduction.

Keywords: Confidentiality; Data Sharing; RPGEH

doi:10.3121/cmr.2013.1176.ps3-13