Coriell Personalized Medicine Collaborative®: a prospective study of the utility of personalized medicine

There is a dearth of large prospective studies to determine if genetic risk factors are useful predictors of health outcomes and if reporting them to individuals or physicians changes health behavior. The Coriell Personalized Medicine Collaborative® (CPMC) is a prospective observational study with three cohorts – community, cancer and chronic disease. Participants provide detailed medical history through a dynamic internet-based portal. DNA is tested and personalized risk reports are provided for potentially actionable health conditions. To date, the community cohort has enrolled 4372 participants. The internet-based portal supplies educational content, captures phenotypic data and delivers customized risk reports. The Informed Cohort Oversight Board has approved 16 health conditions to date, and risk reports with genetic and nongenetic risks for six conditions have been released. The majority (87%) of participants who completed requisite questionnaires viewed at least one report. The CPMC is a cohort study delivering customized risk reports for actionable conditions using a web interface and measuring outcomes longitudinally.

KEYWORDS: biobank • genetic education • genome • health outcomes • personalized medicine • pharmacogenomics • population • risk

Margaret A Keller, Erynn S Gordon, Catharine B Stack, Neda Gharani, Courtney J Sill, Tara J Schmidlen, Joseph Mintzer, John Pallies, Norman P Gerry & Michael F Christman†

The Coriell Institute for Medical Research, 403 Haddon Avenue, Camden, NJ 08103, USA
†Author for correspondence:
Tel.: +1 856 757 4820
Fax: +1 856 964 0254
christman@coriell.org

Personalized medicine has the potential to significantly improve health management. The human genome represents a wealth of information, including genetic variation, that predicts susceptibility to disease and response to medications. Complex diseases such as coronary artery disease, Type II diabetes and most forms of cancer are caused by an interaction between multiple genetic and environmental factors. Genome-wide association studies (GWAS) have made significant inroads in identifying genetic variants associated with these complex conditions [1]. However, the utility of such genetic variants as predictors of disease has only recently begun to be tested. Since the relationships between genes and the environment are not well understood, the ability to create a single personal risk assessment based on this information is complex and requires additional information. Importantly, large prospective cohorts with rich phenotypic datasets, dense genome information and the ability to follow study participants longitudinally are needed to dissect these interactions and inform risk algorithms [2]. As these genotype–phenotype data are generated, the understanding of how multiple genetic risk factors interact with clinical parameters can be used to estimate disease risk, stratify patients by risk and target more aggressive prevention and screening to those with the highest risk. In addition to the potential use of genomics in the management of common complex diseases, there is probably an even larger impact to be made in the area of pharmacogenomics (PGx). It is estimated that 60% of prescription drugs are ineffective [3] and that adverse drug reactions are in the top ten leading causes of death in the USA [4]. Individual response to some commonly prescribed medications can now be predicted using PGx testing, and this information could be used to match the most effective drug and dose to the patient.

Direct-to-consumer genetic testing companies

In the last several years, personal genetic tests for complex disease risk have been offered direct to the consumer for a fee by several companies. Risk interpretation is provided 'in isolation', meaning that there is no attempt to adjust an individual's risk based on family history, age or other factors. Furthermore, these companies report risk in absolute terms using algorithms that require assumptions that are not often explained, yet have a direct impact on the reported risk and, if unsound, will result in invalid risk estimates [5]. Many of these companies offer no access to genetic counseling to aid consumers and physicians in their interpretation of complex genomic information. In addition, in some instances, risk reporting is not limited to potentially actionable conditions. While these companies have highlighted public interest in personalized medicine, they have also drawn attention to the discrepancies in current approaches to risk estimation [5–7].

Author P

future medicine part of fsg

Genome-wide association studies

It remains to be seen if gene variants identified using GWAS will be useful for risk stratification. Some recent editorials have questioned their utility for identification of critical genes, highlighting the relatively small contribution that known SNPs have made to the heritability of complex diseases [8,9]. While the so-called 'missing heritability' for complex disease poses an important scientific challenge [10], this should not be misinterpreted to mean that genetics is unimportant in the etiology of complex disease. Possible explanations for the missing heritability include rare variants and/or the epistatic relationships among genes. It is expected that ongoing research will probably uncover the causes very soon [1].

Pharmacogenomics

Pharmacogenomics is likely to be the first area of genome-informed medicine that will provide clinical utility, and its clinical implementation is likely to lead directly to improving drug efficacy and safety [11–19]. It has been known for decades that patients have a wide range of responses to medications, some of which are now known to be owing to variation in the genes involved in drug absorption, distribution, metabolism and excretion [20]. For example, certain alleles of cytochrome P450 enzyme genes are used to classify individuals as 'poor', 'intermediate', 'rapid' and 'ultra-rapid' metabolizers of specific drugs [21,22]. This information can be used to personalize drug selection and dosing. The US FDA has established guidelines for the pharmaceutical industry to follow regarding genetic variation in drug response [23] and has relabeled some prescription medications with information regarding the role of PGx in drug response [101]. Innovative programs such as the FDA's voluntary genomic data submission process [24] have prompted pharmaceutical companies to consider the role of genomic biomarkers in drug development.

Need for healthcare provider education

Existing research suggests that specifically primary care providers may not have the knowledge and training necessary to integrate genomic medicine into patient care [25–28]. This limitation may be due to any number of variables, including crowded medical school curricula, failure to integrate genetics across the curriculum, misperceptions about genetics, lack of knowledgeable faculties, a disconnection between basic sciences and clinical experiences during training, inadequate representation of genetics on certifying examinations, and a shortage of genetics professionals for providing ongoing education [29].

Efforts are underway to revamp medical education as it relates to genetics, as is evidenced by recent changes implemented by the Association of American Medical Colleges (DC, USA), the Accreditation Council for Graduate Medical Education (IL, USA) and the American Academy of Family Physicians (KS, USA) [30,31]. In addition to a need to improve basic genetic education and genomic education, interpretation of the vast amount of data that comes with genomic testing provides additional challenges for physicians. As the clinical utility of genomic information is elucidated through studies such as the Coriell Personalized Medicine Collaborative® (CPMC; Coriell Institute for Medical Research, NJ, USA), additional work will be needed to determine the extent of required educational reform and then develop and implement new educational programs.

Impact of personalized medicine

There have been very few studies examining the delivery of complex disease risk information to individuals and the subsequent effects on health behaviors and health outcomes long term. The Multiplex Initiative examined patient interest in receiving genetic testing for complex disease risk and described an approach to identify genetic variants for use in translational research [32]. They found that those who chose to participate in the research study were motivated to learn about their genetic susceptibility and intended to make lifestyle changes to improve their health [33]. The Family Healthware Impact Trial involved participants from primary care practices using an online tool to assess risk of six complex diseases due to family history alone [34]. They found that the majority of participants had a significant family history risk of chronic disease [35].

Unlike previous studies that delivered risk information to participants, the CPMC is unique in its inclusion of genetic, family history and nongenetic risk information. The CPMC is building a large cohort of participants who complete medical, family history and lifestyle questionnaires online, have access to web-based education and receive risk reports containing genetic-, family history- and lifestyle-based risk information. Study participants complete baseline and follow-up surveys to assess behavior and health outcomes. In this article, we present the

Author Proof

construction of a secure internet-based portal for two-way communication between the study and the participant, the development of a multidisciplinary panel to identify potentially actionable health conditions and the development of customized risk reports and initial data on the currently enrolled participants.

Methods

The CPMC is a longitudinal, observational study that is composed of three cohorts:

- The community cohort, which consists of individuals over the age of 18 years from the general population;
- The cancer cohort, which consists of patients with breast and prostate cancer who are recruited into the study through their oncologists;
- The chronic disease cohort which consists of participants with congestive heart failure and hypertension recruited through their primary care physician or cardiologist.

Preliminary data from the community cohort is included in this article.

Human subjects

The CPMC research study protocol was approved by the Coriell Institute Institutional Review Board (IRB) for direct recruitment of community members and employees at Coriell Institute and Cooper University Hospital (NJ, USA) for inclusion in the community cohort of the study and for enrollment of cancer patients recruited by Cooper University Hospital. The study protocol was approved by the Virtua Health System IRB for recruitment of employees of the Virtua Health System, a large community-based hospital in southern New Jersey (USA). The recruitment of breast and prostate cancer patients into the cancer cohort was approved by the Cooper University Hospital IRB. Recruitment and enrollment of breast and prostate cancer patients into the cancer cohort has been approved by the Fox Chase Cancer Center (PA, USA) IRB. Ohio State University Medical Center (OH, USA) enrollment of chronic disease patients will begin pending IRB approval.

Recruitment, consent & enrollment

Potential participants for the community cohort are made aware of the research protocol via the Coriell [102] or CPMC [103] websites or through partner internal websites, news articles and print material. No paid advertising is used.

The inclusion criteria for all participants are:

- 18 years of age or older;
- A unique (nonshared) personal email address;
- Access to the internet;
- Willingness to complete web-based questionnaires throughout the study period.

Individual and group informed consent sessions are conducted at the Coriell Institute and partner institutions, during which potential participants learn about the study through a formal presentation. At the end of the presentation, individuals are given ample time for questions. Initial consent is obtained in person and in writing; if and when the consent document undergoes a significant revision, re-consent is obtained electronically via the internet-based portal.

Participants consent that their ability to receive genetic information is contingent on their yearly completion of online questionnaires pertaining to demographics, medical history, medications, family history and lifestyle factors. In addition, participants are asked to voluntarily complete follow-up surveys regarding what they did with their results approximately 3 and 12 months after they view each disease-specific result. Furthermore, the Coriell IRB has approved an incentives program, in which participants are eligible for a random drawing to win a gift card if they are among the first participants to complete their questionnaires. The study protocol includes the potential for genotyping via multiple platforms during the course of the study. Participants are informed that they will not receive all results generated from the genotyping platforms used. The consent document includes an opt-in option for release of de-identified genetic and phenotypic data to either researchers at for-profit or non-profit institutions. For participants who opt-in to de-identified data release, Coriell will deposit the dataset into a national database (Database of Genotypes and Phenotypes [dbGaP]) where it will be archived and where access will be overseen by a data access committee. Finally, participants consent to receive information through their internet-based portal account about other studies for which they are eligible, but for which participation is voluntary.

Saliva collection, DNA extraction & genotyping

Saliva samples (2 ml) are collected at the time consent is obtained using Oragene® DNA Sample Collection Kits (DNA Genotech,

Author Proof

Ontario, Canada). Samples are processed at Coriell in the Clinical Laboratory Improvement Amendments (CLIA)-certified Genotyping and Microarray Center (NJ, USA) with all processes tracked using a custom-designed biorepository management system. DNA extraction is performed using an Agencort Biomek® NX liquid-handling robot (Beckman Coulter Inc., CA, USA), and DNA concentration is determined using a NanoDrop® ND-1000 (Thermo Fisher Scientific Inc., MA, USA) spectrophotometer. Following normalization, DNA is stored in 2D barcoded tubes at -80°C. For genotyping using the Affymetrix® (CA, USA) Genome-Wide Human SNP Array 6.0, DNA samples are prepared and hybridized following the manufacturer's instructions or as previously described [36]. After processing, chips are scanned and genotype calls are made with the Affymetrix Genotyping Console using the Birdseed 2 algorithm [36]. All CPMC samples must achieve a chip-wide call rate of at least 97%. Individual SNPs are excluded from consideration in the study if they have a 'no-call' rate in either a Hapmap validation panel or saliva validation panel of more than 5% and/or a concordance rate of less than 98%. In addition, all samples are processed on the Affymetrix DMET™ Plus array using the manufacturer's protocols and analyzed using Affymetrix DMET Console.

Health condition & genetic variant selection process

Peer-reviewed scientific and medical literature is curated to identify health conditions for which genetic variants have been associated for potential inclusion in the CPMC. Resources such as the National Office of Public Health Genomics' Human Genome Epidemiology Network (HuGENet™) [104], the National Human Genome Research Institute GWAS catalog [105] and PubMed [106] are used for this purpose. In addition, medical literature and medical society policy statements, such as those from the American Heart Association or the American Diabetes Association, are reviewed to determine if candidate health conditions are potentially actionable. For the purposes of this study, a potentially actionable condition is defined as a condition for which the risk is likely to be mitigated by individual action (behavior or lifestyle) or by medical action (screening, preventative treatment or early intervention). Currently, one SNP has been selected for genetic variant risk reporting per health condition. This approach has been adopted for two reasons. First, given that one of the primary goals of the CPMC is to educate and, in order to understand the level of comprehension of the study participants, we opted to start by reporting genetic risk in the simplest format, as the risk associated with one of potentially many risk variants contributing to the total genetic risk for the disease. The multifactorial nature of common diseases and the fact that the risk variant reported accounts for only a small fraction of the total genetic risk for the health condition is highlighted in the 'Genetic Variant' genotype results tab and disease 'Causes' section under the 'About' tab of the risk report (see Figure 1). Second, since the CPMC reports risk estimates directly from published peer-reviewed papers, the ability to report a combined multivariant risk is limited by the availability of such publications. As the study moves forward, we plan to increase the number of risk variants reported per health condition. In addition, as multigenic risk models are developed and published in peer-reviewed literature, these will be considered for release to study participants. Selection of the specific genetic variant for risk reporting per health condition is based on the amount of supportive evidence. The minimum inclusion criteria for candidate genetic risk variants is a documented association with the disease in more than one cohort of the same race, either replicated within a single peer-reviewed publication or published in separate peer-reviewed publications. SNPs are then prioritized according to the greatest amount of supportive evidence. SNPs with multiple independent published reports and preferably additional supportive meta-analysis are preferentially chosen. If more than one equally well-supported risk variant is available, the SNP with the greater relative risk (RR) for disease will be selected. If the RR estimates are equivalent for more than one genetic variant, then an arbitrary selection is made of one of the remaining risk variants. Currently, only genetic variants contained on the Affymetrix 6.0 or DMET Plus GeneChip® platforms have been considered. Additional variants not included on these platforms will be considered in the future, with the plan to add custom-designed genotyping assays to the project. Coriell prepares brief synopses of candidate health conditions and associated genetic variants for review and voting by the Informed Cohort Oversight Board (ICOB).

Informed Cohort Oversight Board

The Informed Cohort Oversight Board is an independent advisory board that includes scientists skilled in genetic research, medical

Author Proof

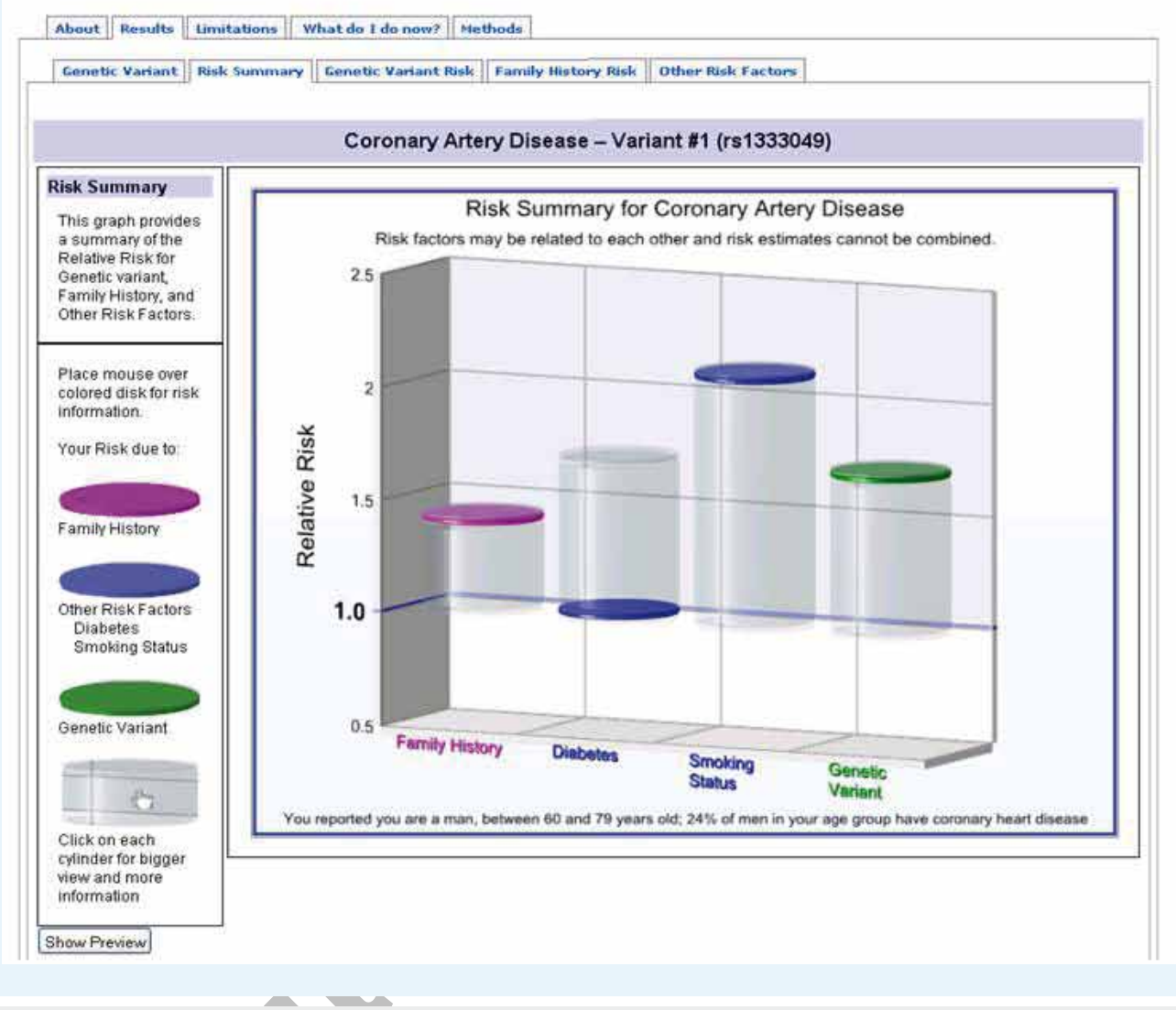

Figure 1. Sample Coriell Personalized Medicine Collaborative® (CPMC) risk summary report. A sample risk report is shown for coronary artery disease. The risk summary includes relative risk assessments for family history (magenta cylinders), nongenetic factors (blue cylinders) and genetic variant risk (green cylinders). The height of the gray cylinders depicts the range of risk. In the case of family history risk, those with a family history of coronary artery disease (RR: 1.2 for females, 1.4 for males) are compared with the reference group (no family history of coronary artery disease, RR: 1.0) [47]. In the case of co-morbidity with diabetes, those with diabetes (RR: 1.7 for males, 2.4 for females) are compared with the reference group (no diabetes, RR: 1.0) [48]. In the case of smoking risk, smokers (RR: 2.1 for males, 2.6 for females) are compared with the reference group (nonsmokers, RR: 1.0) [48]. Finally, in the case of the genetic variant risk due to SNP rs1333049, those with one (RR: 1.3) or two (RR: 1.7) copies of the risk allele are compared with the reference group (no copies of the risk allele, RR: 1.0) [46]. The placement of the colored disc represents the risk based on the information supplied in the Medical History/Medication, Family History and Lifestyle Questionnaire or genotyping data. The tabs for individual 'Genetic Variant Risk', 'Family History Risk' and 'Other Risk' contain only the cylinders within that category with supporting text. The 'About' tab contains information on the disease prevalence and heritability, and the 'Genetic Variant' tab contains information on the genotype frequency. The 'Methods' and 'Limitations' tabs contain technical information, and the 'What do I do now?' tab supplies links to supporting web pages or view a short video about the condition or through which to request genetic counseling.

professionals familiar with the use of genetics in medical care, ethicists and community members. The ICOB meets twice a year to review conditions and genetic variants selected by the CPMC staff. The ICOB is charged with assessing whether or not the condition is potentially actionable and whether the associations between the genetic variants and disease conditions are robust. The ICOB votes on condition and genetic variant submissions, and the majority vote rules. Submissions not approved by the ICOB can be resubmitted in the future if newly published peer-reviewed data strengthens the evidence for potential actionability or genetic association. Submissions approved by the ICOB are developed into CPMC risk reports that are released

to CPMC participants who have completed the required web-based questionnaires. The ICOB model was adopted from that proposed by Kohane *et al.* in 2007 [37].

Pharmacogenomics Advisory Group

The CPMC has convened a separate Pharmacogenomics Advisory Group to advise how PGx-relevant gene information could best be structured in pharmacogenomic reports. This group is made up of pharmacists, pharmacologists and clinicians with experience in PGx. As with the ICOB, the Pharmacogenomics Advisory Group will convene twice a year to review information compiled by CPMC staff on medications where efficacy or adverse events are known to be affected by variation in PGx genes. Those approved for inclusion in the study will be the topic of future reports delivered to participants through their internet-based portal account.

CPMC internet-based portal

The CPMC internet-based portal is freely accessible to the public and provides information about the study, how to enroll, health conditions included in the study, genetic education and sample risk reports [101]. The 'How It Works' section explicitly describes the steps involved in participation, and the 'How to Enroll' section allows for electronic registration for upcoming enrollment events for the community cohort. Two additional sections of the internet-based portal, 'Genetic Education' and 'Health Conditions', have educational content and links to other web-based educational resources for participants and the public. The 'Health Conditions' web page is dynamic, with pages added as health conditions are approved by the ICOB and as risk reports are developed. A portion of the website contains detailed genetic education pages written for the medical professional.

In addition to providing general information about the study and detailed educational material, the internet-based portal also allows participants to access their secure account using their username and password. This allows participants to update their contact information, change their password and complete demographic, medical history, family history, lifestyle and medication information as well as optional questionnaires. Participants can also view any available personalized results from this page. Furthermore, participants can input contact information for physicians to whom they want their results released, schedule an appointment with a CPMC genetic counselor free of charge or register for an upcoming educational session. Availability of newly released reports is announced to participants via email. However, no results are shared by email. Participants must log on to the secure internet-based portal to view their results. Participants view their risk reports '*à la carte*', choosing to view or not view each risk report or to speak to a genetic counselor prior to deciding whether or not to view their risk reports. Prior to initial viewing of a risk report, participants are directed to a web page where there is both written and video-based material describing the health condition, risk factors, diagnosis and treatment. These materials emphasize that these reports are not diagnostic and encourage participants who have concerns about their risk to discuss their report with a healthcare provider or a CPMC genetic counselor. In addition, a network of CPMC project pharmacists will be available by telephone or email to discuss pharmacogenomic reports with participants and doctors though a partnership with the American Pharmacists Association. Once released to participants, risk reports remain listed in the 'My Account' section for future viewing. Results of newly approved conditions are released every 4–8 weeks. Sample results are publically available from the CPMC home page [101].

Coriell has taken several steps to minimize barriers to participation due to lack of internet access or computer skills. Coriell has partnered with a large community church that makes its computer center available to participants. In addition, Cooper University Hospital offers participants access to computers in its Health Resource Center. Finally, Coriell has a CPMC telephone helpline, through which participants can ask a CPMC information technologist questions about managing their account, establishing a free email account with an outside provider or completing online questionnaires.

Web-based questionnaires

The Medical History/Medication, Family History and Lifestyle Questionnaire (MFLQ) is a detailed questionnaire, designed specifically for use in the CPMC, that captures demographics, medication information, current/past diseases, history of cancer screening, such as colonoscopy, prostate-specific antigen testing (men) and mammograms (women), pregnancy, participant lifestyle behaviors, such as smoking, exercise and alcohol use, as well as current/past diseases in first-degree relatives and grandparents. The MFLQ must be completed by participants before their sample enters the genotyping

Author Proof

queue and before they receive any risk reports. The family history section is the most extensive and represents the biggest barrier to completion of the MFLQ. Responses to the MFLQ inform the nongenetic disease risks included in the individual risk reports. Annual updates to the MFLQ are required to remain in the study and will be utilized to detect changes in health status and lifestyle.

Disease-specific outcome surveys are questionnaires designed to assess changes in health behaviors, risk perception due to genetic, family history and other risks, sharing of information, medical interventions, and anxiety in participants who have viewed a risk report. Participants are requested via email to complete an outcome survey through the internet-based portal every 3 and 12 months after viewing a risk report. There are specific outcome surveys for each health condition for which the CPMC has released a risk report. The primary end point assessed with the outcome surveys is change in health behaviors. Data from disease specific outcome surveys in combination with the MLFQ baseline and annual updates will be used to detect changes in health behaviors and health outcomes. Each participant will serve as their own control (baseline MFLQ), with short-term changes in health behavior detected on the 3-month disease-specific surveys and long-term changes detected using the 12-month disease-specific surveys and annual MFLQ update. Comparisons will be performed between risk groups to assess the association between behavior changes and each risk variable (e.g., genetic, nongenetic and family history), as well as various combinations of risk categories. Additional insights to be gained from the disease-specific surveys include data on risk perception and frequency of sharing of risk reports with healthcare providers or family members, motivations for sharing or not sharing (such as privacy concerns) and outcomes of sharing information such as referral to specialists.

Confidentiality & security

Saliva and DNA specimens are identified by barcode only. The database that contains the link from the sample barcode to participant personal information can only be accessed using trusted, secured and encrypted internal connections and is not accessible via any wireless connection. Consent documents are stored in a secure nonlaboratory storage area. Separation of CPMC datasets is used to enhance security and CPMC databases are designed to keep genetic and phenotypic data separate from consent and personal information. Certain personal information, such as login credentials and passwords, are stored in an encrypted format. All communications and connections to and from the CPMC internet-based portal are encrypted using industry-standard, secure sockets layer technology via the Hypertext Transfer Protocol Secure network protocol. The CPMC internet-based portal uses public-key pairs, private key and a certificate, issued by VeriSign® (London, UK).

To access personal information via the CPMC internet-based portal, study participants sign in using a username and password, which they create when they activate their account. A participant may retrieve a forgotten username. Passwords are stored in an encrypted format; participants may reset their passwords through the CPMC's secure system. CPMC technical support staff are available to assist individuals with these processes.

Saliva and DNA are stored at Coriell. For participants who consent to release de-identified data to the research community, de-identified datasets will be generated and uploaded to dbGaP; follow-up datasets will be added to generate longitudinal datasets via a code kept by Coriell Institute. Participants who wish to release information to a healthcare provider via the internet-based portal enter the contact information, and the provider is notified with an invitation to establish an account. Through the provider account, they can review risk reports from multiple patients. This mechanism is used for CPMC genetic counselors to view participant risk reports as part of genetic counseling sessions.

Genetic risk estimates

Risk estimates provided within participant risk reports are given as RR, and are derived or reported from a valid and representative peer-reviewed publication. The publication used for risk reporting is determined by first selecting potential studies with designs most likely to give valid results (i.e., prospective preferred over case–control; meta-analysis preferred over a single study), and next assessing study quality and whether or not the relevant risk estimates are reported in the paper. Aspects of study quality considered incorporate current published recommendations [38,39] and include the disease definition, genotyping methods and, when relevant, population stratification.

Once a peer-reviewed publication is selected for risk reporting, RRs are either reported directly or estimated based upon reported odds

Author Proof

ratios (ORs). For prospective studies, estimates of RR (including hazard ratios from survival analyses) are generally reported directly. In rare situations where a prospective study reports estimates of absolute risk, absolute risk is given either instead of or in addition to RR. When the selected study is a case–control study and ORs are reported, the degree to which the OR overstates the RR is estimated based upon an estimate of the underlying disease prevalence in the population that produced the cases and controls [40]. If the OR overstates the RR by an estimated 10% or less, then the OR is given as an estimate of RR. If the OR overstates the RR by more than 10%, then a conservative estimate for RR is calculated using the relationship [41]:

$$RR = OR/[(1-p_0) + p_0 \times OR]$$

where p_0 is the estimated overall disease prevalence.

Owing to limitations imposed by the available literature, risk estimates are generally determined for Caucasian populations. However, if racial-/ethnic-specific published risk estimates are available, these results will be reported based upon the participant demographic information in the MFLQ. All results reports provide clear indication of the source population from which risk estimates are derived. For example, an African–American participant's risk report for a given variant might include the statement 'These results are based on studies in Caucasian populations' if there was no risk information derived from one or more African–American cohorts.

Nongenetic risk estimates

Risk estimates provided within participant risk repots for nongenetic factors, such as family history and lifestyle or environmental factors, are also derived or reported directly from valid and representative peer-reviewed papers. Nongenetic factors are included if they are collected by the CPMC Medical Family History and Lifestyle Questionnaire, and accepted disease risks, based upon a review of the clinical and epidemiologic literature. For risk factors that meet these criteria, a peer-reviewed publication is selected based upon study design and study quality features, following the same strategy used to select the publication used for genetic risk estimates. Since most of the nongenetic risk factors have been studied prospectively, RRs are generally reported directly from the selected publication. In cases where RRs are not reported, they are estimated from ORs using the same methods used when estimating genetic risk.

Web-based risk reports & their release

Risk reports include information on heritability, prevalence, methods, references and limitations, organized in a tab structure (see Figure 1). The gray cylinders indicate the range of possible risks for each factor, and the solid disc represents the RR for individuals in that risk group. The 'What do I do now?' tab includes the ability to review the educational video or health condition summary web page, register for an upcoming educational session focused on the reported disease, or schedule an appointment with a CPMC genetic counselor.

When risk data are available in non-Caucasian populations, it is utilized in risk reports; when it is not available, RR derived from Caucasian populations is presented and noted as such. For some health conditions, there are quantifiable, nongenetic risk factors that are not available through the MFLQ for inclusion in the CPMC risk report. These other risks are described in the educational material and limitations pages. As it is believed that participants previously diagnosed with conditions included in the CPMC may be interested in their genetic risk status, risk reports are provided for 'information purposes only'.

When information entered through the MFLQ is missing or incomplete or when the genotype is not known (due to a genotype failure or 'no call'), the risk report includes an empty gray cylinder for the factor with missing information along with an explanation for the missing data.

The genetic variant risk report is reviewed and approved by the CLIA laboratory director. The generation of the risk report is tested in-house with predefined combinations of genetic and nongenetic characteristics. The concordance of genotype calls and risk report genotypes are reviewed prior to release.

Educational components

The CPMC internet-based portal contains educational content relating to genetics and complex disease for the lay public and participants and, enriched materials for medical professionals. Participants are invited to attend educational sessions focused on health conditions for which the study has released risk reports. These sessions are free and are hosted by a CPMC genetic counselor and a physician from a CPMC hospital partner who specializes in the condition. Video or audio from these sessions will be made available through the CPMC internet-based portal.

Author Proof

Results

Recruitment

Between December 2007 and December 2009, enrollment of community members and employees of Cooper University Hospital and Virtua Health System reached 4372 individuals. To date, 88% of participants opted to allow release of de-identified genetic and phenotypic data to non-profit entities and 78% to for-profit institutions.

Of all those enrolled, 2809 participants have completed the demographic portion of their MFLQ. The characteristics of these individuals are shown in TABLE 1. The community cohort is largely female (63%), well educated and white (92%). The distribution of professions (TABLE 2) shows that over a quarter (26.%) of these participants are health professionals, including 8% nurses and 5% physicians, and 12% are employed in the field of education.

ICOB actions

The ICOB has met four times, and a compiled summary of their actions is shown in TABLES 3 & 4. The committee approved 16 health conditions and one or more variants associated with each condition as well as six genes associated with prescription drug response. They rejected two health conditions, atrial fibrillation and psoriasis, as not potentially actionable. The committee based their decision on atrial fibrillation on the limited ability to screen for the condition, lack of clear guidelines for risk mitigation and no reduction in time to diagnosis. For psoriasis, the rejection was based on the consensus that presymptomatic screening would not be likely to lead to an earlier diagnosis or an improved prognosis.

Genetic testing

Of the CPMC saliva specimens processed for DNA, 1.3% yielded insufficient quantity or quality of DNA for further processing. In all such cases, participants were contacted and asked to submit a second specimen. All but one of the resubmissions received to date yielded DNA suitable for further analysis. Of the DNA samples genotyped using the Affymetrix 6.0 GeneChip, 2% failed 'contrast quality control' and were rehybridized. Of the genotyped specimens, the average call rate is 99.33%, exceeding the required 97% call rate. Two specimens had one 'no call' for a SNP utilized in the CPMC risk reports. These participants received reports that read 'Your result for this genetic variant could not be determined due to technical limitations'.

Internet-based portal & risk reports

As of December 2009, six health condition risk reports have been released to participants. TABLE 5 lists the health condition, the gene and variant, as well as the RR values due to the genotype, family history and other risk factors. The RRs shown in TABLE 5 are for Caucasian individuals.

The participation status of study participants enrolled during the approximately 2 years from the inception of the study to date is shown in TABLE 6. Of the 4372 participants, 3247 (74%) responded to an email message from the CPMC and activated their internet-based portal account. More than half of those enrolled (2809 participants or 64%) completed at least the demographic section of the questionnaires, with 1917 (43%) having completed all prerequisite questionnaires. Those who have not completed their questionnaires or activated their accounts are currently being contacted via email, then by telephone to inform them that if they wish to continue their participation, they are required to complete this information prior to receiving their

Table 1. Demographic characteristics of the Coriell Personalized Medicine Collaborative® participants†.

Characteristic	N	%
Gender		
Female	1775	63.2
Male	1034	36.8
Age range (years)		
18–29	353	12.6
30–49	921	32.8
50–69	1257	44.7
70–99	276	9.8
Did not want to answer	2	0.1
Education level		
Less than high school	10	0.4
High school	214	7.6
Some college	416	14.8
Associates degree	256	9.1
Bachelors degree	885	31.5
Graduate degree	1021	36.4
Did not want to answer	7	0.3
Race		
White	2575	91.7
Black or African–American	95	3.4
Asian	86	3.1
Native American or Alaska Native	5	0.2
Native Hawaiian or other Pacific Islander	1	0.04
Did not want to answer or no entry	47	1.7

†Includes enrolled participants as of December 2009 who have completed the demographic section of the online questionnaires; of these, 892 participants have not yet completed all online questionnaires.

Author Proof

Table 2. Professions reported by the Coriell Personalized Medicine collaborative participants†.

Profession	N	Percentage of total
Other	372	13.2
Education, training and library occupations	333	11.8
Healthcare practitioner and technical occupations:		
• Nurse	236	8.4
• Physician	134	4.8
• Technician	77	2.7
• Therapist	38	1.4
• Other	189	6.7
Management	223	7.9
Office and administrative support	180	6.4
Sales	129	4.6
Financial specialists	108	3.8
Homemaker	100	3.6
Computer and mathematical occupations	94	3.3
Life, physical and social science occupations	91	3.2
Legal occupations	78	2.8
Healthcare support	70	2.5
Architecture and engineering	57	2.0
Arts, design, entertainment, sports and media occupations	56	2.0
Unemployed	56	2.0
Community and social service	45	1.6
Construction trades	33	1.2
Food preparation and serving	19	0.7
Did not want to answer or no entry	17	0.6
Business operations specialists	16	0.6
Manufacturing	12	0.4
Unemployed due to disability	12	0.4
Military specific occupations	8	0.3
Transportation and material moving	8	0.3
Building and grounds clearing and maintenance	7	0.2
Protective services	6	0.2
Personal care and service	5	0.2
TOTAL	**2809**	**100.0**

†Includes enrolled participants as of December 2009 who have completed the demographic section of the online questionnaires; of these, 892 participants have not yet completed all online questionnaires.

customized risk reports. Of those who completed all prerequisite questionnaires, 97% had customized risk reports available and 87% have chosen to view one or more risk reports to date.

Genetic counsellor–participant interactions

Of the 1674 study participants who elected to view one or more of their risk reports as of 7 December 2009, 238 participants (14% of participants who viewed reports) requested information or counseling regarding their report(s). Participants interested in accessing genetic counseling can either submit a request for counseling through the internet-based portal (with options for telephone- or email-based contact), use the 'Email a Genetic Counselor' portal function or call a toll-free hotline. A total of 58% of requests were submitted through the CPMC internet-based portal, of which 34% requested contact by email, 24% requested contact by telephone and 42% of requests were submitted through the 'Email a Genetic Counselor' option. The reasons for requesting genetic counseling can be categorized into five themes: 'understanding risk' (5.9%), 'complex disease genetics' (11.8%), 'what do I do now?' (14.7%) 'basic genetics' (15.1%) and 'other' (52.1%). The 'other' category encompassed requests for result status, conditions or traits participants would like to receive results for in the future (e.g., celiac disease, Alzheimer's disease and alcoholism), and requests for information outside of the scope of the study, including testing for rare, Mendelian diseases or clinically available genetic testing for high-penetrance genes (e.g., *BRCA*). Review of genetic counseling inquiries to date has revealed that participants struggle with understanding basic concepts in genetics, commonly confuse RRs with absolute risks and often mistakenly attribute a greater role and risk burden to individual genetic variants in the etiology of common complex diseases. Despite these initial genomic literacy hurdles, an informal assessment of CPMC participants who have pursued genetic counseling demonstrated that these individuals recognize that they can mitigate their risk of common complex diseases through lifestyle and behavioral changes [42].

Ancillary studies

Through the internet-based portal, CPMC participants are made aware of ancillary studies in which they may choose to participate. An ancillary project and collaboration with Fox Chase Cancer Center is assessing baseline knowledge and level of change in knowledge of genetics of complex disease during participation in the study. Another ancillary project, this one conducted in collaboration with researchers from the University of Pennsylvania School of Medicine (PA, USA), examines individuals' risk perceptions and reasons for choosing whether to enroll in the CPMC study (NIH # 1RC1HG005369-01). A number of other ancillary projects are under review and development.

Discussion

CPMC Cohort

The CPMC cohort is made up of community, cancer and chronic disease groups. The

Proof

community cohort includes both community members recruited directly by Coriell Institute and employees of some CPMC healthcare partners. As the success of personalized medicine will hinge on the ability to integrate genetic data into health management strategies, we focused on recruitment of employees of Cooper University Hospital and Virtua Health System with the notion that a personally engaged population would be more motivated to learn about personalized medicine benefiting both themselves and their patients. More than a quarter of the cohort are health professionals, and the majority of participants have a college degree. While this population presents opportunities in terms of utilization of personalized medicine, it is not representative of the general public in terms of education level and interest in health issues. The cohort is overwhelmingly Caucasian (92%), in spite of efforts to recruit minority participants. Recruitment efforts include hosting recruitment events at local churches with high African–American attendance and at Philadelphia's premier science museum, The Franklin Institute, (PA, USA) during their free-admission community nights, which are primarily advertised to minority communities. We are developing additional strategies to engage minority communities to increase participation. Recruitment of participants into the cancer cohort began in January 2010 and recruitment of participants into the chronic disease cohort is planned for later in 2010.

Given the rise in direct-to-consumer genetic testing companies, individuals may obtain genetic information with or without their doctor's involvement. The CPMC cohort involves all the major stakeholders and will allow studies of both individuals and healthcare providers. Our partnerships with academic medical centers for recruitment of primary care and chronic disease patients will allow us to examine how genomic information is used in a clinical setting. Recruitment of a cancer cohort through cancer clinics will not only allow for studies of breast and prostate cancer susceptibility and outcomes but will also allow for an examination of the differences in understanding of genetic information and perception of risk between individuals with and without cancer. In addition, the CPMC has established an infrastructure within which ancillary studies can be proposed and performed. Finally, the CPMC will generate a rich genotypic and phenotypic dataset for research by the broader scientific community. As participant records will be associated with an enrollment site, the heterogeneity of the entire cohort can be managed at the time of the analysis. The availability of data for use by researchers outside of the CPMC is contingent upon the individual release of data by participants to for-profit

For Proof

Table 3. Health condition-related outcomes of the Informed Cohort Oversight Board's deliberations†.

Health condition	Potentially actionable	Genetic variant	Gene name/region	Association approved
Age-related macular degeneration	Yes	rs10490924	*ARMS2*	Yes
Atrial fibrillation	No	rs2200733	Intergenic region near *PITX2* and *ENPEP*	NA
Bladder cancer	Yes	rs9642880	*MYC*	Yes
Breast cancer	Yes	rs2981582	*FGFR2*	Yes
Chronic obstructive pulmonary disease	Yes	rs1828591 rs8034191	*HHIP* Intergenic region near *CHRNA3* and *CHRNA5*	Yes
Coronary artery disease	Yes	rs1333049	Intergenic region near *CDKN2A* and *CDKN2B*	Yes
Colorectal cancer	Yes	rs6983267	Intergenic region on 8q24.21	Yes
Inflammatory bowel disease	Yes	rs11209026	*IL23R*	Yes
Iron overload or hemochromatosis	Yes	rs1800562	*HFE*	Yes
Melanoma	Yes	rs910873	*PIGU*	Yes
Obesity/high BMI	Yes	rs9939609	*FTO*	Yes
Psoriasis	No	rs12191877	*HLA-C*	NA
Prostate cancer	Yes	rs16901979	Intergenic region on 8q24.21	Yes
Rheumatoid arthritis	Yes	rs6920220	Intergenic region between *OLIG3* and *TNFAIP3*	Yes
Systemic lupus erythematosis	Yes	rs3821236	*STAT4*	Yes
Testicular cancer	Yes	rs995030	*KITLG*	Yes
Type I diabetes	Yes	rs9272346	*HLA-DQA1*	Yes
Type II diabetes	Yes	rs7754840	*CDKAL1*	Yes

†*As of December 2009.*
NA: Not applicable.

Table 4. Pharmacogenomics-related outcomes of Informed Cohort Oversight Board deliberations†.

Gene	Potentially actionable
CYP2D6	Yes
CYP2C9	Yes
CYP2C19	Yes
VKORC1	Yes
UGT1A1	Yes
CYP4F2	Yes

†*As of December 2009.*

and not-for-profit companies as indicated in the consent form. While the majority of participants have agreed to release their data to at least one of these groups, this proposition raises an independent research question regarding the public perception of privacy and security as it relates to personal health information and genomic data. These questions will be investigated through our first ancillary study on the early adopters of personalized medicine, a mixed-methods study of motivations and perceptions of participants in the CPMC.

Genetic research using the internet

The CPMC study is designed around the use of a secure internet-based portal for information exchange between the study and the participant. The study utilizes information supplied via web-based questionnaires to customize risk reports with demographic, medical history, family history and lifestyle information. The web-based design makes the study scalable, as communication of risk reports occurs via a secure internet-based portal account with access to genetic counselors via telephone and email as well as in person. The internet-based portal also allows the cohort to view and consider involvement in ancillary research studies. Approximately 74% of participants who enrolled in the CPMC activated their web account, and 44% of CPMC participants completed web-based questionnaires (a prerequisite for obtaining personalized risk results). This rate of follow through exceeds rates seen in the single existing similar study, the Multiplex Initiative (a study of factors predicting interest in update of genetic susceptibility testing), in which 31.2% of participants responding to a baseline survey logged on to the study website for more information and 13.6% proceeded to genetic testing (which required a blood draw) [33]. Although the value of the cohort would be increased if participation were higher, the rate of uptake observed to date is neither surprising nor unusual.

CPMC risk reporting

Strengths of the CPMC risk-reporting system include the criteria that the variant be validated in the peer-reviewed literature and that candidates are vetted by an external oversight board. The strength of the risk report is that it includes nongenetic risk factors that are customized for the participant based on self-reported demographics, medical history, family history and lifestyle information. The participant can use this additional nongenetic risk information to put their genetic risk in perspective and have the opportunity to take action in areas where they can mitigate risk, such as in lifestyle choices like smoking. Emphasis of the role of family history in disease risk and health management is a positive outcome of participation in this study and may help individuals review their family history with their family members and communicate this information to their healthcare providers. At this time, validated, predictive models based upon well-studied, prospective populations with known genetic and nongenetic risk factors have not been developed. As participants consent to long-term participation in the study, which currently has no end date and will continue for a minimum of 5 years, the CPMC cohorts will provide valuable prospective data to advance this area of research.

Conclusion

We have begun a unique, web-based study that involves enrolling participants for longitudinal, phenotypic data collection and personalized risk assessments for complex diseases and drug responses. The CPMC cohort will be useful for testing the predictive nature of genetic tests for complex disease and drug response and for assessing the impact of genetic and nongenetic risk on health behaviors and health outcomes. It includes public enrollment as well as recruitment and enrollment from primary care, chronic disease and cancer care settings. The infrastructure of the CPMC will allow collaborations, release of de-identified datasets and support of ancillary studies. The study includes educational components for both participants and healthcare providers, and engages physicians in the study as both participants and web content reviewers. Participants are offered access to genetic counseling at no cost, via telephone, email or in person.

The CPMC provides a model in which to examine the impact of personalized medicine in the context of an observation study that takes place in a real-world setting. The importance of examining how genome-informed medicine

Author Proof

will impact behavioral and clinical outcomes in the context of nonacademic centers and community hospitals has recently been noted [43]. The CPMC has been recognized as a model for personalized medicine research, as shown by the inclusion of Coriell researchers in a recently funded implementation planning grant for educational, behavioral and social studies for translation of genetic factors in common diseases (1U34DK084548–01). The education of healthcare providers in genomics, PGx and translational medicine is essential in order for the full potential of personalized medicine to be born out. While the CPMC has taken steps to educate healthcare providers, additional work is needed.

Leroy Hood coined the term 'P4 medicine' to describe health management in the era of personalized medicine, where care would be predictive, preventive, personalized and participatory [44,107]. The CPMC is implementing P4 medicine, where scientists are partnered with healthcare professionals and individuals interested in identifying personal risks and implementing strategies in an attempt to mitigate those risks. Recently, a workshop sponsored by the NIH and the Centers for Disease Control and Prevention reviewed the use of personal genetic information for disease risk assessment and prevention [45]. Their recommendations included a call to assemble large cohorts to be studied by multidisciplinary teams of researchers and include examination of individual and population subgroups' perceived risk based on personal genomic information. The CPMC represents one such cohort well poised to address the growing number of questions surrounding the utility of personal genetic information. Through targeted collaborations and ancillary studies, the CPMC has the potential to advance research in the areas of translational genomics and PGx. Analysis of CPMC cohort data may contribute to development of new methods to detect gene–gene and gene–environment interactions, moving our understanding of the causes of complex disease forward. In addition, since genetic variation at more than 2000 sites of variation in 225 genes involved in drug transport, absorption and metabolism will be studied in all CPMC cohorts, data from CPMC participants will be used to characterize the overlap of important

Proof

Table 5. Genetic variant, family history and other factor relative risks for the Coriell Personalized Medicine Collaborative risk reports released as of December 2009.

Health condition	Gene or region	SNP	Risk allele, nonrisk allele	Genotype relative risk†	Family history relative risk‡ (yes vs no)	Other factors relative risks	Ref.
Coronary artery disease	Region between *CDKN2A* and *CDKN2B*	rs1333049	C,G	1.3 (CG vs GG), 1.7 (CC vs GG)	1.4(m), 1.2(f)	Diabetes§: 1.7(m), 2.4(f) Current smoker: 2.1(m), 2.6(f)	[46–48]
Type 2 diabetes	*CDKAL1*	rs7754840	C,G	1.2 (CG vs GG), 1.3 (CC vs GG)	1.9	BMI¶: 2.3 (25–29.9 vs <25), 5.9 (≥30 vs <25)	[49,50]
Iron overload/ hemochromatosis	*HFE*	rs1800562	A,G	AA: 33–57%(m), 3–16%(f), AG: 0–5%(m), 0–1%(f) GG: 0–4%(m), 0–1%(f)	NR	NR	[51]
Melanoma	*PIGU*	rs910873	T,C	1.7 (CT vs CC), 3.0 (TT vs CC)	2.2	NR	[52,53]
Prostate cancer	8q24.21 intergenic region	rs16901979	A,C	1.5 (CA+AA vs CC)	1.9	NR	[54,55]
Age-related macular degeneration	*LOC387715*	rs10490924	T,G	2.4(GT vs GG), 6.0(TT vs GG)	3.9	Smoking: 1.4 (former vs never), 2.1 (current vs never)	[56–58]

†Absolute risk reported for iron overload/hemochromatosis.
‡Positive family history defined as one or both parents dying owing to coronary artery disease, one or both parents with Type II diabetes, one or more first-degree relatives with melanoma, biological father or any brothers diagnosed with prostate cancer, and one or more first-degree relatives with age-related macular degeneration.
§Comorbidity with diabetes.
¶BMI in kg/m^2.
m: Male; f: Female; NR: None reported.

Table 6. Breakdown of participant activities.

Participant description	N	Percentage relative to total enrolled participants†	Percentage relative to total participants with activated web accounts	Percentage relative to total participants with completed questionnaires
Those enrolled from study start to 7 December 2009†	4372	100	NA	NA
Those with activated web account‡	3247	74	100	NA
Those with partial web-based questionnaires completed§	2809	64	87	NA
Those with all required web-based questionnaires completed	1917	44	59	100
Those with available customized risk reports on web¶	1861	43	57	97
Those who viewed one or more customized risk reports	1674	38	52	87

†Activities are presented for participants enrolled from study inception (4 December 2007) to the time the manuscript was drafted (7 December 2009).
‡The Coriell Personalized Medicine Collaborative® internet-based portal was launched approximately13 months after inception (30 January 2009).
§This number refers to those who have partially completed the Coriell Personalized Medicine Collaborative medical history/medication, family history and lifestyle questionnaire, completing at least the demographic portion of the questionnaire.
¶There is a several weeks long lag time between completion of questionnaires to completion of DNA isolation and genotyping.
NA: Not applicable.

pharmacogenetic variants within individuals and contribute to the understanding of the potential for personalizing treatments.

Future perspective

Personalized medicine is likely to evolve in the next 5–10 years in the following ways:

- A single genetic test, most probably whole-genome sequencing, that catalogs genetic variation in an individual will be performed at birth;
- An independent research study examining physician understanding of genomic information, clinical utility and educational resources needed by the medical community will be conducted. The results of this will be used to inform personalized disease prevention and screening strategies and to optimize treatment. Expert software systems will be developed that capture quantitative information from electronic records to provide personalized risk estimates;
- Guidelines for curation of genomic data will be developed, as well as a system for widespread implementation of these guidelines for use in reporting of genomic information. As smaller groups begin this work, some effort will need to be made to unite these groups under a lead organization;
- A person's genomic information will be referenced when medications are prescribed such that drug–gene interactions are taken into account at the time of drug selection and dosing;
- Electronic health records will be adapted to access this genomic information as needed;
- Education and continuing education of health professionals will be adapted to incorporate the growing role of genomics in health management;
- Release of de-identified, longitudinal datasets from large cohorts will allow for the development of multigenic models of complex disease risk;
- Whole-genome sequencing will probably uncover the missing heritability that GWAS have been unable to find;
- Pharmaceutical companies will adapt drug development and clinical trial management to include genomic information such that new drugs will be optimized for smaller groups of patients, with increased efficacy and decreased adverse events.

Acknowledgements

We are extremely grateful to the participants of the Coriell Personalized Medicine Collaborative® for their ongoing commitment to the study. We acknowledge the physicians of Cooper University Hospital (NJ, USA) and Virtua Health System (NJ, USA) who assisted with health condition web and survey content. We thank the staff of the Coriell Genotyping and Microarray Center and of the Coriell Information Systems Department for their significant contributions to the study. Finally, we acknowledge the members of the Informed Cohort Oversight Board (Erin O'Shea, Jennifer Hoheisel, Marc Lenburg, Steve Murphy, Kenneth Offit, David Pellman, Charles Rotimi, Reverend

Author Pro

Floyd White and Ellis Neufeld), for their advice and suggestions concerning the project. We thank Scott Megill and Steven Madore for critical review of the manuscript.

Financial & competing interests disclosure

This work is supported by grants from the Rohrer Foundation, the RNR Foundation, private philanthropy and the Coriell Institute for Medical Research. NIH grants 1RC1HG005369-01 (Erynn S Gordon) and 1U34DK084548-01 (Margaret A Keller and Erynn S Gordon) support the ancillary studies on CPMC participants risk perception and on the development of web-based tools for physician education. The Genetic Knowledge Assessment tool was developed in collaboration with Mary Daly, MD (Fox Chase Cancer Center, PA, USA). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

Introduction

- There is a need to build large cohorts with which to study the utility of genomic information for improving human health.
- Studies of personalized medicine require partnerships between research institutions, medical centers and the community with a commitment to education at all levels.

Method

- The Coriell Personalized Medicine Collaborative® (CPMC) has been developed to recruit a population cohort, collect saliva and perform genome-wide and targeted genotyping, and assess the utility of personalized medicine.
- An external advisory panel, the Informed Cohort Oversight Board, has been established to review candidate health conditions and SNPs associated with the conditions proposed for inclusion in the study.
- A separate external board with appropriate expertise will advise on which gene–drug pairs merit inclusion in the pharmacogenomics aspect of the study.
- A secure internet-based portal is used by participants to enter medical history, family history and lifestyle information and to receive personalized risk reports. The internet-based portal contains educational material for both participants and medical professionals and includes anticipatory guidance videos and a system to enter requests for genetic counseling free of charge.
- A systematic approach to estimating relative risks of complex diseases based on genetic variants, family history, comorbidities and lifestyle factors has been developed.

Results

- The CPMC has enrolled over 4372 participants into the community cohort, which includes employees of healthcare systems.
- To date, the cohort is mostly Caucasian, well-educated and enriched for educators and health professionals.
- The Informed Cohort Oversight Board has convened four times and has approved 16 health conditions and six pharmacogenomic genes.
- The study has received requests for genetic counseling from 14% of participants who have viewed their results.

Conclusion

- The infrastructure is in place and a large cohort has been enrolled with which to study many aspects of personalized medicine, from genetic knowledge and risk perception, to intent to change health behaviors and long-term health outcomes.
- The internet-based portal, email and in-person communication between the study and the study participants allows for a multimedia approach to education focused on how genomics influences a person's risk of complex disease.

Bibliography

Papers of special note have been highlighted as:
- ▪ of interest
- ▪▪ of considerable interest

1 Manolio TA, Collins FS: The HapMap and genome-wide association studies in diagnosis and therapy: *Annu. Rev. Med.* 60, 443–456 (2009).

▪▪ **Encompasses the successes of genetic association studies as well as the privacy issues surrounding release of genetic information and how it might be used by patients and physicians.**

2 Collins F: Francis Collins interview. Departing U.S. genome institute director takes stock of personalized medicine. Interview by Jocelyn Kaiser. *Science* 320, 1272 (2008).

3 Spear BB, Heath-Chiozzi M, Huff J: Clinical application of pharmacogenetics. *Trends Mol. Med.* 7, 201–204 (2001).

4 Lazarou J, Pomeranz BH, Corey PN: Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA* 279, 1200–1205 (1998).

5 Yang Q, Flanders WD, Moonesinghe R, Ioannidis JP, Guessous I, Khoury MJ: Using lifetime risk estimates in personal genomic profiles: estimation of uncertainty. *Am. J. Hum. Genet.* 85(6), 786–800 (2009).

6 Mihaescu R, van Hoek M, Sijbrands EJ *et al.*: Evaluation of risk prediction updates from commercial genome-wide scans. *Genet. Med.* 11, 588–594 (2009).

▪ **Discusses the limitations of disease risk estimates based on genetic variants that continue to be investigated in discovery and replication studies.**

7 Ng PC, Murray SS, Levy S, Venter JC: An agenda for personalized medicine. *Nature* 461, 724–726 (2009).

8 Goldstein DB: Common genetic variation and human traits. *N. Engl. J. Med.* 360(17), 1696–1698 (2009).

9 Kraft P, Hunter DJ: Genetic risk prediction – are we there yet? *N. Engl. J. Med.* 360(17), 1701–1703 (2009).

10 Maher B: Personal genomes: the case of the missing heritability. *Nature* 456, 18–21 (2008).

11 Burczynski ME, Oestreicher JL, Cahilly MJ *et al.*: Clinical pharmacogenomics and transcriptional profiling in early phase oncology clinical trials. *Curr. Mol. Med.* 5, 83–102 (2005).

12 Chung WK: Implementation of genetics to personalize medicine. *Gend. Med.* 4, 248–265 (2007).

13 Lesko LJ, Woodcock J: Pharmacogenomic-guided drug development: regulatory perspective. *Pharmacogenomics J.* 2, 20–24 (2002).

14 McCarthy AD, Kennedy JL, Middleton LT: Pharmacogenetics in drug development. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1579–1588 (2005).

15 Relling MV, Hoffman JM: Should pharmacogenomic studies be required for new drug approval? *Clin. Pharmacol. Ther.* 81, 425–428 (2007).

16 Roses AD: Pharmacogenetics in drug discovery and development: a translational perspective. *Nat. Rev. Drug Discov.* 7, 807–817 (2008).

17 Tonon G, Anderson KC: Moving toward individualized cancer therapies. *Clin. Cancer Res.* 14, 4682–4684 (2008).

18 Vizirianakis IS: Clinical translation of genotyping and haplotyping data: implementation of *in vivo* pharmacology experience leading drug prescription to pharmacotyping. *Clin. Pharmacokinet.* 46, 807–824 (2007).

19 Winston A, Hatzimichael E, Marvin V, Stebbing J, Bower M: Host pharmacogenetics in the treatment of HIV and cancer. *Curr. Drug Saf.* 1, 107–116 (2006).

20 McLeod HL, Evans WE: Pharmacogenomics: unlocking the human genome for better drug therapy. *Annu. Rev. Pharmacol. Toxicol.* 41, 101–121 (2001).

21 Belle DJ, Singh H: Genetic factors in drug metabolism. *Am. Fam. Physician* 77, 1553–1560 (2008).

22 Ingelman-Sundberg M, Rodriguez-Antona C: Pharmacogenetics of drug-metabolizing enzymes: implications for a safer and more effective drug therapy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1563–1570 (2005).

23 Ratner M: FDA pharmacogenomics guidance sends clear message to industry. *Nat. Rev. Drug Discov.* 4, 359 (2005).

24 Lesko LJ, Woodcock J: Translation of pharmacogenomics and pharmacogenetics: a regulatory perspective. *Nat. Rev. Drug Discov.* 3, 763–769 (2004).

25 Greendale K, Pyeritz RE: Empowering primary care health professionals in medical genetics: how soon? How fast? How far? *Am. J. Med. Genet.* 106, 223–232 (2001).

26 Scheuner MT, Sieverding P, Shekelle PG: Delivery of genomic medicine for common chronic adult diseases: a systematic review. *JAMA* 299, 1320–1334 (2008).

27 Lubin IM, McGovern MM, Gibson Z *et al.*: Clinician perspectives about molecular genetic testing for heritable conditions and development of a clinician-friendly laboratory report. *J. Mol. Diagn.* 11, 162–171 (2009).

28 Watson EK, Shickle D, Qureshi N, Emery J, Austoker J: The 'new genetics' and primary care: GPs' views on their role and their educational needs. *Fam. Pract* 16, 420–425 (1999).

29 McInerney JD: Genetics education for health professionals: a context. *J. Genet. Couns.* 17, 145–151 (2008).

30 Riegert-Johnson DL, Korf BR *et al.*: Outline of a medical genetics curriculum for internal medicine residency training programs. *Genet. Med.* 6, 543–547 (2004).

31 American Academy of Family Physicians: *Recommended Curriculum Guidelines for Family Medicine Residents: Medical Genetics.* American Academy of Family Physicians, KS, USA (1999).

32 Wade CH, McBride CM, Kardia SL, Brody LC: Considerations for designing a prototype genetic test for use in translational research. *Public Health Genomics* 13(3), 155–165 (2010).

33 McBride CM, Alford SH, Reid RJ, Larson EB, Baxevanis AD, Brody LC: Characteristics of users of online personalized genomic risk assessments: implications for physician–patient interactions. *Genet. Med.* 11, 582–587 (2009).

▪ **Examines the motivations of those who chose to participate in the Multiplex Initiative, a pilot study that delivered genetic risk information for complex diseases to individuals.**

34 Yoon PW, Scheuner MT, Jorgensen C, Khoury MJ: Developing family healthware, a family history screening tool to prevent common chronic diseases. *Prev. Chronic Dis.* 6, A33 (2009).

35 O'Neill SM, Rubinstein WS, Wang C *et al.*: Familial risk for common diseases in primary care: the Family Healthware Impact Trial. *Am. J. Prev Med.* 36, 506–514 (2009).

36 McCarroll SA, Kuruvilla FG, Korn JM *et al.*: Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat. Genet.* 40, 1166–1174 (2008).

37 Kohane IS, Mandl KD, Taylor PL, Holm IA, Nigrin DJ, Kunkel LM: Medicine. Reestablishing the researcher–patient compact. *Science* 316, 836–837 (2007).

▪ **Describes the need for oversight surrounding genetic studies and many of the principles presented here were operationalized in the Coriell Personalized Medicine Collaborative® Informed Cohort Oversight Board.**

38 Attia J, Ioannidis JP, Thakkinstian A *et al.*: How to use an article about genetic association: C: what are the results and will they help me in caring for my patients? *JAMA* 301, 304–308 (2009).

39 Ioannidis JP, Boffetta P, Little J *et al.*: Assessment of cumulative evidence on genetic associations: interim guidelines. *Int. J. Epidemiol* 37, 120–132 (2008).

▪ **Dissects genetic association studies and ranks the factors that influence their quality.**

40 Davies HT, Crombie IK, Tavakoli M: When can odds ratios mislead? *BMJ.* 316, 989–991 (1998).

41 Zhang J, Yu KF: What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *JAMA* 280, 1690–1691 (1998).

42 Schmidlen T, Gordon ES, Christman M: Genomic literacy: emerging themes among genetic counseling inquiries from participants of the Coriell Personalized Medicine Collaborative. *J. Genet. Couns.* 18, [abstract] (2009).

43 Ray T: Comparative effectiveness and personalized medicine can unite in community-based research, Woodcock says. *Genome Web* 4th November (2009).

44 Hood L: A doctor's vision of the future of medicine. *Newsweek* 13th July (2009).

45 Khoury MJ, McBride CM, Schully SD *et al.*: The Scientific Foundation for personal genomics: recommendations from a National Institutes of Health-Centers for Disease Control and Prevention multidisciplinary workshop. *Genet. Med.* 11, 559–567 (2009).

46 Schunkert H, Gotz A, Braund P *et al.*: Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. *Circulation* 117, 1675–1684 (2008).

Author Proof

47 Myers RH, Kiely DK, Cupples LA, Kannel WB: Parental history is an independent risk factor for coronary artery disease: the Framingham Study. *Am. Heart J.* 120, 963–969 (1990).

48 D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P: Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA* 286, 180–187 (2001).

49 Meigs JB, Shrader P, Sullivan LM *et al.*: Genotype score in addition to common risk factors for prediction of Type 2 diabetes. *N. Engl. J. Med.* 359, 2208–2219 (2008).

50 Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB Sr: Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med.* 167, 1068–1074 (2007).

51 Allen KJ, Gurrin LC, Constantine CC *et al.*: Iron-overload-related disease in HFE hereditary hemochromatosis. *N. Engl. J. Med.* 358, 221–230 (2008).

52 Brown KM, Macgregor S, Montgomery GW *et al.*: Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat. Genet.* 40, 838–840 (2008).

53 Cho E, Rosner BA, Feskanich D, Colditz GA: Risk factors and individual probabilities of melanoma for whites. *J. Clin. Oncol* 23, 2669–2675 (2005).

54 Zheng SL, Sun J, Wiklund F *et al.*: Cumulative association of five genetic variants with prostate cancer. *N. Engl. J. Med.* 358, 910–919 (2008).

55 Giovannucci E, Liu Y, Platz EA, Stampfer MJ, Willett WC: Risk factors for prostate cancer incidence and progression in the health professionals follow-up study. *Int. J. Cancer* 121, 1571–1578 (2007).

56 Conley YP, Jakobsdottir J, Mah T *et al.*: *CFH*, *ELOVL4*, *PLEKHA1* and *LOC387715* genes and susceptibility to age-related maculopathy: AREDS and CHS cohorts and meta-analyses. *Hum. Mol. Genet.* 15, 3206–3218 (2006).

57 Smith W, Mitchell P: Family history and age-related maculopathy: the Blue Mountains Eye Study. *Aust. N. Z. J. Ophthalmol* 26, 203–206 (1998).

58 Cong R, Zhou B, Sun Q, Gu H, Tang N, Wang B: Smoking and the risk of age-related macular degeneration: a meta-analysis. *Ann. Epidemiol.* 18, 647–656 (2008).

Websites

101 Drugs@FDA
www.accessdata.fda.gov/Scripts/cder/DrugsatFDA

102 Coriell Institute for Medical Research
www.coriell.org

103 Coriell Personalized Medicine Collaborative® (CPMC)
http://cpmc.coriell.org

104 Human Genome Epidemiology Network (HuGENet™)
www.cdc.gov/genomics/hugenet/default.htm

105 Genome.gov: A Catalog of Published Genome-Wide Association Studies
www.genome.gov/26525384

106 PubMed
www.ncbi.nlm.nih.gov/pubmed

107 Institute for Systems Biology: Predictive Preventive Personalized and Participatory
www:systemsbiology.org/Intro_to_ISB_and_Systems_Biology/Predictive_Preventive_Personalized_and_Participatory

Author Proof