Personal Genomics: Establishing the Scientific Foundation for Using Personal Genome Profiles for Risk Assessment, Health Promotion, and Disease Prevention
December 17 – 18, 2008
Bethesda, MD
DRAFT

DECEMBER 17, 2008

Session I: Genome Profiles, Risk Assessment, and Personalized Health: The Basics
Moderator: Gregory Downing, Ph.D., D.O.
U.S. Department of Health and Human Services (DHHS)

Welcome
Gregory Downing

The goal of this meeting was to discuss how genomic profiles can facilitate personalized health care by improving and personalizing risk assessment, health promotion, and disease prevention. Use of this information may lead to higher value health care by reducing costs, improving prevention and outcome prediction, and more precisely delivering interventions.

Successful use of genomic profiles depends on translation of scientific discoveries to clinical applications. Researchers and clinicians must create a clearer, more efficient pathway from discovery to clinical utility. Developing an accurate and understandable language for communicating risk based on genomic information is highly important both for clinicians who will communicate risk to their patients and for consumers who will need to understand their risk and possible options for mitigating it.

The scientific community is charged with providing leadership to move discoveries in the field of personal genomics to clinical utility. This will require innovation and creation of new pathways to efficiently incorporate the most up-to-date scientific discoveries into health care.

Robert Croyle, Ph.D., Division of Cancer Control and Population Sciences (DCCPS), National Cancer Institute (NCI), National Institutes of Health (NIH)

This meeting brought together investigators from numerous disciplines, private industry and nonprofit organizations, and federal agencies including NCI, NIH, and the Centers for Disease Control and Prevention (CDC) to discuss the use of personal genomics in healthcare. This meeting was designed to complement previous discussions held by the Institute of Medicine regarding the clinical utility and validity of cutting-edge genomic information. The NIH's strength lies in the field of basic discovery, with smaller efforts in translation and dissemination, although where the ultimate responsibility for synthesis, application, and dissemination of this information lies is unclear. Incorporating genomic information into clinical practice will require creation of clinically useful information that is meaningful to clinicians, providers, and patients.

Alan Guttmacher, M.D., National Institute for Human Genome Research (NHGRI), NIH

NHGRI led the highly successful federal effort to sequence the human genome, but the applications of this work to human health are less clear. The Institute is currently exploring ways to understand the contribution of genomic information to human health and diseases and has initiated a planning process to define research questions upon which to focus. NHGRI also plans to work with its sister agencies to determine ways to combine research efforts and also to identify issues of importance that may lie between the missions of NIH's Institutes and Centers but nonetheless affect public health.

Denise Simons-Morton, M.D., Ph.D., National Heart, Lung and Blood Institute (NHLBI), NIH

Investigators and clinicians must work together to determine the types of research needed to move from identification of genetically defined risk to understanding gene × environment interactions, and perhaps also interactions between genetics and treatments to permit more accurate targeting of interventions. This meeting focused on the utility of genetic information in clinical practice, aimed at improving patient and public health.

At this conference, research that will generate and test hypotheses relevant to practice was discussed. The types of hypotheses to test and study designs to use are complicated, and use of genetic information in practice requires special consideration because of privacy issues and other areas of sensitivity. Before genetic information can be used to improve health, the clinical importance and potential real world use of this research need to be determined. The goals of this conference included finding ways to advance the research needed to use genetic information to improve both individual and public health.

Personal Genomics: Establishing the Scientific Foundation for Using Personal Genome Profiles for Risk Assessment, Health Promotion and Disease Prevention
Muin Khoury, M.D., CDC and NCI

In 2007, *Science* magazine named genome wide association studies (GWAS) for the study of human genetic variation as the Breakthrough of the Year. In 2008, *Time* named the retail DNA test as the Invention of the Year. Although platforms employed for discovery work have been used to provide the public with their own genetic information, many are skeptical about the use of these platforms by retailers and consumers. The discovery aspect of personal genomics is well under way, but the translation of this information to clinical use has lagged.

A number of evidence gaps exist, such as establishing the amount of evidence needed to move a discovery into the public sphere, and development of definitions of clinical validity and utility for these discoveries. Improving use of genetic information for public health will require participation from investigators in a number of disciplines including epidemiology, clinical trials, communication, behavioral and social research, economics, and outcomes research. This meeting featured presentations on risk assessment, epidemiology, use of genomic profiles, and applications of genomic profiles to conditions such as cancer, heart disease, Alzheimer disease, and diabetes. Different models of translational research also were explored.

Personal Genomics: Review of Current Practices

Kenneth Offit, M.D., M.P.H., Memorial Sloan Kettering Cancer Center

Integration of genomic information into health care will help to tailor treatment and prevention strategies to the individual. However, accuracy, clinical validity, and clinical utility of genetic tests are necessary to realize this goal. A review of current practices reveals that the capacity to test a person's genome for disease-risk variants has raised more questions than answers, suggesting that continued research is needed before widespread dissemination of these medical technologies.

An evidence base to determine the clinical utility of the use of genomic information in practice is needed to avoid inappropriate treatments resulting from false positive results or a misunderstanding of risk. In the 1990's, the NIH provided leadership in the translation of cancer genetic testing in the United States; the Cancer Genetics Working Group recommended creation of the Cancer Genetics Network as well as the Cooperative Registries, one of whose purposes was to collect prospective data and create databases to inform clinical translation. Professional societies such as the American Society of Clinical Oncology provided training in cancer genetics to thousands of practitioners, and also called for increased regulation of laboratories to ensure quality of the testing. Similar federal and professional initiatives may be needed at the current time to most responsibly translate to practice current research in cancer genomics.

In light of increasing amounts of data generated by genome wide association studies (GWAS) for cancer risk, understanding the impact of each cancer-associated locus on risk is important. Most GWAS have identified variants that by themselves confer a very small increase in cancer risk, although these variants are found at high frequency in the population. The discriminatory accuracy of the genetic variants may be low and the best intervention for prevention may be unclear. Many of these loci are located in intronic regions, and thus functional significances and a putative role in cancer development are more difficult to discern. Variants may also have different abilities to predict risk in different population groups, or may have opposite effects in some populations.

Genetic associations have been found for several common diseases, including breast and prostate cancer, type 2 diabetes, age-related macular degeneration, schizophrenia, and myocardial infarction. Commercial entities have increasingly used published research findings to market direct-to-consumer (DTC) genetic tests for these conditions. Providing these tests via the Internet bypasses the medical community. Consumers may misunderstand or misinterpret the tests, which could lead to lifestyle choices or medical interventions. Ethical and legal issues may be raised by individuals participating in these studies without a full understanding of risks or without follow-up. Three instances were reviewed in detail where individuals tested by multiple commercial laboratories received widely conflicting genomic testing results for serious medical diseases. Genetic testing results also are prone to clinical misinterpretation, which could lead to error and injury; patients may not understand that a negative result does not always mean he or she will not develop a particular condition. The uncertainty and difficulty associated with conveying the results of genetic tests could lead to loss of trust in caregivers and added expense arising from unnecessary further testing.

Thus, a review of current practices of genomic screening for disease risk concludes that further prospective, population based, behavioral as well as laboratory-based research is needed. Additional needs have been stated by the Secretary's Advisory Committee on Genetics, Health, and Society (SACGHS), which advises the Secretary of Health and Human Services (HHS) on the human health and societal issues raised by the development and use, and potential misuse, of genetic technologies. SACGHS has recommended proficiency testing of genetic tests for which the Centers for Medicare and Medicaid Services (CMS) will provide reimbursement and that the Food and Drug Administration (FDA) oversee all laboratory tests, regardless of how they are produced. SACGHS also recommends guidance on regulation of clinical decision support systems; funding of a mandatory, publicly available, Web-based registry for tests by HHS; funding of a public-private partnership to evaluate the clinical utility of genetic tests; and identifying education or training deficiencies.

How Much Do SNPs Improve Models to Predict Breast Cancer Risk?
Mitchell H. Gail, M.D., Ph.D., Division of Cancer Epidemiology and Genetics (DCEG), NCI

Questions regarding the use of genetic data for risk prediction include whether the new information will help make treatment decisions and improve allocation of prevention and treatment modalities in a way that results in the best possible outcomes. The Breast Cancer Risk Assessment Tool (BCRAT) is a well calibrated tool, its discriminatory accuracy is somewhat low, i.e. it does not clearly distinguish between those who will and those who will not develop breast cancer. Recently seven SNPs associated with breast cancer have been identified and confirmed, but individually these SNPs have small impacts on overall risk. The purpose of this study was to determine how much adding these SNPs could improve the discriminatory accuracy of BCRAT and improve its performance for making various clinical and public health decisions.

To combine these SNPs with the BCRAT, I assumed no interactions among the SNPs or with factors in BCRAT and that the SNPS and factors in BCRAT were independently distributed in the population. Including the seven SNPs in the BCRAT increased discriminatory accuracy, measured by the area under the receiver operating characteristic curve, but the increase was less than from adding mammographic density. The addition of SNPs to the BCRAT also was tested to determine if it could improve decision-making regarding the use of tamoxifen. Among women aged 50-59 years, a woman's breast cancer risk must be above three times the average risk in order that the use of tamoxifen will confer more benefit than harm. Including the SNPs in the BCRAT improved the ability to predict beneficial use of tamoxifen by only 0.07 percent among women between the ages of 50 and 59 and by only 0.8 percent for women between the ages of 40 and 49. A related calculation showed that the decision whether or not to have a mammogram was only very modestly improved by adding SNPs.

The ability of the BCRAT plus SNPs to more precisely and effectively allocate mammography to women also was analyzed under the assumption that there was only enough money to provide mammograms to half the female population. Using the BCRAT to ration mammograms

prevented 63 percent of the maximum number of deaths that would occur if all women received mammograms. Including the SNPs prevented 66.7 percent of these deaths, a modest improvement. However, this calculation did not take into account the costs associated with collecting DNA and genotyping, which would reduce funds available for mammography.

Thus, including the seven SNPs in the BCRAT provided only very modest improvements over the BCRAT for providing discriminatory accuracy, deciding whether to use tamoxifen, deciding whether to have a mammogram, and allocating scarce mammogram resources. Moreover, the model with BCRAT plus seven SNPs needs to be validated using independent data on individuals. To achieve high discriminatory accuracy in this model would require hundreds of SNPs.

Genomics: What Kind Of Information Do Consumers Want?
Susan Friedman, D.V.M., Facing Our Risk of Cancer Empowered, Tampa, FL

When considering genetic testing, consumers want to know what conditions can be tested for, what the information will mean for both themselves and their families, whether they will have access to the information, where they may be able to go for more information, and whether the information is actionable.

Merely presenting a list of available tests may be confusing because many consumers choose which tests to take based on their perception of their own risk for a particular condition. Consumers also may not understand the differences between types of test, such as those that test for known mutations in known genes (e.g., BRCA1), and those that test for SNP variants that modify risk. Consumers are likely to decide on a testing protocol based on cost or availability rather than what is useful; in addition, if consumers believe they are at low risk for a particular trait, they are less likely to wish to pay for a test for it.

Most consumers want to know their risk for a disease as indicated by these tests, but do not clearly understand the difference between relative and absolute risk. They also are interested in knowing the risk for passing on a condition to their children, and whether their relatives may also be at risk. Genetic testing may be of use for treatment decisions, and consumers have indicated a willingness to take genetic tests that might predict response to treatment or help determine the best medications for treatment, prevention, or avoidance of adverse effects. Consumers also want to know if the results of a test will affect their quality of life or indicate lifestyle changes that could be made to mitigate risk.

Because of DTC marketing, consumers may have more information on potential genetic tests than their physicians, but lack the ability to interpret the results of the tests. Consumers (and in some cases physicians) also lack basic information on cost, coverage, and how to order the tests. They also may lack information concerning whether the information provides new options or affects future health-related decisions. Consumers also have interest in knowing the risks and benefits of options indicated by the test, whether experts agree on the best course of action, and whether insurance will pay for prevention or treatment options based on test results. Education

is needed to ensure that health care providers ordering genetic tests can understand and interpret the results, or can refer consumers to experts.

At the present time, consumers have not been asking how test results and health care options work together to affect personal health choices. There also is a poor understanding of at what stage of life different aspects of one's genomic profile should be explored and of interventions that might be most effective at particular ages. For example, there is little information for women testing positive for the BRCA1 mutation regarding whether they should take hormone replacement therapy to lower their risk of heart disease and whether oophorectomy in response to a positive BRCA1 test is advisable, considering the negative impact on cardiovascular disease (CVD). However, consumers are making health care decisions based on these tests, and steps should be taken to ensure they fully understand the test results.

Sharon Terry, M.A., Genetic Alliance, Washington, D.C.

Interest in genetic testing is evident in both diagnosed and not yet diagnosed groups of consumers, but at present there are no scientific or industry standards regarding use of and access to genetic information. Recently, consumers have had increasing access to information and also increased ability to share information easily. Given the privacy issues associated with genetic information, and possible misuse of and discrimination based on test results, consumers and producers of genetic information must decide whether free, unlimited access or a more moderate approach to information gathering is preferable.

To help consumers determine whether they should trust the information provided to them, a "toolbox" was created in partnership with the National Center on Birth Defects and Developmental Disabilities and the CDC. This effort will help consumers judge the trustworthiness of information by asking them to consider who provided the information, when it was provided, and how the information was gathered. Although credibility is a significant issue, consumers also wish to know whether the information will be clinically useful for them, which goes beyond strict definitions of clinical utility.

Personal Genomics: What Kind of Information Do Primary Care Providers Need? What Is the Role of Evidence-Based Guidelines?
Greg Feero, M.D., Ph.D., NHGRI

Approximately $2.26 trillion is spent on health care in the U.S. in 2007 (16% of gross domestic product), but life expectancy is lower than in a number of other industrialized nations. A significant portion of this spending pays for treatment of common, chronic diseases, which account for seven of every 10 deaths and $0.75 of every dollar spent on health care. Five percent of the population incurs 49 percent of all health care spending. One solution to this problem is to provide better primary care, which could prevent worsening of conditions such as diabetes, hypertension, and high cholesterol.

Despite the large number of genomic discoveries relevant to common disease diagnosis and management, the genomics revolution has not had a significant impact on health care,

particularly primary care. Although questions about analytic and clinical validity and clinical utility remain, the current approach to translation of genetic discoveries favors unfiltered access to these technologies. This approach has failed to include significant education of primary caregivers in the understanding and communication of genetic risk. Most patients will consult with their primary physicians about genetic testing; only approximately 0.18 percent of physicians in the United States are medical geneticists, and genetic counselors are similarly scarce. Despite educational outreach attempts by the genetics community, most internists, family practitioners, pediatricians, and obstetricians/gynecologists rated their knowledge of genetics of medical conditions as average or poor. Although most medical workers agree that familiarity with statistics and risk communication are important for their practice, only a small percentage felt that their training in these fields was sufficient and less than half felt confident about communicating quantitative risk. Establishing clinical guidelines for communicating genetic and risk information may mitigate this situation, but guidelines have different standards and are not always adhered to.

As more and more genetic information becomes available, the health care community must take care to avoid "gizmo idolatry," which refers to the idea that a more technological approach is always better than a less technological approach, even in the absence of evidence. Primary care physicians are vulnerable to "asymmetric knowledge," meaning they can be convinced that a test is more useful than it really is if they are informed about the test by a more knowledgeable person. Guidelines for use of genetic information based on health outcomes are critical, but further efforts are needed to effectively incorporate this information into health care.

DISCUSSION

Dr. Geoffrey Ginsburg commented that incorporation of the seven or eight risk alleles into the Gail risk model may identify different patients who need increased surveillance than would be identified using only the model, although the absolute benefit of using the SNPs appears small. He asked whether patients would be reclassified regarding the need for surveillance if the markers were incorporated into the model. Dr. Gail answered that the area under the curve (AUC) for adding the SNPs into the model was 0.632, and the AUC for the model alone was approximately 0.607. This indicates that the SNPs add information, but their inclusion did not necessarily result in reclassification. In addition, inclusion of the SNPs did not always increase risk; in some cases risk was reduced by their inclusion.

Dr. Sharon Kardia noted that the scientific community appears to be modeling use of risk alleles on models used for the genetic of inborn errors of metabolism, but it is more likely that these SNPs will be involved in interactions. One question to address is whether consumers will be able to understand concepts related to gene × environment interactions. Ms. Terry agreed that it may be difficult for consumers to understand gene × environment or gene × behavior interactions, but in general, consumers understand risk well. Because large, complex studies will be required to understand these interactions, the development and use of large databases, along with attendant privacy issues, must be addressed. Dr. Friedman added that for consumers to understand the larger effects of the environment, they also must understand that the effects of

gene × environment will likely be different for subgroups and individuals, which will complicate the decision-making process.

Dr. John Ioannidis asked how the Gail model was affected if the uncertainty associated with use of the SNPs was taken into account, and whether people's attitudes toward different risks (i.e., hip fracture versus endometrial cancer) would affect their response to risk knowledge. Dr. Gail answered that cost-benefit analyses had not been extensively performed. When using the model to determine the benefits of tamoxifen use, he included only life-threatening risks. Dr. Khoury noted that the AUC for the Gail model is much smaller than the AUC for heart disease risk models. He asked how many SNPs would be needed to warrant incorporation of them into the model and what studies would be needed to show that they improved clinical practice. Dr. Gail said the AUC he reported was age-specific and that if age was credited in computing AUC for the Gail model, the AUC would be larger. He noted that the AUC is not an optimal measure to use because of its retrospective quality. He would prefer to know positive predictive value for a given risk, because this provides a better prospective view. Identifying real decision problems in medicine, addressing the costs and benefits associated with these decisions, and then introducing a risk model to see if it can improve decision making would be a better approach. Development of reclassification tables is an active area of research in risk modeling. It is possible to develop two different estimates of risk when using two different models, and then cross-classify an individual to determine how risk measured in the first model relates to risk measured in the second. Only the margins of reclassification tables are relevant for decision-making accuracy.

The panelists were asked to offer suggestions for short-term research to improve application of genomics to population health. Dr. Offit suggested developing a Request for Applications (RFA) for research to translate genomics into practice. Such an RFA could include studies that use existing trials that have collected genetic information. Dr. Friedman said that research on the information consumers and physicians receive regarding genetic risk are needed. Dr. Feero suggested that federal agencies develop a coordinated approach to addressing this issue. Dr. Offit added that public-private partnerships also should be encouraged. Ms. Terry asked for engagement of all stakeholders and development of a national biobank or database. Dr. Gail said that cohort data were needed to validate and determine the utility of including genetic information in risk models. Clinicians also need to become more involved in medical decision making.

Session II: The Scientific Foundation for Which Genetic Variants Should Be Included in Genome Profiles: The Credibility of Genetic Associations

Moderator: Teri Manolio, M.D., Ph.D., NHGRI

Navigating the Epidemiology of the Human Genome

Marta Gwinn, M.D., M.P.H., National Office of Public Health Genomics, CDC

The HuGE Navigator has been developed to help navigate literature related to the epidemiology of the human genome. This tool was developed because it can be difficult to search PubMed effectively for articles featuring topics such as prevalence, associations, interactions, and genetic tests. The HuGE navigator combines both human and artificial intelligence searching to create a

database of citations on human genetics/genomics and human genome epidemiology. As of December 11, 2008, the knowledge base contained 34,208 genetic association studies, 863 meta-analyses, 243 GWAS, 3,888 genes, 1,958 Medical Subject Heading (MeSH) disease terms, and 5,645 common variant names matched to rs numbers.

Although the extent of this literature is impressive, many of the results reside in databases and are seldom used. The University of California – Santa Cruz Genome Browser allows the user to view an association for a particular SNP and also find other SNPs surrounding the SNP of interest. This is then linked to publications that may contain information on population measurements or disease term associated with a given gene located in the area of interest. The HuGE Navigator has two integrated components, Genopedia, which can be used to look up gene-disease association summaries by gene, and Phenopedia, which can be used to look up gene-disease association summaries by phenotype. Integration of genetic and phenotypic information also will be useful for analyzing gene × environment interactions, although currently there are few publications on this subject matter.

Questions

Dr. Kardia asked if there were ways to integrate the HuGE Navigator with other databases, for example, those that have information on signaling pathways. This might help identify sites where the environment might have an impact on biological processes. Integration with gene expression information also might be useful. Dr. Gwinn answered that because the nomenclature for genes is more standardized than that for diseases or SNPs, it is possible to connect to other databases at the level of the gene. Links to other gene-based databases exist, but a main focus of this project is to integrate genetic and epidemiologic points of view.

Dr. Manolio asked why there are so few articles on gene × environment interactions. Dr. Gwin answered that there are few articles because the research is very complicated; it is difficult to measure environmental exposures, especially over a lifetime. The data also are difficult to analyze because of the small effect sizes. Some GWAS use populations that have associated exposure data, and methods should be developed to more accurately measure these exposures and determine their analytic validity.

Genome Wide Meta-Analysis: Promise and Pitfalls
John Witte, Ph.D., University of California at San Francisco

To make the best use of genome profiles, relative risk estimates rather than p values are needed. However, it is unclear whether the estimates should come from the results of the initial GWAS, the replication studies, or pooled or meta-analyses. Using pooled data from all GWAS of a particular condition would probably be best, and therefore a database to help organize the data to permit meta-analyses is needed; something similar to the HuGE Navigator would be useful. The goals of a meta-analysis include combining findings across studies to develop the "best" estimates of association and determining if and why differences across studies exist. However, meta-analyses of GWAS are complicated. Publication bias can be problematic because negative studies likely will not be published. However, if those performing the meta-analyses use only

the GWAS "discovery phase" data and are aware of all existing GWAS, published or unpublished, on a particular condition, publication bias is less of a concern. If investigators wish to make use of GWAS data and focused replications, they will need to search for all data related to the condition.

The alleles identified by most GWAS tend to be common in populations but have low penetrance or are associated with small effects on risk. Low frequency variants with intermediate penetrance may be more important for determining risk, and often multiple GWAS must be combined to find these variants. For example, analysis of SNPs associated with prostate cancer identified a SNP on 10q11 with high statistical significance; however, this SNP was not in the top 10,000 hits of the initial GWAS. It was identified because a large scale replication of the GWAS was performed. This SNP also would have been identified if data were combined across GWAS. Pooled analyses of individual-level data permits analysis of independent, interacting, and multi-phenotypic effects, but these data are rarely available.

The need to impute data across different genotyping platforms and then combine the data complicates meta-analyses of GWAS. Population stratification adjustments and analyses to distinguish between observed and imputed data also will be needed. Logistical issues, such as ensuring that all SNPs are correctly oriented, also require attention. Standard approaches for combining GWAS results include using Z scores weighted by sample size and inverse variance weighted odds ratios. These assume no variation between studies (fixed effects model). However, one goal of GWAS meta-analyses is to determine if and why differences exist across studies. Tools are needed to allow for heterogeneity of results. For example, although the association of 8q24 with prostate cancer appears fairly similar across studies, heterogeneity could be tested for if a random effects model was used. Most GWAS meta-analyses have focused on a fixed effects model, but both fixed and random effects models need to be used to assess differences across studies.

Questions

Dr. Witte was asked to explain whether researchers should have confidence in a meta-analysis that identifies a strongly associated SNP when that SNP was not a strong hit in the initial GWAS and whether this pointed to problems intrinsic to combining analyses. Dr. Witte answered that this situation arose because of the effect size of the SNP in the initial GWAS. Many true associations may have small effect sizes. Dr. Manolio asked if the initial study was designed to replicate the top 25,000 SNPs because of anticipated small effect sizes. Dr. Witte answered that most GWAS identify SNPs with small effect sizes, and access to larger amounts of data permits modeling to improve risk estimates. Meta-analyses may not significantly change estimates of risk, but the best possible estimates should be used if the goal for use of these SNPs is disease prediction.

Assessing Cumulative Evidence in Genetic Associations
John P.A. Ioannidis, M.D., Ph.D., University of Ioannina School of Medicine, Ioannina, Greece

It is difficult to assess the cumulative evidence generated by genetic association studies, particularly given small effect sizes. Guidelines for assessing cumulative genetic association evidence have been developed that involve grading the evidence as strong, moderate, or weak across three criteria: amount of evidence, replication, and protection from bias.

The amount of evidence can be defined in terms of sample size of the least common genetic group among those compared (which could reflect participants or alleles, depending on the model), study power, false discovery rate, or Bayesian credibility. Defining whether sufficient replication has been performed is difficult. A recent analysis of data in the NHGRI GWAS catalog found 233 associations for binary outcome phenotypes with p values of less than 10^{-5}; 142 with a p value less than 10^{-7}; and 87 (39%) with a p value of less than 10^{-10}. Most loci discovered in GWAS need further exact replication with more large scale evidence before they can be considered sufficiently reliable to serve even as simple markers. Consistency of replication also must be considered. The highest ranking should be given to replications that include at least one between-study meta-analysis with little between-study inconsistency; heterogeneity of the data must be considered if the data are to be used for testing. Comparing heterogeneity in candidate gene studies to GWAS found less heterogeneity for the candidate gene studies, but the amount of heterogeneity relative to effect size was similar across both study types.

Protection from bias can be difficult to determine. Typical biases include bias in genotyping, phenotype definition, population stratification, and selective reporting. For the most highly ranked studies, bias, if present, would affect only the magnitude but not the presence of the association. A research finding cannot reach credibility greater than 50 percent unless all biases are less than the pre-study odds. Bias checks for retrospective meta-analyses include effect sizes less than 1.15-fold from the null effect; association lost with exclusion of the first study or with adjustment for Hardy-Weinberg equilibrium; or evidence of excess single studies with formally statistically significant results. Bias checks for a prospective consortium analysis include the magnitude of the effect size and small study effects; an excess of studies with significant findings is not an issue, provided there is no selective reporting. Once the evidence has been graded, calibration of the credibility for this association must be performed. Most epidemiological associations have low credibility based on the rules used to analyze GWAS credibility. Determining whether effects are different among individuals, e.g., between males and females, also must be addressed.

Other issues hampering assessment of cumulative evidence in GWAS include the lack of publicly available data and the less than optimal reproducibility and repeatability of many GWAS. Another problem is determining how to handle conglomerate evidence, given the existence of scattered studies and scattered single GWAS and questions concerning whether to consider only the highest level of evidence or all evidence. Determining who will summarize the evidence also must be addressed. This is best done by experts in a given field, and creation of a "network of networks" of such experts may be beneficial.

Assessment of the cumulative evidence on genetic associations focuses on amount of evidence, consistency of replication, and protection from bias. The evidence generated by GWAS is often

uncertain and tenuous, and the uncertainty is underappreciated. The evidence is likely to become more reliable when its integration is transparent and anticipated prospectively by all involved partners. Discovery and integration of evidence ideally should proceed in parallel.

Questions

A participant noted that for conditions such as Parkinson disease, highly powered GWAS have not been performed, nor have solid replications. The ability to replicate should be part of the criteria for validity. Dr. Ioannidis agreed that there is a large range of credibility and replicability across GWAS. Replication of the study should be significant, independent of the original study.

Dr. Khoury asked how sound results should be before they are included as part of genomic profiling. Dr. Ioannidis answered that regulators must make these decisions. Considering efficiency and likelihood of utility, any associations included in a genomic profile should have extremely high credibility. Ms. Terry asked how existing heterogeneity could be incorporated. Some large studies have less power than small ones, and there is no way to incorporate factors such as age differentials with respect to power and effect size. Dr. Ioannidis answered that it is difficult to incorporate covariates unless they are known to be essential to the study. It would be preferable to perform meta-analyses one SNP at a time during the discovery phase. Once a SNP with high credibility is discovered, questions can be asked about the amount of heterogeneity associated with it, from what the heterogeneity arises, and whether there are any population features to test as part of the exploratory analyses.

What Variants Are Included in Genome Profiles and How Are Disease Risks Calculated? Is There an Industry-Wide Standard?

Amy Miller, Ph.D., Personalized Medicine Coalition, Washington D.C.
Jeff Gulcher, M.D., Ph.D., deCODE Genetics
Andro Hsu, Ph.D., 23andMe, Inc., Mountain View, CA
Michelle Cargill, Ph.D., Navigenics, Redwood Shores, CA

The Personalized Medicine Coalition represents patients, health care providers, and industry, all of whom are interested in targeted therapeutics. Navigenics, 23andMe, Inc., and deCODE have worked together to develop standards for genomic studies. Analytical and clinical standards include selection of a SNP genotyping platform in a Clinical Laboratory Improvement Amendments (CLIA) regulated laboratory with high accuracy; a requirement for replication of SNPs chosen for annotation of risk in multiple powered studies and derivation of their odds ratios (ORs) from large datasets; use of methodologies to convert from the reported OR (or genotype-specific OR) to risk compared to the general population; and assumption of a multiplicative model for both the allelic risk at each marker and when combining markers to define overall risk unless the data support a better model.

All three companies have found variants within 9p21 that are associated with myocardial infarction (MI)/coronary heart disease (CHD); this is the only region to show significant

association with this condition in the four GWAS published to date. All markers cluster within a single linkage disequilibrium block. This result has been replicated in 25 Caucasian and five East Asian populations; no effect was found in African populations. The replication included over 30,000 patients and 60,000 controls and included several prospective studies. Twenty-one percent of the population was homozygous for the variant, and these individuals have a 2.0-fold risk for early MI compared to non-carriers; the association is independent of known risk factors including family history, low-density lipoprotein (LDL), hypertension, diabetes, obesity, and smoking. All three companies converted the allelic OR to relative risk to the general population. Addition of the 9p21 variant to the Atherosclerosis Risk in Communities (ARIC) and National Population Health Survey (NPHS) prospective cohorts resulted in a significant increase in the accuracy of MI prediction. Eighteen percent of patients in the intermediate and intermediate-high risk categories were reclassified; inclusion also resulted in a change in LDL target. Because these companies have access to large datasets, they can analyze the effect of multiple variants together and determine whether the variants interact. For example, a multiplicative model was used to determine relative risk for prostate cancer based on eight validated genetic markers. Increases in risks for prostate cancer ranged from 0.4- to 5-fold.

To calculate genotype-specific risk, all three companies begin with single SNP ORs from existing GWAS and then incorporate later studies and meta-analyses when possible. Despite different starting populations, similar estimates of genotype-specific risks are generally obtained. deCODE, 23andMe, Inc., and Navigenics intend to work together to standardize the presentation of genetic risk information. The companies will investigate dissimilar risk numbers more thoroughly. However, the scientific community must work to establish standardized baseline numbers. Transparency with regard to references used, backend calculations if number is not reported in the text, and explanatory text should be maintained.

DISCUSSION

When asked about the impetus for these companies to work together to develop standards, Dr. Gulcher described a SACGHS meeting at which major differences in estimated lifetime risk based on published results were discussed; this meeting represented a starting point for coordination and ensuring consistent statistical calculations. Dr. Khoury noted that translating association data for use in clinical models can be done only for diseases with registries; for other diseases, investigators will need to depend on publications, which are not standardized with respect to disease reporting. Given the small effect sizes of the SNPs, uncertainty around the estimates of lifetime risk may become large. In addition, because diseases change over time, case ascertainment varies, and other risk factors may have a role, it is difficult to see how weak association data can improve predictive models. Dr. Gulcher explained that deCODE does not present risks based on only one variant. He agreed that other risk factors must be explained when presenting lifetime risk. He emphasized that these companies are defining genetic risk as part of lifetime risk; lifetime risk is presented to provide perspective on how genetics affect it.

Dr. Miller noted that providing medical information on the Internet, through sites such as WebMD, has improved consumer care. A participant countered that seeking medical information comprises 35 percent of Web searches, but no studies have proved that access to

genetic information improves health care. Dr. Cargill said that all three companies are interested in outcome testing. Dr. Witte said that he was concerned about how company data and results from meta-analyses are controlled, which may hinder followup. Dr. Cargill noted that the three companies also have trouble obtaining data sets for analyses. Dr. Witte said that the ORs and standard deviations for every SNP should be available, if not the individual data. All agreed that managing data access can be difficult when a study involves large numbers of collaborators.

Dr. Ioannidis noted that providing an uncertainty estimate along with the main risk estimate could be useful. Dr. Cargill said that her company considered including confidence intervals when presenting risk results, but raw data are needed for this and it can be difficult to obtain such data. Dr. Gulcher said that deCODE makes confidence interval estimates available in the published literature.

Dr. Offit asked how population heterogeneity issues were addressed, given that most data are generated using Northern European populations. Dr. Cargill referenced a 2004 publication on ORs for validated genetic associations across ethnic groups that found generally low heterogeneity in the OR, but considerable heterogeneity in allele frequency. More work on populations with mixed ancestry is needed, as well as further genetic population analyses concerning how risk is affected by mixed ancestry. Dr. Hsu said that 23andMe, Inc. includes SNPs if they were replicated in populations that correspond roughly to ancestry. Dr. Gulcher said that deCODE does not provide risk based on specific ethnic group unless the GWAS was replicated in that population.

Session III: The Scientific Foundation for Establishing Clinical Validity and Utility of Genome Profiles—Part 1

Moderator: Kay Wanke, Ph.D., M.P.H., Office of Behavioral and Social Sciences Research, NIH

Inter-Disciplinary Evaluation of Genomic Profiles of Clinical Validity and Utility
Steven Teutsch, M.D., M.P.H., Merck & Company, Inc., West Point, PA

Translating genetic discoveries to improvement in health care will require evidence-based guidelines to connect research and practice. The information needed for multidisciplinary evaluation of genetic tests can be defined by ACCE—Analytic validity, Clinical validity, Clinical utility, and associated Ethical, legal, and social implications.

Clinical validity is defined as the degree to which a laboratory test accurately distinguishes between those with and without a health condition; it is characterized by sensitivity, specificity, positive predictive value, and negative predictive value. Clinical validity also defines the ability of a test to detect or predict a phenotype or particular clinical outcome. Sensitivity is defined as the proportion of positive test results in individuals who have the phenotype, while specificity refers to the proportion of negative test results in individuals who do not have the phenotype. Discriminative accuracy combines sensitivity and specificity and determines their accuracy. Clinicians often refer to positive predictive value (sensitivity) and negative predictive value (specificity). Predictive values depend on the definition and prevalence of the phenotype,

characteristics of a tested population; the penetrance of the gene(s) involved in the phenotype; and genetic heterogeneity. As an example, a screening test with 99 percent sensitivity and 95 percent specificity for an allele with 1 percent prevalence will have a positive predictive value of only 17 percent; there is a less than 1 in 5 chance that this test will predict phenotype for this disease.

Clinical utility refers to whether incorporation of a test into practice improves clinical management, taking into account the risks and benefits associated with including the test in practice and the likelihood of improved health outcomes. Ethical, legal, and social issues are considered a part of clinical utility, as they take into account issues such as the severity of a condition and available therapy and how these impact offering a test to the public. This part of evaluation of a test also considers whether negative consequences (such as discrimination, stigmatization, or health disparities) could result from testing and whether effective safeguards have been established. This area also covers issues related to consent and ownership of samples or discoveries arising from use of samples.

The evidence threshold for implementing a practice also must be considered. Setting the bar fairly low allows discoveries to be translated to practice sooner, but often there is inadequate information on validity or utility. Because of this, payers may not wish to cover testing costs and there also is potential for harm. However, early introduction of new discoveries can stimulate innovation. Setting the bar high is more likely to result in a test with validity and utility, but there is lower incentive for innovation and potentially diminished benefits because the test may not be optimally used.

Data on the use of CYP450 testing in adults with non-psychotic depression to optimize treatment with selective serotonin reuptake inhibitors (SSRIs) were evaluated to assess clinical utility. Recommendations were made by rating data sources and using research conclusions to determine the level of certainty in the results; the magnitude of the net benefit also was considered. CYP450 genotypes were not consistently associated with the outcomes of interest, including clinical response to SSRI treatment or adverse events occurring as a result of treatment. CYP450 testing thus was not recommended for this clinical situation.

To make the best use of genomic discoveries, the clinical validity and utility of genetic tests must be understood to inform decision-making. Standards for evaluating these tests must be established. The ethical, legal, social and economic implications of using these discoveries must be included in test evaluation.

How Do We Assess the Added Value of Genetic Information in Predicting Disease?
Cecile Janssens, Ph.D., Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands

Because the public has shown significant interest in using genetic tests to predict their risk for common diseases and implement interventions to mitigate risk, the results of genetic testing must be presented in a way that is understandable. Information on the accuracy of the risk estimate

(calibration), the disease risks of others (risk distribution), and the risk change compared to prediction without the test result (risk difference and reclassification) should also be provided.

Calibration and risk distribution are aspects of clinical validity. Calibration helps judge if the predicted risks are correct; predicted and observed risks should agree. Analysis of calibration in recent empirical studies of genetic profiling found poor calibration at the high end of risk distribution. Calibration is especially important when predictions are based on models. Issues that should be addressed include determining whether a multiplicative model is the correct assumption, whether effects are independent, and whether effect sizes (OR) obtained from various studies apply to the population tested (of higher importance when ORs are generated from case-control series, rather than prospective population-based studies). Validation investigates the predictive value of a test using an independent dataset. Validation is always important, but less so when risk estimates are obtained from other studies (in this case, calibration is equal to validation).

Risk distribution refers to the utility of knowing one's risk dependent on the risks of others; if all predicted risks center around average, the test is not useful. A fairly broad distribution of risks is needed to provide good discrimination; a large amount of overlap in the distribution provides limited or no discrimination. AUC is generated by plotting all sensitivity-specificity combinations for all possible cut-off values of the predicted risks; the higher the AUC, the better the discrimination and prediction provided by the test.

Risk distributions often are used in clinical practice to create categories for clinical decisions (such as to treat or not). Reclassification refers to the percentage of individuals who change risk categories when prediction models are updated, for example, when genetic variants are added to a model based on traditional risk factors. If patients do not change treatment categories, updating the prediction model is not useful. Prediction of type 2 diabetes based on three models was analyzed. The first model predicted risk based on TCF7L2 polymorphisms, the second used 18 different polymorphisms, and the third used age, sex, and body mass index (BMI). When the second model was included with the first, 32 percent of participants were reclassified; 28 percent were reclassified when the third model was included, some back to their original category. The usefulness of learning about every risk update and reclassifying patients accordingly must be considered and the impact of updating risks assessed.

Scientific Evaluation of the Impact of Interventions Associated with Genetic Risk Factor Information

Barry R. Davis, M.D., Ph.D., University of Texas School of Public Health

A well-established hypertension treatment algorithm exists, along with compelling indications for different treatment options. Clinical experience has shown that different treatments work better or worse in specific population groups; this is an issue that needs to be addressed further with regard to genetic testing. The Genetics of Hypertension Associated Treatment (GenHAT) study was created to analyze the use of genetics in guiding hypertension treatment. GenHAT is an ancillary study of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT), which enrolled 42,418 participants and compared three different

hypertension treatment options: calcium channel blockers, angiotensin-converting enzyme (ACE) inhibitors, and alpha-blockers to the standard treatment, diuretics. The primary outcome was fatal CHD and non-fatal MI. Secondary outcomes include mortality, stroke, and heart failure.

GenHAT used blood samples taken as part of ALLHAT to isolate DNA for genetic analysis. Gene-treatment interactions were analyzed for three genes: α-adducin, ACE insertion-deletion (I/D) gene, and NPPA. The gly460trp polymorphism in α-adducin was not an important modifier of hypertension treatment choice on cardiovascular risk, but results also suggested that women Trp460 carriers may have increased CHD risk if treated with chlorthalidone versus amlodipine plus lisinopril. Similarly, no strong association between type of antihypertension medication and CHD or other major ALLHAT outcomes was found across different ACE I/D polymorphisms. Carriers of a minor NPPA allele had more favorable outcomes when treated with chlorthalidone versus amlopidine. However, the associations were not strong.

The clinical utility of genotyping to better target hypertension treatment remains to be determined. Pharmacogenetic studies of complex conditions such as hypertension and heart disease are challenging, but may identify clinically useful genetic markers. GenHAT findings need to be replicated and gene × gene and gene × environment interactions more thoroughly explored. Cost-benefit analyses also should be conducted, as well as determining the best ways to move these findings into practice, which may be complicated by resistance from industry and clinicians to changing long-established treatment guidelines.

DISCUSSION

Dr. Lauer asked whether quantitative ways of describing the calibration ability of a model existed. Dr. Janssens answered that simple reclassification tables can be used for this purpose. She further explained that for risk prediction research, models initially are developed using derivation datasets. The same datasets can be used to evaluate discrimination and calibration. If the model discriminates and calibrates well, a different data set is used to test the prediction model. Dr. Ioannidis noted that if AUC does not change, thresholds can be adjusted to provide better reclassification. He asked Dr. Janssens to explain how strict thresholds for reclassification are. Dr. Janssens replied that genes with small effects can result in significant reclassification if the threshold is in the middle of the risk distribution, since most people have average risk for a condition.

Dr. Khoury noted that the meanings of clinical validity and utility need to be more clearly defined to improve outcomes, particularly for genome profiles that have little discriminatory power and little effect on reclassification. Dr. Davis noted that there is a great deal of research in progress to determine how to incorporate genetic information into clinical trials, such as was done for GenHAT. To apply genetic information to a specific population, decisions must be made concerning whether to treat participants according to genotype or stratify to interventions based on genotype. Better ways of using existing data are needed, as is information on biological plausibility and mechanism of action. Dr. Mark Greene suggested that prospective cohorts that have banked DNA dating from the time of study enrollment may be useful. Because

the effects of genetic variants are small and many of the events studied are rare, it will be necessary to have large study populations; risk factor and covariate information also will be needed. Dr. Ioannidis suggested developing simple randomized trials for serious and long term outcomes and to consider trials linked to registries. Dr. Kardia suggested using free living clinical populations to gather outcomes information and also explore whether knowledge of genetic risk influences outcome.

Dr. Gail speculated on whether use of genetic technologies would increase exponentially in the absence of evidence of benefit, similar to what has occurred in the field of cardiovascular care. A systematic assessment method is needed, along with a fairly high standard for evidence of benefit; this standard may need to be set differently for different genetic diagnostic technologies. Dr. Greene agreed that a way to perform timely assessments of utility of genetic tests and provide the results to providers and consumers is needed. A number of cancer-related diagnostic and screening tools are on the market, but there is no way for consumers and providers to determine if the tests have been validated or have clinical utility. NCI does not systematically evaluate and review these tests, and although the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) initiative is well-designed, the turnaround time is long.

A participant suggested developing a prioritization scheme for evaluation of genetic tests. He also asked whether providers might be more willing to prescribe innocuous drugs, such as those for CVD risk, without rigorous demonstrations of utility. Dr. Davis said that this was a significant public health issue because the ability to choose drugs that would better mitigate CVD risk would have a large effect on public health. Dr. Gail noted that combining information on somatic mutations found in tumors with information on the effects of cancer interventions also could help better target treatment therapies. Genetic information also would be useful for determining the impact of tamoxifen on a wide range of health outcomes. Dr. Greene said that the current clinical trials system could provide many opportunities for studying genetic modifiers of treatment outcomes, provided that DNA is collected. A participant asked whether a change in assessment of risk would provide information important to the patient. Reclassification tables are academically interesting, but because thresholds currently seem to be arbitrarily set, changing these thresholds might not be meaningful.

Dr. Downing noted that consideration of clinical utility could apply to payment and coverage decisions; federal reimbursement decisions will need to address quality and practice guidelines and effectiveness. He suggested developing a framework that professional societies could build on to assess the use of genetic testing for risk prediction and treatment decisions. Dr. Teutsch said that he has been working with several American health insurance plans on coverage decisions to understand how concepts such as level of certainty and magnitude of effects will affect how different technologies can contribute to improvements in health care. Establishing an evidence standard will be important, with the understanding that the evidentiary bar may be higher or lower for different conditions depending on the severity and possibility of interventions. Of most interest are interventions that can make substantial differences at a reasonable cost.

Session IV: The Scientific Foundation for Establishing Clinical Validity and Utility of Genome Profiles—Part 2

Moderator: Mark Greene, M.D., DCEG, NCI

Personalizing Genomic Information

Angela Trepanier, M.S., C.G.C., Wayne State University School of Medicine

Genetic counseling practices will change as counselors begin to counsel patients about risk based on genetic profiles rather than single genes. Current counseling approaches include the teaching model and the counseling model. The teaching model contends that patients can make their own decisions about risk management options once they are educated. Assumptions about human behavior and psychology are simplified and minimized, and cognitive and rational processes are emphasized. The task of the counselor is to provide information as impartially as possible. The counseling model includes complex assumptions about human behavior and psychology that are addressed. Clients are perceived to seek counseling for complex reasons, including information, validation, support, and reduction of anxiety. The goals of the counseling model include bolstering a sense of competence and feeling of control over one's life, relieving psychological distress, and helping find solutions to specific problems. Education is viewed as a means to achieving these goals. Leaders in the field of genetic counseling recommend a combination of both approaches, and call for flexibility to apply the appropriate model for any given client. The key goal of genetic counseling is to facilitate decision making. These ideas will need to be integrated into primary care, given that people other than those with a history of genetic diseases will be interested in obtaining information about their genetic profiles.

Nondirectiveness has historically been a guiding principle of genetic counseling because of the need to promote informed, autonomous decisionmaking and distance genetic counseling from eugenics. Problems with incorporating nondirectiveness include difficulties with remaining completely impartial, the limits it places on full use of counseling techniques and engagement, its questionable efficacy with regard to informed decisions, and a lack of applicability to certain situations. At a recent genetic counseling workshop, a need for a flexible approach to genetic counseling with varying adherence to nondirectiveness based on client/family needs, values, clinical circumstances, and desired counseling outcomes was recognized.

The best approach for communicating risk based on genomic information needs to be determined. Genomic profiles contain large amounts of information, in contrast to genetic information, which tends to be targeted. Genetic counseling is usually sought when there is a family history of a condition, whereas risk conferred by a genomic profile may not be associated with family history. Genomic profiling also tends to assess risk for conditions for which there may be interventions to mitigate risk, which is not often the case for genetic information. An effective way to communicate the information to promote informed decision-making, healthy behaviors, and perceived control and low distress must be developed. Effective communication of genomic profiles will likely require a blended teaching/counseling approach. As part of determining clinical utility, studies that evaluate how to maximize perceived value and intent to act are needed, which will contribute to development of evidence-based counseling strategies.

Major Trends in Health Marketing: How Do They Apply to Genomics Communication?
Jay M. Bernhardt, Ph.D., MPH, CDC

Evidence-based, strategic, consumer-centered, and cutting edge approaches should be applied to health marketing to improve public health. Several current trends will affect health communication related to genomics.

First, the number of places from which people can seek information has increased significantly. The average informed person reads or listens to seven sources of information daily, and approximately 25 percent of all media use time is spent using multiple sources simultaneously. This multi-tasking means the consumer accesses more media in less time. Much health information is received from mass media, although other people are considered a more credible source of information. This can result in health consumers feeling that they have more information than ever but also feeling overwhelmed. This can lead to a change in information-seeking behavior; information-seeking tends to increase, but so do the numbers of information blockers or avoiders. These factors create challenges for disseminating information to people in a way that influences their decisions. To increase personal genomics awareness, multiple channels should be used to reach the public; the message should be mediated through partners; information should be accessible; demand should be generated; and the most effective communication channels should be researched.

A second trend reflects the shift in information sources from experts and authorities to communities and peers (e.g., social networking sites and blogs). This reflects a horizontal, more egalitarian exchange of information that is attractive to those who wish to receive their information from someone they can relate to. Expert communication needs to be combined with peer-to-peer communication; an example of this might be expert information from a pharmaceutical manufacturer combined with consumer comments from those who have used the drug. Personal genomics awareness can be increased by providing information through both traditional media and use of experienced peers as spokespeople. Online video content also is becoming increasingly important as a way of providing good content that is shareable. Online sources also can provide qualitative data by creating opportunities to learn about people's opinions about a product or treatment approach.

The third trend is characterized by tailoring and personalization of communication. Mass media-based approaches to behavior change have been found to have limited effectiveness to prompt behavior change. In contrast, tailored messages are more likely to be perceived as relevant and salient and have effectively changed diet, physical activity, immunization participation, and other behaviors. Mobile health communication, in which tailored messages related to health are sent to a person's mobile phone, is becoming increasingly important as a way to change behavior, including risky behaviors such as alcohol use. Genomics information, tests, and tools should be provided where, when and how people want and need them. Information should be tailored based on demographics, psychosocial beliefs, communication factors, and individual abilities and preferences.

Does Genetic Information Change Behavior?

Colleen M. McBride, Ph.D., Social and Behavioral Research Branch, NHGRI

Genetic information may increase perceptions of susceptibility and increase motivation to implement behavior changes to mitigate risk. Conversely, genetic information may decrease a person's sense of control and confidence in his or her ability to change in a way that will change risk. Cognitive abilities, disposition, attitudes, and beliefs influence how people perceive information. Context, e.g., family history, also affects perception.

Many genetic variants associated with a particular condition have only a small affect on risk; thus, presenting this information might not greatly affect motivation. Large, randomized controlled trials on smoking performed in the 1990s that included information about gene variants that affected lung cancer risk did not affect participants' abilities to quit smoking. A new study, the Family Risk and Lung Cancer Study, contacted smokers who had family members diagnosed with lung cancer to explore their attitudes about genetic testing. Some believed the results of testing would motivate them to quit smoking, whereas a smaller but still significant number believed that results showing they were at low risk for developing lung cancer (based only on a single gene test) could enable them to continue smoking. Smokers who reported the strongest motivation to quit smoking were the most likely to seek testing.

Confidence about managing weight based on the presence or absence of an "obesity mutation" was explored in obese women. Those without the mutation felt a slightly greater confidence in their ability to lose weight. Slightly more women with the mutation felt that this knowledge would de-motivate them to change their diets. This and the smoking studies suggest that increased fatalism, or relief about "low risk" results, might validate continuance of less healthy behaviors; however, this has not been extensively tested. The Multiplex Initiative is an observational study using a test with 15 genes for eight common health conditions that will analyze what prompts people to seek genetic information, such as family history or a desire to change behaviors.

The possibility of tailoring interventions for conditions such as obesity based on a person's genomic profile also should be explored. A small study of perimenopausal women in Japan found that some women with a polymorphism in the α-adrenergic receptor gene had more difficulty losing weight using certain behavioral interventions. Deconstructing behavioral phenotypes to identify and measure pathways that might affect behavioral adherence also will help determine how genetic information might influence behavior. For example, physiological factors such as perceived exertion and the ability to feel a positive effect of physical activity may affect a person's ability to lose weight. Developing a multidisciplinary approach that incorporates behavioral research and moves beyond the psychological effects of genetic risk communication is crucial for optimal use of genetic information to motivate behavioral change.

DISCUSSION

Dr. Khoury noted that the research agenda for establishing clinical validity and utility for personal genomics needs to include behavioral and social marketing research as well as risk modeling and reclassification. To justify funding such an effort, how genomic information can

influence the public health agenda needs to be determined. Dr. McBride suggested moving away from generic public health recommendations because they are perceived to be ineffective. Risk communication in itself also is not sufficient; transdisciplinary approaches are needed. Dr. Bernhardt said that epidemiological information could be used to determine target populations, which could lead to better resource allocation and better ways to segment populations for interventions. Gaps between knowledge and behavior or action are not predicted by perceived risk based on genotype. Markers for risk-taking and other behaviors may help determine the best ways to deliver information to individuals.

Dr. Downing commented on the shift from using experts to social networking for obtaining information. Health information should be treated differently than consumer information, and it is difficult to understand why the average person would trust other people rather than experts. Dr. Bernhardt responded that institutions have been losing credibility since the 1950s; in addition, access to information from other sources has increased. He argued that health information is not significantly different than other consumer information—health information tends to be accessed or presented more haphazardly, but sites where people can exchange health information much as they do consumer information exist. Consumers exchange information on diagnoses and treatments with their peers, and this influences decision making. A participant asked about the impact of DTC pharmaceutical marketing on health communication. Dr. Bernhardt acknowledged that pharmaceutical advertisements are prominent in people's minds and may dilute some of the other messages.

Dr. Feero asked whether the effects of genetic variation on behavioral interventions could be considered similar to genetic effects on drug response. If this is the case, the effect sizes likely would be small and determining the appropriate intervention for a given risk profile will be complicated. Dr. McBride countered that this could also mean that focus could be limited to a small number of genes. Dr. Simons-Morton said that information about risk factors for CVD can motivate people to lose weight, decrease salt intake, and increase physical activity. Whether adding genetic risk factors to existing profiles will increase motivation to change remains to be seen. Dr. McBride said that the objective for risk communication needs to be clarified; risk communication alone will not directly affect behavior change or adherence. Including genetic information in studies of behavior change may motivate people to learn more about their risk but will not necessarily motivate behavior changes. The information could be used to tailor prevention programs or interventions such that people are more likely to make changes and adhere to them. Dr. Bernhardt suggested that genetic information about risk may prompt people to form social networks that become intervention points.

Ms. Trepanier asked if the point in life at which genetic information about risk is offered has an effect. Dr. McBride acknowledged that primary prevention will require genetic testing in children, but research is needed to determine if learning about risk earlier in life will have an effect. Dr. Friedman suggested working with disease-specific advocacy groups and networks because they might be able to suggest strategies for reaching people who are not yet at risk.

A participant noted that many of the presenters seemed optimistic about using genetic information to manipulate behavior and asked if there were any harmful consequences to

providing this information, such as misunderstanding of risk. Ms. Trepanier answered that genetic counseling reduces anxiety, but most people who come for counseling are highly motivated and educated. Dr. McBride said she had found no consistent evidence of a negative psychological impact of genetic information. Dr. Bernhardt said that some terms used when communicating genetic information are perceived as stigmatizing, particularly in certain populations.

Dr. Ioannidis said that there is evidence that experts do not promote their messages effectively; perhaps rather than providing information, experts should allow people to find this information on their own. Dr. Bernhardt responded that communication from experts still is needed but is insufficient by itself for increasing awareness and motivating behavioral change. The impact of expert communication can be increased by combining it with information disseminated through peer networks.

DECEMBER 18, 2008

Session V: Case Studies and General Discussion of Clinical Validity and Utility

Moderator: Michael Lauer, M.S., NHLBI, NIH

Clinical Validity and Utility of Genome Profiles in Risk Assessment and Control of Colorectal Cancer

David F. Ransohoff, M.D., Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill

Genomic profiles might be of use for prevention, screening, and early detection of colorectal cancer (CRC). Approximately 150,000 cases of CRC are diagnosed each year in the United States, resulting in approximately 50,000 deaths. Only one-third of cases are detected at a curable stage, and chemoprevention is of limited use. Genomic profiles could be informative for both a person's current and lifetime risk.

High risk but rare conditions associated with CRC include adenomatous polyposis coli (APC) and familial adenomatous polyposis (FAP), which have provided a great deal of information about the biology of CRC. For these conditions, action is based on family history; given a family history of APC, screening is implemented earlier in life. A combination of family history plus screening results is adequate for diagnosing APC; thus, genomic information does not add greatly to screening for this condition.

Far more people have an average risk for CRC; current preventative approaches include screening (sigmoidoscopy, colonoscopy, or fecal occult blood test [FOBT]) for people 50 years of age or older. Genomics could be used to more precisely quantify risk among those at average risk, classifying them into higher or lower than average risk groups and tailoring screening approaches accordingly. Risk currently is quantified to some degree based on family history, but recommendations regarding degree of risk, features that indicate risk, and the degree of risk that warrants action (and what action) vary widely among groups. This situation has arisen because of insufficient data and disagreement about what the data mean; there is no quantitative conceptual framework for handling family history data. Before developing such a framework,

whether it will provide actionable information must be determined. In the United States, screening is overused; quantitative information about risk (e.g., genomics or tailoring) could improve this situation. Genomic data that could identify people in the population at very low risk for CRC could result in more efficient allocation of screening resources.

Identifying a person's current (rather than lifetime) risk informs screening for early detection. Genomics could be of use in this regard; for example, DNA mutations can be detected in CRC cells shed in stool. A test was developed that classified risk based on detection of a progression of mutations in a canonical series of genes involved in CRC development. This test had a sensitivity of only 51 percent and a specificity of 95 percent; this is better than FOBT, but the test was expensive. The test was improved by including detection of "long DNA" (people with colon cancer tended to have longer lengths of DNA present in stool) and methylated vimentin, which were not previously believed to be part of the progression of mutations occurring during the course of CRC development. These data underscore the importance of understanding the biology of a condition when developing risk prediction tools.

The Impact and Utility of Personalized Genomic Information: Insights from the REVEAL Study

Robert C. Green, M.D., M.P.H., Boston University Schools of Medicine and Public Health

Alzheimer disease (AD) is the most common dementia worldwide and becomes more common as people age. Risk currently is estimated based on family history; the cumulative risk for those with family members with AD is 41 percent. Polymorphisms in the APOE4 gene are associated with increased risk of AD, with risk increasing threefold for one variant allele and 10 to 15-fold if two copies of the variant are present. When this information became publicly known, many people wanted to know their APOE4 genotype, despite the lack of treatment or prevention options for AD. In 2000, there was reluctance to genotype people for APOE4 because of the possibility of psychological harm or discrimination and a lack of treatment to prevent AD.

The Risk Evaluation and Education for Alzheimer's disease (REVEAL) study explored the use of APOE4 genotype for AD risk assessment. APOE4 genotyping could be used to define at-risk persons to enrich clinical trials, explore responsive or vulnerable subpopulations, respond to interested family members, and develop clinical paradigms for the use of susceptibility markers in common disorders. APOE4 testing provides a good model for exploring clinical utility and ethical, legal, and social implications of genetic testing because the test has excellent analytic and well-documented clinical validity; the lack of treatments alleviates market pressures; the disease is serious; and people want to know their risk. The main question asked by the trial was whether the information provided by APOE4 testing was beneficial or harmful. Challenging IRB issues arose because AD is untreatable.

REVEAL asked how risk information based on genetics could be communicated. Because genotype is not always clearly associated with risk, risk curves were generated and participants were shown risk for the general population, risk for people with a family history of AD, and risk for those with a particular APOE4 genotype. APOE4 test results have been disclosed to approximately 700 people; a number of these people were satisfied to learn that their risk of

developing AD was around 52 percent because they had previously believed their risk to be 100 percent.

REVEAL also sought to learn why people would want to know their APOE4 genotype. People with first degree relatives with AD were asked if they wished to participate in the study; 24 percent agreed to participate and did in fact participate. Among those who self-referred themselves to the study, 64 percent agreed to participate. Reasons for testing included preparing their families, obtaining information for family planning, arranging personal affairs, and arranging for long-term care. These results point out that "personal utility" must be considered along with clinical utility, given the participants' desire for genotype results in the absence of treatment options.

The impact of learning personal risk was assessed in a randomized controlled trial in which the first group received risk information based on family history and the second group received information about their APOE4 genotype and was divided into two groups based on high risk or low risk genotype. There was no difference in anxiety or among validated measures of psychological distress among the groups. All groups reported satisfaction with the information they received and most said they would participate in risk assessment again. Participants also reported fairly good recall of the genetic information presented to them. Analysis of behavioral changes showed that, despite the lack of options for mitigating AD risk, those with the high risk APOE4 genotype were more likely to improve their diets, increase their physical activity, and purchase long-term care insurance; they also were more likely to use vitamins or herbal supplements. The REVEAL study also explored attitudes of African Americans toward genetic testing for AD and developed a process that might be useful when applying genetic information across ethnic groups.

The REVEAL study also asked whether preparatory genetic counseling was necessary for safe disclosure and found attending counseling before receiving negative information may help with short-term coping. Participants' willingness to pay for testing also was examined, with 29 percent reporting willingness to pay at least $100 for genetic testing for AD. Thus, a fairly significant value was placed on testing despite the lack of treatment or prevention options. Whether genetic testing changed self-perceived risk was analyzed; among those who accurately recalled their risk, 47.5 percent continued to believe their risk was different, demonstrating people's natural resistance to changing their perceived risk. When asked who they told about their risk, 81.5 percent said they had discussed their results with a spouse or other family member or a health care provider.

REVEAL has shown that people find "personal utility" in risk information, regardless of whether the information is medically actionable. Additionally, as treatments are developed, what was once inactionable may become actionable. REVEAL will continue to explore issues concerning the effects of telephone or online disclosure of results with minimal counseling; the effects of receiving risk information about a disease the participant did not expect to learn about; and the effects of combining genotype and phenotype information (such as early memory loss for AD) to offer more imminent risk information.

How Can We Assess Clinical Validity and Utility of Genome Profiles in Risk Assessment and Control of Cardiovascular Disease?

Philip Greenland, M.D., Northwestern University

Defining clinical validity and utility of a test for risk assessment involves assessment of the predictive capability of the test itself; assessment of the clinical utility of the test; determining the effects of the test on patient outcomes; and cost-benefit analyses.

Tailoring interventions according to CVD risk has occurred for a number of years. Tailoring is based on validated risk factors and tools, including the Framingham Risk Score, which incorporates risk factors such as family history, cholesterol levels, blood pressure, and smoking, among others. Experience in the CVD risk prediction field has shown that use of multiple markers results in better predictive capacity and better discriminatory ability than use of a single marker. The AUC for the Framingham Risk Score is 0.80, but addition of coronary artery calcium score improves it to 0.84 and provides better discrimination between those who will develop CVD and those who will not.

Assessments of clinical utility involve asking whether physicians can predict an event just as well without using the risk assessment tools, whether the physicians use the tools, and whether the tools improve patient outcome. In the CVD field, physicians understood the relative importance of specific risk factors well and were confident in their ability to estimate CVD risk. However, although they estimated relative risk for a patient (compared with an average adult) well, they overestimated the absolute baseline risk of developing CVD and the risk reductions associated with specific interventions. To have clinical utility, a risk assessment tool also must point to a specific risk reduction intervention, including behavioral change, which will affect patient outcome. There is some evidence that inclusion of coronary artery calcium score with Framingham Risk Score results in reclassification of patients and changes in outcomes. However, a randomized controlled trial showed that although coronary artery calcium scoring provided important information regarding risk, it was not sufficient to change patient behavior. In contrast, providing support to patients did result in better management of risk factors. A systemic review of the routine use of global CHD risk scores was performed to determine if this translates to clinical benefits (or harms). This review found that physicians' knowledge of CHD risk scores translates to modest increases in prescription of cardiovascular drugs and modest short-term reductions in CHD risk factors; however, it is unclear whether this translates to improved long-term CHD outcomes.

Prospect theory allows description of how people make choices in situations where they have to decide between alternatives that involve risk and how they evaluate potential gains and losses. One study found that at low predicted risk, people behaved according to the information they were given. At very high risk, people also behaved accordingly. However, over a broad range of medium levels of risk, people behaved identically and not according to their own risk. This implies that placing people in low or high risk categories will be more effective for prompting rational behavioral changes.

The evidence has shown that better ways of making decisions about CVD risks are needed, including better tests. New breakthroughs in the predictive capability of tests are needed to result in better treatment and improved outcomes. Better ways of communicating risk also are needed.

Predict Prostate Cancer Risk Using SNPs: Promising but Complex
Jianfeng Xu, M.D., Dr.P.H., Wake Forest University School of Medicine

Risk prediction using genetic variants is promising but complex. Recent GWAS for prostate cancer have identified 8q24, two regions on chromosome 17, and several other loci located throughout the genome as associated with prostate cancer risk. Several of these SNPs have been consistently replicated. Most of these risk variants are fairly common and although the ORs for each variant are small, the combined OR is significant. The cumulative effect of five SNPs was examined and the analysis found that risk increased as the number of variants a person had increased.

Although the association of these variants with prostate cancer risk appears significant, it is too soon to use them to predict prostate cancer risk. More research is needed, including larger sample sizes and prospective studies, analysis of race-specific effects, and analysis of aggressive versus non-aggressive disease. Work also is needed to address issues related to prostate specific antigen (PSA) detection bias, determine how the variants can supplement PSA for predicting positive biopsy results, and analyze how the variants can aid targeted chemoprevention.

Recent work has shown that six risk variants were significantly associated with PSA levels in men without prostate cancer. This appears to be related to detection bias; if a man positive for prostate cancer risk SNPs also has high PSA levels, he is more likely to undergo a biopsy, and cancer thus is more likely to be detected. No difference has been found in distribution of SNPs in aggressive versus non-aggressive prostate cancer cases, although men with lower grade cancer appeared to be positive for more risk alleles. There is some evidence that a SNP within 8q24 is more tightly associated with prostate cancer risk in African American men than in Europeans. Work is also underway to determine if inclusion of SNPs in prostate cancer risk assessment can improve the predictive value of a positive biopsy and reduce the need for multiple biopsies. Use of these risk variants also may better target finasteride treatment, which has been shown to be effective in reducing prostate cancer diagnosis and also diagnosis of aggressive prostate cancer.

DISCUSSION

Dr. Ioannidis asked if better outcomes could be achieved by adding SNPs to the high end of risk curves. Dr. Xu answered that because prostate cancer is a polygenic disease, the effects of addition of SNPs to risk curves are not clear. In response to a question, Dr. Xu explained that his group has requested data from finasteride trials to assess PSA detection bias.

Dr. Khoury noted that the personal utility found in testing for AD implies that personal genomics might be useful even for deadly diseases for which there are no treatments. He asked if randomized controlled trials could be developed to analyze personal utility in the absence of

interventions. Dr. Ransohoff noted that a conceptual framework is available through United States Preventive Services Task Force (USPSTF), which was created because prevention research was being performed without including proof of outcome and there were concerns about possible harms of these trials to otherwise healthy people; for example, whether knowing their AD risk would cause more harm than good. Research on risk assessment protocols in healthy people can be analyzed using the USPSTF framework to quantitatively assess the evidence. It is important to remember that the goal is to improve outcomes and determine that the benefits are commensurate with risk (e.g., stigmatization, insurance issues) in healthy people.

Dr. Greenland said that some outcome information could be obtained from modeling, and then could help inform trials. He described modeling work for CVD which compared outcomes resulting from risk assessment based on standard care versus Framingham Risk score versus the score plus various other biomarkers (carotid artery intima, coronary artery calcium) and compared this to unconditional treatment with statins for all participants over a certain age. This study showed that for CVD outcomes, unconditional treatment with statins was cost-saving compared to standard care and produced the best clinical outcomes. Although this raises some controversial issues, following up on these results with a clinical trial might be useful.

Dr. Green noted that there is a range of actionability for a number of conditions. Many genetic and environmental risks are being discovered, and sentinel clinical trials to understand how categories of conditions are affected by risk information might be useful. For example, REVEAL studied a late onset disease with no actionable outcome; the findings of this trial could stand in for those for similar conditions. If a treatment for AD is found, it would be moved to a new outcome category. In addition, because the knowledge level of physicians and patients is increasing, and they are developing better understanding of genetic information, clinical trials could be performed in various domains.

Dr. Gulcher suggested that more prostate cancer studies are needed to accurately calculate genetic risk. He was concerned about the variability of assessments using risk alleles; it may be more efficient to use a tool like the Gail risk model and recognize that some risk factors have a stronger effect than others. Dr. Xu responded that each risk variant is taken into account when the AUC for these variants is calculated. Detection bias is a major complication, as is distinguishing between aggressive versus less aggressive disease.

Dr. Offit asked how people who are negative for the APOE4 risk variant but have a positive family history of AD are advised. He also asked if there was evidence that use of prostate cancer risk alleles would be of use for people with a family history of early onset prostate cancer, because this could affect screening practices. He asked if recent information on cyclins and other biomarkers had proven useful for assessing CVD risk. Dr. Xu answered that no risk variants have been associated with early onset prostate cancer. Dr. Green answered that most participants in REVEAL had a family history of AD; about 30 to 40 percent who tested negative for the APOE4 risk alleles were falsely reassured about their risk. Dr. Greenland noted that a number of published SNPs actually appear to be independent in terms of their role in CVD risk. Many biomarkers appear to minimally influence overall risk; they can be used to show reclassification changes, but there are errors in the estimate of rates. He proposed that

information on error and changes in AUC be used to model the effects of change, and then model the likely effect on patient outcome and choose the most effective model for developing a trial.

Session VI: Models for Conducting Translational Research on Genomic Profiles
Moderator: Gregory Feero, M.D., Ph.D., NHGRI, NIH

Scientific Evaluation of the Clinical Validity and Utility of Genetic and Genomic Risk Factor Information
Geoffrey S. Ginsburg, M.D., Ph.D., Duke Institute for Genome Sciences and Policy

Biomarker discovery, clinical confirmation, and assay development have been achieved for a number of genetic variants. Moving these biomarkers into clinical trials and practice will be challenging. The steps to move biomarkers forward into practice also are not well defined for genetic markers of risk. Moving biomarkers into practice will require development of an infrastructure that includes improvements in biobanking, genomic technologies, informatics, biostatistics, and decision making. The last is particularly crucial because most physicians are untrained in decision making.

Several examples of genome risk factor clinical utility studies exist, all of which depend on access to samples from large trials. An RNA signature was used successfully in a retrospective study to predict recurrence and death in patients with early stage lung cancer. The results informed design of a trial testing the benefit of chemoprevention based on risk prediction using the signature. A breast cancer neoadjuvant trial used genomics to guide treatment options; use of the genomic signature predicted chemotherapy response with an accuracy of 75 to 80 percent.

A primary care-based randomized controlled trial for the clinical utility of TCF7L2 was performed to assess the ability of a genetic test for type 2 diabetes risk to alter behavior and health measures in a general clinic population. Secondary goals were to measure whether changes in perceived risk and beliefs about genetics correlated with behavior change following genetic testing and to determine whether a genetics-guided clinical trial would change primary care physicians' beliefs and understanding of genetics and their role in practice. This trial involved researchers from a number of different disciplines, including clinical research coordinators trained in genetics.

To enable scientific and clinical evaluation of genomic markers, patient registries of both rare and common diseases, with longitudinal followup and robust phenotyping, are needed. Creation of a "Genomics Trials Cooperative Group" also would be beneficial, as well as public-private partnerships including sample collection from phase II-IV trials performed by the pharmaceutical industry. A national virtual sample biorepository linked to research and clinical data also is needed. As part of the effort to improve translation, a National Clinical and Translational Science Awards Consortium is being created to serve as a resource for personal genomics research. This consortium will address many of the stated needs for improved translation of genomic information and also will promote community engagement in the process.

NIH GEI Genomics Translation Research: Recent Initiatives
Paul L. Kimmel, M.D., Kidney Translational Genetics Program, National Institute of Diabetes and Digestive and Kidney Diseases

NIH's Genes, Environment, and Health (GEI) Initiative supports research on the involvement of genetics and exposures—environmental, psychological, and treatment-related—in complex disease. Another goal of this initiative is to support efforts to move GWAS results into practice and improve public health. A workshop was held in March 2008 to explore the challenges of translating genomic findings related to complex diseases to improve public health. The consensus developed at the meeting stated that development of diagnostics and approaches to development of therapies using genomic information will be useful, but randomized clinical trials using information from GWAS may be premature. In addition, dissemination of accurate information to patients and practitioners was deemed essential, but evidence concerning the correct way to use genomic data, whether genomic data can be used to improve patient behaviors and health, and patient response to receipt of information on risks determined from GWAS is lacking. Clinical utility was discussed in terms of whether genetic risk information would prompt behavior change in patients, given that many do not modify their behaviors (e.g., exercise and diet) in response to information about traditional risk factors. Whether there are adverse consequences to receipt of genetic information about common diseases also was discussed, as was patients' and consumers' abilities to assess risk and make decisions based on genetic information.

An outcome of this meeting was the development of two trans-NIH GEI initiatives. The first, Translation of Common Disease Genetics into Clinical Applications, will support clinical studies using information from GWAS or other genetic studies in common diseases; development and assessment of diagnostic, clinical trial, epidemiologic, and risk analysis tools for use in clinical research or practice; and studies of the cost-effectiveness of clinical applications of genetic information. The conditions covered by this RFA include asthma, diabetes, obesity, and atherosclerosis, among others. The second, Implementation Planning Grants for Educational, Behavioral, or Social Studies for Translation of Genetic Factors in Common Diseases, is a U34 planning grant that will support multicenter research on educational and communication initiatives for health care providers and consumers regarding interpretation of findings from genetic studies of common diseases and the results of their dissemination, and behavioral or psychosocial aspects of clinical application of genetic findings.

CDC Genomics Translation Agenda
Ralph J. Coates, Ph.D., National Office of Public Health Genomics, CDC

Although the United States spends approximately 16 percent of its GDP on healthcare, it lags behind many other developed countries in health. Only 55 percent of Americans receive recommended care for acute or chronic conditions and 50 percent receive recommended preventive care. Approximately 30 to 40 percent receive contraindicated care, and 30 to 40 percent of health care dollars are spent on overuse, underuse, or misuse of services. Translation of personal genomic information into practice may help to address these issues and improve health care spending and public health. However, many questions about genomic information

exist, including the validity and reliability of genetic tests; their efficacy in predicting outcomes; benefits and harms associated with the tests; and the actions that should be taken based on genetic results.

The CDC has developed a continuum for translation that begins with discovery, followed by application of the discovery to evidence-based guidelines, development of guidelines for clinical and public health practice in communities, and practice to impact on health in communities. There is limited research on evaluation and implementation. The purpose of EGAPP is to establish and test a systematic, evidence-based process for evaluating genetic tests and other applications of genomic technology in transition from research to practice. EGAPP has produced systematic evidence reviews evaluating analytic and clinical validity and clinical utility; recommendations on appropriate use of genetic tests and other applications; and recommendations for research to fill specific evidence gaps. CDC also is developing the Genomic Applications in Practice and Prevention Network (GAPPNet), which will work with EGAPP to help increase sharing of information among stakeholders and help foster collaboration, knowledge synthesis, and dissemination of evidence-based guidelines.

CDC has decided to support analysis of the clinical utility of family history because family history and genomics information will contribute to the development of personalized medicine. Family history risk factors exist for a number of conditions and family history can be viewed as a low cost "omics" tool. This study will ask whether risk notification and tailored messages using a family history tool will change behaviors (e.g., use of clinical services, lifestyle changes, family communication). CDC also is supporting research to evaluate the clinical utility (including improved health outcomes and value in clinical decision making) of pharmacogenomics for warfarin use, gene expression profiling for treatment of early stage breast cancer, and Factor V Leiden testing for pregnant women with clotting or adverse pregnancy outcomes.

Public Engagement, Policy, and Oversight While the Science Accumulates
Kathy Hudson, Ph.D., Genetics and Public Policy Center, Johns Hopkins University

Effective implementation of genetic medicine requires a robust, responsive, and responsible research enterprise; improved guidelines development and adoption; prepared patients and providers; fair reimbursement; safeguards for genetic information; and safe and effective tests.

Currently, two types of genetic tests exist: laboratory-developed tests and test "kits." Regulatory oversight for the laboratories that develop these kits rests with CMS, rather than FDA, which would seem a more appropriate agency. The differences in qualities of these kits are not readily apparent to patients and healthcare providers. There is also inadequate oversight of claims made by the manufacturers of these tests. Safe and effective genetic testing requires a level of oversight based on risk and not mode of manufacture. These tests should deliver correct results nearly all the time, and data on their analytic and clinical validity and clinical utility should be publicly accessible. High risk tests also should be subject to independent review before being marketed.

Congress intended laboratories to perform proficiency testing, but the list of analytes requiring proficiency testing has not been updated for 20 years. CMS has argued against new proficiency testing requirements, citing an absence of sufficient proficiency testing materials and programs. Accordingly, proficiency test providers cite an absence of demand sufficient to develop new materials and programs. CMS regulation of genetic testing laboratories has no mandate to perform proficiency testing, no evaluation of clinical validity, little public access to information, and no authority over claims and labels, and is buried in an agency with a different mission and expertise. FDA regulation of test kits calls for evidence of clinical validity for intended use, and FDA has authority over manufacturer or distributor claims. Only a few human genetic tests have been approved by FDA as kits. FDA retains more discretion for regulation of laboratory-developed tests, which has led to inconsistencies in the quality of these tests. In December 2008, the first requirement for FDA to regulate test kits in the same manner regardless of how they were manufactured was enacted.

Safe and effective use of genetic information also should require that pharmacogenetic data be rapidly incorporated into labels, and pharmacogenetic tests are subject to post-market surveillance. Retrospective studies have shown that monoclonal antibodies to epidermal growth factor are not effective for cancer patients with somatic K-RAS mutations; FDA is considering including this information on the drug label. It would therefore be beneficial to have confidence in the validity of the test for K-RAS mutations.

Information also is needed concerning how people view and will use personal genomic information. When asked, a majority of those questioned said they would be willing to participate in a large cohort study to further explore the impact of personal genetic information on common diseases and conditions. Most people expressed an interest in knowing their own genetic risk factors for any number of conditions, regardless of type or severity. Approximately two-thirds of those questioned also do not believe that genetic data require extra privacy protection, viewing it on a par with medical records with respect to privacy protection.

DISCUSSION

A participant noted that at least one-third of those questioned were concerned about privacy regarding genetic information. Dr. Hudson responded that more than 90 percent of respondents said that privacy protection was necessary when participating in research, but only one-third viewed genetic information as needing extra protection.

Dr. Witte noted that family history is obtained more cheaply than genetic information, but it tends to underestimate risk. Dr. Coates said that research is underway to test the use of family history and genetic information, but at this point it is difficult to say which will give a better risk estimate. Dr. Witte responded that the relative usefulness of each will probably vary depending on the condition being assessed. Dr. Kari Stefansson said that most sequence variants associated with common disease are separate from family history, and one set of information cannot be substituted for the other. Dr. Greene said that for cancer, there is a great deal of information on family history and its validity for risk prediction. Cancer site and the closeness of the relationship between the reporter and family member influence predictive ability, which is highly

accurate for some sites but not others, particularly metastatic disease. Unverified reported family history has been used for every epidemiological case-cohort study to estimate risk, but few studies have confirmed whether this approach is valid. A study is under way to analyze the efficacy of using family history to predict risk by comparing these data with data in a population-based registry in Connecticut. Family history is a low cost, low technology way of predicting risk that could be powerful if used properly.

Dr. Friedman asked if there was any way for CMS to respond to concerns about claims made by tests. Dr. Hudson answered that most claims issues arise over clinical validity rather than analytic validity, and CMS alleges that it has no authority over clinical validity issues. The Federal Trade Commission is involved with this process, but only one staff member oversees genetic testing. He has recently launched two investigations of genetic testing companies based on allegations of false claims.

Dr. Kardia noted that determining the best study designs for evidence generation in this field is necessary. Generalizability of inferences is desirable, but randomized controlled trials and trials performed in hospital settings have serious generalizability issues. Natural cohort studies might be more generalizable, but these studies often take too long to generate data. Many study design choices are generated by available funds and peer review, and reviewers must determine if they have funded the best possible approach for generating evidence. Dr. Ginsburg suggested that the way in which the data will be used might affect trial design. For example, data used to make high risk decisions should be generated using prospective clinical trials. For lifestyle modifications, observational data might be sufficient. Trials also could be prioritized based on how quickly outcomes can be implemented; trials including pharmacogenetics of response to therapy or infectious disease trials might fall in this category. Dr. Hudson asked Dr. Ginsburg if existing assays that distinguish between chemotherapy responders and non-responders will be used as point-of-care kits and if plans were in place to market such kits. Dr. Ginsburg answered that talks have been held with diagnostics companies to discuss how to commercialize the assays, but more information is needed concerning the amount of data required and regulatory hurdles.

Models for Conducting Translational Research on Genome Profiles

George Church, Ph.D., PersonalGenomes.org and the Broad Institute

DNA sequence, rather than SNP chips, is the basis for some older, well-used genetic tests. For example, tests for mutations in BRCA1 and BRCA2 are based on sequencing information rather than SNP variants. Tests for BRCA1 and BRCA2 mutations are not available as DTC tests because of technical and intellectual property issues related to sequencing. A number of next-generation sequencing platforms have been developed that will permit sequencing to be performed quickly and at lower cost, which may have implications for genetic tests for risk. As sequencing becomes easier, questions concerning whether whole genome or other forms of sequencing infringe on intellectual property rights will arise.

PersonalGenomes.org is the first and only group to provide open access data on a number of parameters including health records, environmental exposures, gene × environment interactions, and immune response. Because of this open access policy, the company developed a strict informed consent policy, which required participants to score 100 percent on an exam designed to assess their understanding of the policy. PersonalGenomes.org will generate low cost coding sequence and regulatory region data; imaging data; and stem cell RNA; and will provide cells for personal functional genomics. The company has received IRB approval for 100,000 diverse volunteers.

Jeffrey Gulcher, M.D., Ph.D., deCODE Genetics

Prostate cancer is the second leading cause of cancer death in men; 3 percent of diagnosed men will die of this disease. Currently, a family history of early prostate cancer is the only risk factor known for the Caucasian population. Based on analyses using eight validated genetic markers, approximately 10 percent of the average male population has a twofold risk of developing prostate cancer, independent of family history; this accounts for 30 percent of prostate cancer deaths. Improved early detection is needed.

Genetic profiling may help direct screening and treatment for individuals. For example, a 48-year-old Caucasian man whose father was diagnosed with localized prostate cancer at age 68 was recommended to begin PSA screening at age 50. Performing risk assessment using the deCODE prostate cancer variants determined a relative risk of 1.88-fold compared to the general population risk for white men, and a calculated lifetime risk of 30 percent, with a modestly higher risk for aggressive disease. These results prompted an early PSA test (before age 50); PSA levels were in the high normal range (2.0 ng/mL). Because of the higher risk suggested by the genetic profiling results, the patient's urologist decided to perform a biopsy, which found cancer with a Gleason score of 6 (3/3). The patient underwent a radical prostatectomy with nerve sparing, and final pathology on the resected prostate showed a Gleason score of 7 in both lobes.

Prostate cancer utility studies are under way to determine if the genetic test can increase the specificity of PSA tests and to determine whether the markers correlate with aggressiveness at diagnosis or long term. Other ongoing studies of clinical utility include determining whether genetic testing can increase the specificity of breast imaging based on biopsy outcome, improve the Gail risk model, predict responders to tamoxifen/raloxifene prevention therapy, or change patient behavior. Studies are also under way to determine the effect of risk variants for atrial fibrillation on diagnosis of atrial fibrillation-related strokes and the influence of markers for type 2 diabetes on the ability of prediabetics to lose weight.

Linda Avey, 23andMe, Inc.

Translational research is an important part of the mission of 23andMe, Inc., which seeks to provide access to genetic information and also facilitate and conduct research. 23andMe, Inc. currently is working with the Parkinson Disease Institute in Sunnyvale, CA on a project to determine if a range of diagnostic tools and surveys can be provided online to facilitate

diagnosis. IRB approval has been received to use the web-based tools to survey people who were previously diagnosed using paper-based tools to determine if the same results are obtained. Genotype information also will be collected from participants.

23andMe, Inc. has found that consumers are highly interested in participating in genetic research but also want access to their own information. 23andMe, Inc. plans to make their platform and customer base (with disease information) available to researchers.

Amy DuRoss, M.B.A., Navigenics, Redwood Shores, CA

Navigenics is a clinically focused company that plans to offer genetic tests for 25 different conditions. The Navigenics Health Compass analyzes genetic predisposition for a number of common health conditions, a whole genome scan on 1.8 million genetic risk markers, and access to genetic counselors to explain the results.

Navigenics also has launched a number of studies to help understand the hopes and expectations of patients and providers with respect to predictive genetic information. The first study, the Mayo Clinic Proof of Principle Trial, involves 125 people enrolled in the Mayo Executive Health Program. The study will examine people's expectations regarding the significance of genetic testing and whether they understand its limits; determine whether and how genetic information impacts family history risk information; and compare different methods of delivering genetic risk information (DTC, DTC with genetic counselor follow-up, or direct to Mayo physician). Physician expectations and understanding of the limitations and benefits of genetic information on decision making in an outpatient health screening environment also will be assessed. The Cleveland Pilot Study will assess the capability of the Cleveland Clinic's Genomic Medicine Institute's hereditary cancer risk assessment independently and together with Navigenics' whole genome SNP analysis to detect individuals at risk for monogenic cancer syndromes, including breast, colon, and prostate cancers. The Scripps Large Scale Longitudinal Study will determine whether participating in consumer genomic testing can improve health by motivating positive lifestyle changes to mitigate the impact of conditions for which the consumer is at risk.

DISCUSSION

Panelists were asked to speculate on the efforts to increase public-private collaborations. The panelists agreed that clinical utility studies of various genetic tests are needed and that large trials would be needed. Because such trials are expensive, ways to creatively structure financing to support the goals of both the public and private research enterprises are needed.

Ms. Avey noted that the companies have placed more emphasis on the consumer and have learned that members of the public are willing to participate in genetic research and share their results. Gathering genetic data on large numbers of healthy people is essential for sound prevention and health promotion research. Dr. Church noted that providing consumers with access to their own medical records and genomic information could help reduce the cost of studies. In addition, this might provide a way to aggregate studies; a single person's information is relevant to multiple studies.

In response to a question about funding sources, Ms. DuRoss explained that Navigenics' Scripps Large Scale Longitudinal Study is funded by Affymetrix, Microsoft, TSRI, and Navigenics; participants also pay a deeply discounted rate for the genotyping service. Navigenics has direct shared funding with the Mayo and Cleveland Clinics for the other studies. Ms. Avey said that 23andMe, Inc.'s Parkinson study is funded in part by the Michael J. Fox Foundation.

Dr. McBride asked how impressions regarding the utility of genetic tests were managed by the companies. Ms. DuRoss explained that there is no management; data will be released periodically and collected long term. Ms. Avey said that 23andMe, Inc. would publish both negative and positive data. Dr. Ioannidis noted that trials sponsored by companies are more likely to find favorable results, despite similar data quality. He asked if the companies would consider having different people involved in study design, analysis, and interpretation of results.

Dr. Khoury suggested a randomized controlled trial of a few hundred people to explore the benefits and harms of knowing one's own genetic risk information. At the March 2008 workshop, the general opinion was that most GWAS results are not ready for use in clinical trials. Trying to determine the utility of these tests at the same time they are marketed may not be the best approach. Ms. Avey responded that 23andMe, Inc. will open its data to review by other researchers; 23andMe, Inc. customers can download their genotyping results and provide them to any researcher they choose.

Dr. McBride noted that the populations studied by the companies seem rather limited and asked how the external validity of the results could be increased. Ms. DuRoss said that participants in the Scripps study were drawn from the Scripps Health Community and also from Microsoft and Affymetrix employees, and there is a fairly broad range of income distribution. Because the sample size of the study is 10,000, it should be possible to draw meaningful conclusions. Dr. Green noted that volunteers for clinical trials are always non-representative of the general population; this does not imply that all information from trials is valueless.

Dr. Offit asked if the companies plan to seek reimbursement for these tests by making personal genomics part of personalized medicine, allow only individual payers to make use of the techonologies, or return to the traditional academic community-to-clinic model of translational medicine. Ms. Avey replied that 23andMe, Inc. was created because a useful translational model did not exist and to facilitate data collection from large populations to better examine the small or rare effects of genetic variants. The current model for performing clinical studies is costly; using Web-based tools to determine which information to collect and to collect that information might help lower costs. 23andMe, Inc. has hypothesized that consumers will pay for genetic data as long as they receive their own data in return. Participants are clearly informed that the information provided for them is research information and is not actionable. Dr. DuRoss said that Navigenics is focusing on educating providers and selling tests through a provider model focused on large scale employers.

Dr. Janssens noted that because effect sizes are small, sound phenotyping is needed and thus relying on self-report may be inadequate. Ms. Avey replied that 23andMe, Inc. hopes to partner

with Google Health to access health information and also is exploring other ways of collecting patient information; however, customers will control access to their data.

Session VII: Panel Discussions and Next Steps for Research and Practice Agenda
Moderator: Muin Khoury, M.D., CDC, NCI

Francis Collins, M.D., Ph.D., Former Director, NHGRI

Given the identification of more than 8,000 genetic variants linked to risk for a number of common diseases, the genetic approach is a powerful one for discovering risk factors. However, the risks conferred by these genetic variants are moderate and the markers themselves are often located in areas of the genome that lack identifiable gene sequences or control regions; the role of sequence elements such as copy number variants also is not well understood. In addition, few less common variants with large effects have been identified. Most researchers anticipate that it will be possible to identify more than 50 percent of the heritability for a given condition; it is necessary to plan for this result because public interest in using this information to predict risk will increase significantly. Researchers will need to determine the clinical utility of this information with respect to possible interventions. In addition, efforts should be made to integrate genetic information with family history to determine an overall prediction of risk. Technology will be able to provide further details about the genome, and sequencing of individual genomes will be facilitated. The ability to estimate risk will improve and will affect how individuals are empowered by this information.

A great deal of discussion has occurred concerning genetic risk, but there has been less focus on the role of and interaction with the environment. Because genetics cannot be changed, clinicians wishing to provide information about interventions will need to consider ways to change an individual's environment to mitigate risk. With regard to collecting information about gene × environment interactions, the current focus on case-control studies involving already diagnosed subjects is problematic. Investigation of environmental effects using this approach is hampered by recall bias and the inability to measure environmental exposures. To circumvent this problem, a large scale, population-based cohort of the U.S. population should be created, balanced for variables such as age, gender, socioeconomic status, education, ethnicity, and geography. Such a study should enroll approximately half a million people over 5 to 10 years to have sufficient numbers of incident cases. A number of analyses regarding creation of such a cohort have been performed, and a study design and detailed proposal were created in 2004-2005, with an estimated cost of $400 million per year. This would be a high cost, but necessary, study for understanding environmental contributions to diseases. However, this study would likely incorporate a number of individual disease studies that ultimately would carry the same cost. A survey found that many Americans would be interested in participating in this study, particularly if they received information on their own risks for the diseases studied.

Another issue arising as the result of increased genetic information related to risk for disease is that of consumer protection. A great deal of public confusion currently exists regarding risk factors based on genetics and there is no centralized site where the consumer can receive clarification. Private companies have attempted to put forth the necessary information and

explain statistical risks, but a third party, user-friendly, objective source of information is needed for interested consumers. Such a source must be kept up to date and must be targeted at the lay public. The field of personal genomics will benefit from attention to future progress, identification of research and policy needs, and attention to the impact the results of this research will have on the consumer.

Kari Stefansson, M.D., deCODE Genetics, Reykjavik, Iceland

Seeking information, including personal genetic information, is an important part of human nature. Geneticists have published extensively on the genetics of common diseases, and the lay public reads about these discoveries in the press. Researchers must consider how to provide the public with access to these data and how best to take advantage of these discoveries. Setting minimal standards for providing this information, including standards related to analytic and clinical validity and clinical utility, should be considered. It also will be important to protect the public from possible adverse effects of these discoveries; for example, no claims except those supported by good science should be approved. It is naïve to expect that a consumer market for these discoveries will not develop, or that tests will not be created until the clinical utility of the discoveries used to develop the tests is proven. There are many examples in the clinic of tests for conditions being implemented before their clinical utility was unequivocally established. It will likely be difficult and expensive to meet these goals of clinical utility, especially for genetic variants that provide only a numerical assessment of risk. If the data supporting the relationship between a variant and risk of a condition is sound, that should be sufficient justification for use of the test in practice.

With regard to population-based studies to further explore genetic and environmental risk, such a study was proposed approximately 12 years ago in Iceland; this study would have created a centralized healthcare database. Most experts considered this study to be invasive and inadequately respectful of an individual's privacy, but more than 90 percent of the population of Iceland supported the study. The United States, Britain, and the Scandinavian countries disapproved; given the proposed nationwide cohort study, attitudes toward such studies have apparently changed. It has become obvious that such large scale studies are needed to realize the full benefit that genetic information can provide with respect to personal and public health. Such a study essentially exists in Iceland; although the centralized healthcare database was not created, much of this information was eventually centralized because most of the population participated in the study.

Experience has shown that even trivial variants, such as a genetic variant associated with a propensity for solving crossword puzzles, may have an impact on common complex diseases and underscore the genetic and environmental interactions that contribute to risk of these diseases. Because the genome cannot be changed, any efforts to reduce risk will need to affect the environment. For example, brain function underlies an individual's tendency to seek or avoid danger; thus, studies of brain function will be important for informing studies of the environmental components of risk.

Small studies have explored the relationship between genetics and environment and how this relationship affects risk. Common diseases have environmental components and may affect the same risk pathway as genetic variants; these components also may be subject to selection. For example, a study of the diversity of pigmentation in a European population analyzed sequence variants affecting pigmentation in the context of basal cell carcinoma and melanoma. A mutation in the melanocortin 1 receptor had no apparent affect on risk of melanoma for individuals living in Iceland. However, this variant was associated with risk for individuals living outside of Iceland. Individuals with this variant tend to avoid sunlight, which is easier to do in Iceland than in, for example, Spain. Individuals with this variant have a nearly threefold risk of developing melanoma if they live in Spain. The frequency of this variant is approximately 20 percent in Iceland, but only 4 percent in Spain. This example shows the ability to analyze sequence variants in a geographic context and allows separation of genetic from environmental effects.

Making full use of genetic information with respect to improving personal and public health will require establishment of large study cohorts. Issues such as ensuring sound analyses, a cooperative regulatory environment, and the sustainability of the research, particularly large, long term projects, also must be addressed.

REACTIVE PANEL

Dr. Stephen Chanock noted that the context of how genetic markers are used must be considered with respect to clinical paradigms. Use of these markers must be modeled in a more clinically useful way, but clinicians must take care when communicating the idea of absolute risk, how risk affects the use of possible interventions, and whether or not environmental contributors can be changed. It has been sobering to learn that the sizes of the effect conferred by many genetic variants are small, and also that knowledge of few of these variants has led to changes in survival or outcome. Researchers must consider how to integrate currently available information into practice. Many existing tests are not well validated, and researchers have a responsibility not to oversell the promise of genetic testing.

Creation of a large cohort for study of gene × environment interactions is highly important; the proposed cohort of half a million participants may not be large enough, given that effect sizes for variants are much smaller than anticipated. The need for such a large cohort also implies a need for nationalized healthcare, which will facilitate the analyses required to make the best use of genetic information.

Dr. Kardia agreed that a cohort for the study of gene × environment interactions would be useful. The passage of the Genetic Information Non-Discrimination Act, creation of companies such as 23andMe, Inc., and work by other researchers has shown that Americans are ready to engage in such a study. She expressed concern about the emphasis on the results of epidemiological and clinical studies over molecular and animal work to learn about the mechanisms that underlie the risks imposed by genetic variants. Cooperation with other disciplines will be necessary to move this work forward.

Ms. Terry commented that shifts in society's willingness to participate in research are evident. Academia tends to be risk averse, and different disciplines do not interact with each other as much as they should. Changes are needed to facilitate research that moves beyond the "publish or perish" paradigm. Thought should be given to creating a regulatory environment in which it becomes easier to perform studies and translate the results of the research, taking into account improvements in technology that allow better protection of subjects' privacy and the increased willingness of the public to participate in research. The public also needs to be shown that improvement in healthcare is iterative, similar to the way that computer software is developed and improved.

Dr. McBride suggested thinking about broader aspects of public health. A prioritized research agenda based on practical research questions is needed. At this meeting, the need for more precise risk estimates and how to evaluate clinical utility were discussed, but "useful" has not been clearly defined. The idea of "personal utility" should be considered; genetic or other health-based information may be useful to an individual even if no intervention to mitigate risk exists.

It also has become clear that a paternalistic reaction regarding protection of the public should be avoided. Sound mechanisms exist for informing the public, based on improvements in technology and recent innovations in social networking. Researchers have tended to underestimate the public's understanding of genetic research. The research community also has experience with informing the public about the results of research; for example, the Cancer Information System has provided sound, understandable information to the public.

Dr. Ioannidis noted that there has been a great deal of success in identifying genetic risk factors; more than 800 variants have been identified for which there are sound data on the risk they confer. Predicting outcomes is difficult; the research community wants to be able to make deterministic comments about risk but must be careful about conveying the uncertainty associated with these data. In addition, environmental factors need to be studied, emphasizing again the need for large scale studies. One approach to consider is to combine global biobanks, which could aggregate data from 10 million people. A database of that size could allow randomized trials of the impact of environment and behavior on disease risk to be performed.

The gaps that exist between academic understanding of these issues and the general public's understanding must be addressed. Neither the public nor most physicians have strong understandings of probability or statistics. This presents an opportunity for education using genetics and a way to increase the public's awareness of clinical epidemiology. The clinical utility of genetic information should be determined, but hundreds or even thousands of diagnostic and prognostic tools are currently in use for which robust proof of utility does not exist. The bar should not be set higher for genetic information; however, sound, well-designed trials to test the impact of these data on clinical practice are needed.

DISCUSSION

The participants acknowledged that GWAS are generating large amounts of information on the genetics of risk for certain diseases, but the best ways to use this information to improve personal and public health, as well as clear definitions of clinical validity and utility for genetic association markers have not been established. Large cohort studies probably would be useful, but given budget constraints, will not likely occur in the near future. A research structure that addresses this issue across multiple disciplines is needed.

Dr. Offit cautioned against trying to implement new discoveries in health care before they have been fully vetted. Interventions such as computed tomography scans for lung cancer, proteomic markers for ovarian cancer, and antioxidants for cancer prevention received a great deal of publicity, but clinical trials showed that they did not greatly improve diagnosis or treatment. He suggested that many questions about the validity of new genetic markers for cancer could be answered using existing clinical trial information; prospective studies might be needed to analyze how genomic information affects cancer screening. RFAs could be funded to use pre-existing cohorts in disease-specific ways. Dr. Collins agreed that this was an alternative to developing a new national cohort. However, these existing populations are not representative with respect to age, gender, geography, and other variables. There also is evidence that data on environmental exposures for most of these studies were not collected effectively. A study using existing cohorts also will not provide information on the sorts of interventions that could change risk. Dr. Stefansson noted that developing a large cohort is not necessarily an impractical idea. In Iceland, 40,000 people have been genotyped, which allows imputation of genotypes for an additional 100,000. If 10,000 more people are genotyped, the genotypes of all Icelanders could be imputed. Therefore, relatively small numbers of people would need to be genotyped to impute allele frequencies for an entire nation; genotyping approximately 2 percent of the population of the United States would be sufficient.

Dr. Chanock agreed that information on environmental exposures is needed to make full use of genetic data. Some prospective cohorts have sound data for environmental exposure, although the best way to collect and assess the quality of these data is unclear. Teaching clinicians how best to combine genetic information with family history to assess risk also will be necessary. Dr. Ioannidis noted that the studies exploring non-genetic components of disease risk have not always been successful; the results from a number of epidemiological studies have not held true. A different approach could be to use cohorts linked to registries to perform randomized controlled trials for various exposures. However, combining data from different cohorts could be difficult because of inconsistencies in the ways that data are collected and variables defined. For each disease outcome, decisions will need to be made regarding biospecimens to collect, and the uniformity of diagnoses must be considered. Such studies will need to be carefully planned to be able to use the data to determine thresholds for intervention and cost-effective screening methods. Dr. Kardia suggested that lessons could be learned from the Children's Study, which showed that an effective leadership model is needed to bring together researchers from different disciplines; allowing each researcher to design his or her own part of a large trial is not optimal.

Dr. Khoury asked participants to evaluate the value of adding genetic information to current practice. Dr. Greenland said that he was optimistic that the public will be able to understand risk information in a rational way, and act rationally upon receiving it, although there is evidence that

most people do not truly understand statistics, probability, and risk. How the information is presented will impact people's reactions to it. Research is urgently needed to develop a better understanding of how best to present risk information to clinicians and patients. Dr. Ioannidis noted that research has shown that people react to and use information in ways dependent on the context in which it is presented. Dr. Chanock said that the scientific community does not do an effective job of conveying its findings to the general public. Care must be taken not to under-express the complexity of the results or to overestimate their impact. The ways in which risks and benefits of vaccines are communicated could be used as a model. Dr. Coates agreed that risk communication needs to be addressed in a global manner, but what is needed clinically varies across conditions. EGAPP evidence reviews found that the value that genomic information adds varies based on the biology of the disease and availability of treatment.

Dr. Simons-Morton suggested developing "sentinel" clinical trials to test principles relative to multiple diseases. She suggested a number of different trials that could inform clinical practice related to genomics including testing patient/public reaction to communication of genetic risk; screening and risk assessment approaches with or without genetic information and their effect on subsequent healthcare delivery, physician behavior, and patient outcome; whether lifestyle interventions can change genetically conferred risk; whether genetic information can be used to target interventions; and whether providing genetic information motivates behavior change. Dr. Green agreed with this approach but suggested a smaller number of trials that would prioritize diseases that impact public health and those with actionable treatments. Novel ways to enroll and communicate with participants in such trials will be needed, along with a flexible infrastructure that can change as technology changes and results are analyzed.

Dr. Stefansson noted that data published in peer-reviewed journals are often considered too uncertain to be conveyed to the public, and questioned researchers' confidence in such data. Dr. Chanock agreed that only sound data should be published, but interpreting the results for communication to the public must be done carefully. Genetic risk variants are valid, but the amount of information they provide must be clarified. Dr. Gail noted that new models have been proposed based on SNPs, and suggested using established cohorts to check the calibration of these models. Specific problems and interventions must be carefully defined to see how SNP information should be fitted in and to determine whether this information influences decision making. This will be difficult for prevention models, because at present, SNPs do not add greatly to the predictive capability of most existing models. Dr. Kardia agreed that before testing the impact of genetic information, confirmation of the validity of existing models is needed. In addition, because the U.S. population is heterogeneous, not all models may be valid in all locations. Better ways to capture population heterogeneity and different environmental exposures are needed.

Dr. Ginsburg recommended consideration of pharmacogenomics, noting that drug responses are in part genetically defined, the effects are quantifiable and have public health implications, and environmental exposures also may be involved. Dr. Ioannidis agreed that this was an important area of research, but could be complicated by miscalculation of prognoses in many studies. Information on the uncertainty of prognostic models should be publicly available.

Dr. Khoury referred to Dr. Coates' stakeholder driven collaboration to define an infrastructure needed to satisfy the needs of both the research community and society. RFAs aimed at building a knowledge base and addressing implementation should be developed. It was suggested that each NIH Institute or Center allocate a portion of its budget to addressing implementation issues.