Rapporteurs: Cecelia Tamburro and Catherine Sillari
Program Analysts, NHGRI

Genomic Medicine XII: Genomics and Risk Prediction
May 6-7, 2019
Sheraton Silver Spring, Silver Spring, MD
https://www.genome.gov/event-calendar/genomic-medicine-xii-genomics-and-risk-prediction
Meeting Summary

Welcome, Introductions, and Goals of the Meeting

At Genomic Medicine XII, the Genomic Medicine Working Group (GMWG) convened leaders in genomic medicine to discuss the role of genomics in risk prediction. The goals of this meeting were to review the state of science of polygenic risk scores and how it can be improved, examine other information sources that could be integrated with genetic variant information in predicting risk, and identify research directions in development and implementation of genomic risk prediction.

Session 1: Risk prediction with and without genomics

The calculation of polygenic risk scores (PRS) provides a method to estimate the impact of multiple genomic variants on disease susceptibility and other phenotypes. The method relies on results from genome-wide association studies (GWAS). In recent years, increased numbers and size of GWAS, along with increased funding for diagnostic testing, has led to the development of more PRS. To create a PRS, SNPs are first selected from a large GWAS or other discovery study. The number of risk alleles can simply be added (unweighted) or each SNP can be weighted by an effect size. The PRS is then evaluated on how well it estimates an individual's absolute or relative risk of disease. Generally, three steps are followed for evaluating the performance of risk prediction models. First, the model is developed and calibrated to determine how well its estimates fit observed events. Next, the model is internally validated to determine how well it discriminates between those who have the disease and those who do not. Finally, the model is externally validated and decision analysis is used to determine how and when these predictions might impact actual decisions. However, the statistical methods used to measure model performance have not been standardized. In addition to using genomic data, studies have suggested that disease prediction may be enhanced by combining PRS with environmental and lifestyle data. These models may be useful for research, but their clinical application is still uncertain. Implementation science research will be necessary to determine the appropriate role of PRS in healthcare.

Though PRS are not widely used clinically, researchers have begun to investigate their use in common disease prediction. The Framingham Heart Study (FHS) coined the term "factors of risk" in 1961 to describe patient features like sex, age, blood pressure, smoking, and other measurements that modulate risk for coronary artery disease (CAD). Early risk calculations summed risk factors to create a score that patients could compare to the average score for their age and sex. Today, the American College of Cardiology and the American Heart Association recommend a model that analyzes pooled cohort data to assess 10-year risk for definite atherosclerotic cardiovascular disease (ASCVD). This risk calculation is computerized and patients with risk greater than 7.5% are typically recommended for

preventive treatment. These pooled cohort equations (PCE) assess risk by summing sex-specific PCE deviations from the mean in a Cox proportional hazards model. Initial studies suggest that adding a 57-SNP PRS to the PCE improves model discrimination, and that the PRS is independent of risk associated with traditional factors. This suggests that PRS may make a significant contribution to risk estimation for ASCVD in the Framingham Heart Study.

A significant limitation of genomic risk prediction is its unknown generalizability in non-European Ancestry (EA) populations. In breast cancer, for example, African American women have different tumor mutation signatures than EA women and are more likely to get triple negative breast cancer, the most aggressive and least understood breast cancer subtype. Data from case-control studies in Nigeria, Cameroon, and Uganda suggest that GWAS and PRS results from white and Asian populations do not replicate in African populations. African ancestry (AA) women also have a different set of mutations in *BRCA1* and *BRCA2* from EA women, which suggests that *BRCA1/2* mutation testing limited to mutations from EA populations is not sufficient to understand the *BRCA1/2* associated breast cancer risk in AA populations. Overall, AA women younger than 45 appear to have a higher breast cancer incidence, are more likely to have aggressive cancer, and may benefit from more frequent, earlier screening than EA women. This example highlights the critical need for further research into population-based differences in disease types, as well as the need to define disease subtypes in risk prediction.

The ClinGen Complex Disease Working Group is currently conducting a literature review of 34 representative PRS papers to better understand the potential impact of PRS on clinical care. These papers have revealed a lack of consistency in the definition of PRS, which can be used for disease risk prediction, disease or disease subtype diagnosis, and disease prognosis. This lack of consistency also applies to the factors that contribute to PRS, including the ways in which different racial groups are defined. Though frameworks such as the Genetics Risk Prediction Studies (GRIPS) statement, which was developed in 2011 by an international group of risk prediction experts, have attempted to provide reporting recommendations, score evaluation also varies widely and inconsistent reporting makes it difficult to judge score quality. These PRS papers reveal the potential value of integrating PRS and traditional risk in clinical care, but adoption of common standards across studies, like those that exist for GWAS, will be necessary for more clearly defined outcomes.

Discussion

Though the field of genomic medicine has not yet implemented PRS into clinical care, some Direct to Consumer (DTC) companies like 23andMe are already using PRS for conditions like Type 2 diabetes. Transparency about how risk scores are calculated will be crucial to PRS publication and implementation moving forward. For the 23andMe risk score, the company used its own GWAS data with self-reported diabetes phenotypes instead of published GWAS data. Because DTC companies are already using PRS, it is important that the field of genomic medicine address research gaps in PRS implementation, such as the need to validate PRS in multiple populations. This research will help determine how causal variants vary based on a patient's background and must account for how linkage disequilibrium can change

effect sizes. A better understanding of these methods may reduce the need to construct PRS for individual populations.

PRS provide an opportunity to stratify patients into the right pathways for disease treatment and prevention. For example, they could be used as a potential diagnostic tool that justifies patients' needs for an expensive treatment or drug. However, there is less consensus on whether PRS can be used to reduce screening for patients at lower risk. Some studies have suggested that patients at lower risk for conditions may not need certain interventions, like a statin or a yearly mammogram. However, others have suggested that the protective effects of lower genomically defined risk are modest and that socioeconomic factors are larger drivers of risk. Physicians are likely to be uncomfortable with decreasing population-based screening recommendations based on lower genetic risk, and evidence is needed to determine if reducing screening based on lower risk has positive health outcomes and is acceptable to patients and clinicians. A reduction in preventive screening is unlikely to happen in the US, but the potential healthcare savings and effect on patients should be considered. Doctors also have incentive to avoid harming patients by administering unnecessary treatment, especially when aggressive interventions like biopsies in many cases fail to prevent disease.

Broader adoption of PRS may be driven by either clinical trial results or strong observational data showing improved outcomes favoring implementation. If PRS are to be implemented, they must be cost effective, produce actionable results, and be easy to use for clinicians. Currently, one of the strongest determinants of PRS use by clinicians is whether the PRS score is supported in the EHR using informatic best practices. Doctors must also be able to effectively communicate risk to patients, who will be more likely to seek out or comply with screening if they understand the value of the intervention.

Session 2: Using informatics and electronic health record (EHR) data in risk prediction

EHR data can be mined in multiple ways to assist with risk prediction. Phenotypic risk scores (PheRS), which aggregate and weight phenotypes that match with a genotype or disease, are one example of this. By automating assessments of phenotypic patterns in the EHR, OMIM and HPO terms can be used to link phenotypes to disease through Phenome-Wide Association Studies (PheWAS). PheRS have been shown to differentiate cases from controls in Cystic Fibrosis and other Mendelian diseases. PheRS can also be used for variant interpretation when applied to populations with sequence data; known pathogenic variants tend to have an elevated PheRS, while known benign variants do not. Finally, PheRS can help to make diagnoses in cases where the PheRS is much higher than expected for a certain condition. This approach demonstrates the benefits of sharing rich phenotypic information amongst multiple cohorts and linking sequencing results to the EHR. While PheRS may help make diagnoses earlier, they are lowest during the early stages of disease development, before clinical signs and symptoms are manifest and documented in the EHR.

Machine learning can be an important tool to develop risk prediction models. Learned models are often more accurate than conventional risk assessment tools and can potentially identify new predictive factors. For example, when predicting the risk of post-hospitalization venous thromboembolism, models

that incorporate all possible risk factors perform better than those that use only known risk factors or conventional risk assessment tools. In some applications, adding patient genetics to clinical variables can create more accurate models. Studies have compared curated representations of PRS with 'kitchen sink' models in which all recorded information is included. Kitchen sink models have provided the best predictions. Additionally, simple, linear machine learning models are often comparable to complex ones and may be easier to use.

The Observational Health Data Sciences and Informatics (OHDSI) program is a multi-stakeholder, interdisciplinary, international collaborative that empowers communities to generate evidence collaboratively. OHDSI developed a 5-step standardized framework for developing and evaluating patient-level prediction models. This process is: 1) Define prediction problem, 2) pick suitable data, 3) select variables, 4) train model, and 5) validate internally/externally. OHDSI's access to data allows clinicians to ask specific questions about risk and test discrimination and calibration in different populations. However, it is important to consider that a patient may have very different calibration results if tested by two equally accurate models. This is because models may assign patients to different subpopulations, which affects their predicted risk. OHDSI is also in the process of creating an integrated panel that presents clinicians with a patient's estimated risk for multiple conditions. Outcomes can be tested in multiple databases which can then be compared to analyze the transportability of risk scores from one database to another. OHDSI's goal is to use precision medicine to identify subpopulations at low and high risk, and to integrate this stratification into the healthcare delivery system rather than doing it "post-hoc".

Family relationships can be mined from the electronic medical record (EMR) using emergency contact information, and these relationships can then be used to study the heritability of clinical traits. Observational heritability studies can be employed within the EMR because they collect a broad range of information, including clinical phenotypes. EMR relationships can then be validated against genetic data. At Columbia, 3.2 million relationships were inferred for 600,000 families across three New York OHDSI data sets. Heritability estimates of various diseases such as coronary artery disease and rhinitis were close to estimates from previous studies. This approach can also help identify family members who should be screened based on a relative's diagnosis and can make estimates of heritability based on relationship pairs, a measure that typically has low accuracy.

Discussion

Clinical Decision Support can be used within the EHR to support clinicians who have questions and need to deliver information. However, increasing clinician burden through workflow disruption (e.g., alert fatigue) and more complex health interventions in addition to currently recommended population health interventions are significant barriers to clinician use of PRS. It is necessary to be selective with the advice that is provided to doctors. The method of information delivery matters as well, as physicians are more likely to follow instructions if an alert confirms something they already expect. Because clinicians may already feel "under siege" from multiple alerts, it can be difficult to change behavior and it is sometimes best to limit physician involvement in implementation and instead restructure the delivery system to optimize physician workflow. As data become more available to patients, it may help to create

a more patient-driven healthcare system, especially as patients are currently the only common actors as they move among healthcare providers. Young, healthy patients often do not have a lot of information in the EMR, and this creates a gap in available data. This gap could be closed through the utilization of wearable smart devices such as Apple Watches and Fitbits. Extraction of data from these devices would also reduce the burden on patients of gathering data. However, there must be a balance between patient and physician control in the healthcare system, and the utility of these types of data extraction remains to be established.

There is a need for increased implementation science research and further consideration of physician workflow when implementing new technology. Additionally, integration of algorithms into the EHR poses a significant challenge. While IT professionals and informaticists should be included in research projects to help with this challenge, it is more difficult to incorporate these algorithms in non-research settings. Building an evidence base that proves the clinical validity of these interventions will help to "tip" the field in favor of integrating complex algorithms that support physicians.

Session 3: Choosing the best models—lessons from diverse complex diseases

There is a strong need to accurately stratify risk in breast cancer, where risk-reducing strategies are tailored to the specific risk of the patient. Using data from the Breast Cancer Association Consortium (BCAC), Mavaddat et al. developed a PRS with 77, 313, and 3820 SNPs. Increasing the number of SNPs to 313 significantly improved risk prediction but scaling up to 3820 had only a marginal effect. The 313-SNP score performed significantly better than the others in ER+ tumors and EA patients. Research suggests that increasing the size of a breast cancer GWAS improves risk prediction through PRS. This idea led to the Confluence project, which aims to double the current size of breast cancer GWAS using an international multi-ancestry sample from 5 different consortia in North America, Asia, Africa, and Latin America. There have also been efforts to integrate the 313-SNP PRS with other risk factors for breast cancer, including family history, hormone use, and breast density. These scores perform best when a combination of genetic and classical risk factors is applied in an integrated healthcare system. It is also important to remember that most cases will still occur outside of high-risk groups, so better preventive and early detection strategies at the population level are still needed to improve disease outcomes. Additionally, current PRS are better at predicting low-risk disease like ER+ tumors because they are more common. There is a strong need for better prediction of aggressive disease.

The need for diagnostic biomarkers in psychiatric illnesses such as schizophrenia persists, but PRS provide a foundation for the building of a risk model. Prediction of psychiatric disorders is particularly important because a long "diagnostic odyssey" often occurs after onset and before diagnosis. As mentioned above, the criteria for assessing PRS are varied. The maximum variance explained (R^2) by a predictor can be affected by the number of SNPs, the true variance explained by the SNPs, and the GWAS discovery sample size. These values can be manipulated to increase the R^2 value.

Atrial fibrillation (AF) is the most common arrhythmia and increases risk of heart failure, stroke, death, and dementia. However, treatments are limited, and the best drug is only 40-50% effective. It would

thus be useful to be able to better stratify risk for AF. GWAS data have suggested that lifetime risk of AF is high and strongly modified by polygenic risk. Additionally, strokes caused by AF are more disabling but also more preventable. PRS for AF can be used as biomarkers in select populations to identify patients at risk for cryptogenic stroke. AF genetic risk is specific for certain types of stroke and these patients may benefit from increased screening, which can be improved through wearable technology.

When comparing monogenic and polygenic risk prediction for coronary artery disease (CAD) in the UK Biobank, Khera et al. found that people with PRS in the top 8% of the distribution had a risk equivalent to monogenic risk, and the PRS predicted much higher prevalence of CAD. Identifying patients in the highest percentiles of polygenic risk works as a predictive approach for other common diseases as well, including those without monogenic risk factors. In obesity, patients with PRS in the highest decile were much more likely to develop obesity, while those with PRS in lower deciles seemed to have genetic protection. In Alzheimer's, those with PRS in the highest decile had the greatest decline in cognitive function by age 65. Evaluating PRS risk at birth or early in life could lead to earlier intervention for high risk individuals.

Discussion

Because PRS can be applied to multiple diseases, it is important to consider the clinical contexts in which these scores will be implemented, including the population being tested, the severity of the disease, and the likelihood that patients will want to know their results. There should be clear instructions for follow-up if someone is flagged as high risk, and implementation of these methods should be validated. There is also a lack of clarity on how to proceed with patients who do not fit expected risk profiles. For example, a young and otherwise healthy patient with isolated AF may not be the best target for stroke-prevention interventions given their low absolute risk, despite their increased relative risk of stroke.

When deciding on the number of SNPs to include in a PRS, researchers need to consider the architecture of the disease and the size of the discovery sample. The ideal sample size for a GWAS also depends on disease architecture. Practically, if PRS prediction is not significantly improved by adding SNPs, then fewer SNPs can be used. As the number of risk factors increases, risk can increase in an additive or multiplicative fashion, but it is often difficult to tell the difference when effect sizes are small. At high exposure levels, risk is likely multiplicative, but it appears additive at low levels of exposure.

Patients with a monogenic variant may need a lower burden of polygenic risk to be considered at high risk. However, it is unclear whether the outcomes that have been defined for monogenic disease are the same outcomes that should be measured for polygenic disease risk. Process outcomes and intermediate phenotypes may be helpful measurements for improving the predictive value of PRS.

Finally, there are scientific and ethical concerns about continuing PRS research with limited diversity from non-EA populations. It is likely that causative variants are not equally causative in people of all backgrounds. Further research is required to establish whether and how PRS may operate differently in different population groups.

Debate: "Genomic information is essential to clinical assessment of complex disease risk"

Pro: Stephen Chanock

It is no longer possible to consider risk prediction without genetics. Genetic information has already been shown to be useful for screening, diagnosis, and treatment, but challenges lie in appropriate application. Genetic information can be used as an effective predictor of future risk, especially for stratifying and modifying responses to diseases like breast cancer. Pharmacogenomics is an example of genomics' proven utility in estimating toxicity and efficacy; the field of genomic medicine should not wait 10-15 years to start using genetic data that are already available. Instead, available knowledge, even if incomplete, can be built upon. Family history is also a useful tool, but it does not need to be used to the exclusion of genetic data.

Con: Isaac Kohane

Genetic data alone do not make diagnoses. For example, in the Undiagnosed Diseases Network (UDN) experts make diagnoses by combining multiple types of data from genetic, model organism, clinical, and metabolomic sources. In fact, 30% of patients that join the UDN have already had exome or genome sequencing, but their sequencing results were either not correctly interpreted or were insufficient for diagnosis. Clinicians should also be wary of the "incidentalome," in which repeated testing is likely to uncover multiple unexpected findings and false positives. Cancer panels have a relatively low concordance rate, and even autosomal dominant conditions like hypertrophic cardiomyopathy are frequently misdiagnosed. The obesity epidemic also illustrates that lifestyle factors are more predictive than genetic indicators. Family history is the most important risk predictor of disease, and further research is needed before PRS can be used clinically. Increasing data on diverse populations and environmental risk factors, as well as fixing medical education and workflow around genetics, may help address some of these research gaps.

Discussion

The group discussed which diseases and conditions would be best for early PRS interventions and agreed that breast cancer would be a good disease for testing PRS efficacy. However, while there was interest in implementing PRS to identify newborns at high risk for obesity, it may make more sense to address key environmental factors associated with obesity in preventive efforts. The best genetic testing implementation occurs when both risks and benefits associated with certain variants are understood. In cases of uncertainty, clinicians risk harming patients with reports of pathogenic variants that may be reclassified as benign. It is also critical to consider the lack of diversity in genetics research and the ways that this may leave behind patients with actionable mutations due to their socioeconomic status and disproportionately non-European ancestry. Poverty is generally a stronger risk predictor than genetics.

Session 4: Other 'omic data

Increasing the number of biomarkers in a risk prediction model can increase the model's discrimination. Metabolomic testing can thus be a useful tool. In the FHS, blood profiling was used to find amino acids

that stratify patients into quartiles of risk for diabetes. Non-targeted metabolomic profiling can also be used to assess over 10,000 compounds in the human metabolome to determine the number of peaks associated with a normal sample or a given phenotype. Virtual proteomics, in which genomic predictors are calculated for proteins, closely mirror directly measured proteins and may be used in cases where directly measuring metabolites is not possible. Overall, risk is best predicted by a combination of metabolomic and PRS data.

Transcriptome prediction is useful because it captures a large proportion of the non-coding association signal. It is also a gene-based test with an easy to interpret direction of effect and allows researchers to iterate between humans and model systems where signals often come at the level of the gene. Expression quantitative trait loci, which are loci associated with variation in mRNA expression levels, can be used in conjunction with PheRS to identify candidate disease genes.

DNA methylation (DNAm) of CpG dinucleotide sequences is the most studied epigenetic mark and plays a critical role in the regulation of gene expression. Epigenetic risk prediction is still new, but Epigenome-Wide Association Studies (EWAS) can be used to discover relationships among methylation, disease risk, genetic variation, and environmental exposures. However, EWAS depend heavily on the time of methylation assessment, and results vary across tissues and cells. Additionally, correlation of methylation across CpGs is not well defined. It is also difficult to determine whether methylation increases risk for a condition or vice versa, and heritability of CpGs is variable. EWAS data can be integrated with other ‘omic data like gene expression, and by applying two-step Mendelian randomization, researchers can establish the causal relationship among exposures, DNA methylation, and outcome. In blood pressure, for example, methylation explains more variance than genetic loci. DNAm can also be used as a proxy of self-reported phenotypes. It has near-perfect discriminatory power for predicting the risk of someone smoking. It is also a good predictor of mortality and age, though epigenetic age sometimes deviates from actual age. DNAm may be helpful for estimating someone’s rate of aging, though it depends on the population and the tissue from which the score was derived.

Environmental exposures are important to consider when estimating risk. In practice, simple linear models where genetic and environmental factors (G and E, respectively) are summed are used most often. This may be because non-linear models tend to require very large training sets, and linear models can pick up most non-linear effects. However, more research is needed to determine whether linear models are good at predicting risk in the tails of the distribution. Analysis of environmental exposures can be useful to spread out risk distributions and to identify subgroups in which G is actionable.

Discussion

Multi-omic measures, which capture genes, the environment, and the interaction between the two may be used to estimate gene by environment interactions to develop proxies for environments that are unknown and difficult to measure. Metabolome, exposome, and intermediate exposure data can be useful when built into prediction models. However, one challenge with ‘omic data is the timing of measurements. In the case of DNA methylation, measurements depend on when they are taken during a patient’s lifetime, so longitudinal measurements of methylation are ideal. For many metabolites and proteins, timing is not as important, but some analytes need to be measured at the same time every

day. It is also important to consider that predictors of disease may not necessarily be the best predictors of response to therapy.

The reasons for DNAm's high discrimination for smoking are unknown. Smoking has a robust and persistent signature on the genome, even if someone has not smoked in years. Smoking and age are important risk factors for many other conditions. It may thus be useful to look for people who have high methylation age or DNAm signatures like those of smokers, as they may be at higher risk for certain conditions. Additionally, adult onset methylation of many CpGs is heritable, and may explain more variability than SNPs.

Panel—Do we need a clinical trial of genomic risk prediction? If so, what should it test, in whom, and with what outcomes? What do we need to know before planning such a trial?

The five panelists first presented their opening statements:

- Multiple clinical trials will be needed depending on the intended use of genomic risk prediction.
- Clinical trials will need to consider who is studied and how the results will translate across diverse populations.
- Clinical trials should emphasize 3 main outcomes: implementation outcomes, service outcomes, and patient outcomes, including clinical and patient-reported outcomes.
- Clinical trials are necessary to build evidence for PRS use in the clinic. The most important considerations are that the correct population is targeted, the results are actionable, and that ethical issues are accounted for.
- The field should consider health outcomes, economic outcomes, and physician uptake of PRS.

Discussion

It is not clear whether a separate trial will be necessary for each disease, or if a general validation of genomic prediction approaches will be sufficient. It may be useful to focus trials on a few major diseases with similar parameters to compare outcomes. One option for identifying diseases is to look at evidence-based guidelines for screening and treatments. Researchers could start with screenings that are already in practice, then evaluate whether PRS add value to these screenings. These can also be stratified by tiers of severity and actionability, where tier 1 diseases are investigated while tier 2 and 3 diseases are explored later. The diseases of focus should have potential interventions. Trials can be embedded into existing research, such as All of Us, focusing on patients with intermediate risk who may benefit most from a shift in their risk prediction.

The potential cost of these trials is a huge obstacle. It is possible that some private sector entities may be interested in helping to fund clinical trials for PRS. However, many companies would only be interested in doing trials for which the projected return on investment is high. A learning healthcare system that does pragmatic clinical trials and electronically based randomization may also help to decrease the price of clinical trials. Another challenge is the lack of hard outcomes from these studies. In

many cases, it may be years before a hard outcome is observed, but hard outcomes are the most persuasive. This does not fit well with a 4-year grant mechanism. One potential solution is to engage large companies with long employment histories to test trials in their employee health systems or health maintenance organizations. NHGRI held a meeting in March 2019 with employers to discuss this topic, recognizing that institutions may have their own vested interests in providing testing to their employees.

There is a limited window of opportunity to generate evidence in clinical trials. Currently, it is not standard of care to return PRS results, so researchers can randomize patients to receive or not receive PRS results. In the future, clinicians may no longer be comfortable with randomization. It is also possible that PRS will be pulled into practice due to their perceived utility before clinical trials can be performed. If trials are necessary, then potential partnerships should be considered to lower the price. For example, drug companies might be willing to pay for certain types of studies that identify new users for their drug.

In summary, the group felt that there do need to be clinical trials of PRS, either independently or as supplements to ongoing studies. Engaging employers and healthcare systems could be important in the creation and funding of these trials. Potential trials will need to measure hard outcomes, as well as implementation, psychological, reproductive, and physician behavior outcomes. The number of participants and their backgrounds will also be essential factors to consider. There is the potential to worsen outcomes or widen health disparities in underrepresented groups if they are left behind.

Research directions in genomic risk prediction

Several prioritization considerations need to be made to determine the role of genomics in risk prediction. One of these prioritizations is the type of disease on which the field should focus its efforts. Researchers need to decide on the complexity of the diseases to be studied, and whether there should be an emphasis on diseases with large environmental components or other confounders. Data have shown that there is value in implementing PRS in monogenic diseases where polygenic components can further stratify disease risk. The availability of hard endpoints might lead researchers to prioritize one study over another, as well as the amenability of the disease to implementation in real-world healthcare systems. Additionally, the best way to combine PRS with other factors is still unknown. Other ‘omics likely have a place in the future of risk prediction, alongside environmental exposures and family history.

Researchers will need to clarify how they characterize risk scores and whether they are focused on prevention or treatment once a disease is manifest. Clinicians will likely be more willing to use scores to identify individuals at higher risk rather than to reduce frequency of screening in persons at lower risk. Genomic risk prediction will likely be most valuable in situations where therapy is dangerous or expensive, or where monitoring is frequent, such as mammograms for breast cancer screening. It also could help in the case of neuropsychiatric disorders which often entail a long diagnostic odyssey and are a burden on the healthcare system and the patient’s quality of life. However, more research must be done on these illnesses to apply PRS effectively.

The amount and availability of data in existing cohorts should also be considered. Researchers may want to prioritize investigating scores where clinical validity data have been more established. Effect size also needs to be relevantly defined, as it depends on the outcome being measured and can affect the size of the population needed to have a powered study. However, often the data that PRS are being built upon were collected decades ago. It is possible that these data have become less relevant as time and health practices have changed.

There is disproportionate value that comes from studying under-represented racial and ethnic minorities, especially populations of recent African ancestry that have lots of genetic variation. This helps to identify both causal variants and variants that contribute to disease risk. Though it would be useful to create a risk score for each ancestry group, race is often poorly defined and may need to be built into the risk score methodology to ensure the best score for each individual.

One of the more difficult aspects of getting clinicians to adopt PRS is that the data will continue to change and scores will shift as variant discovery continues. There also needs to be further research into why patients sometimes hesitate to receive PRS results. Broader use of decliner surveys could be effective. Finally, research should investigate the best ways to present this topic to patients. If patients are told that an actionable finding might be identified, they might be more likely to engage.

Summary and next steps (see Executive Summary for more specifics)

There are many opportunities for genomics to add value to current risk prediction methods, but significant research gaps must be addressed before PRS are broadly implemented into clinical care. It will be critical to continue to find risk factors and biomarkers that increase risk stratification, as well as finding ways to integrate genomic risk data with clinical factors, environmental exposures, and other 'omic data. Increasing the size of current GWAS studies and focusing on risk variants in diverse populations may help address some of these data gaps. Additionally, PRS need to be disease subtype- and ancestry-specific, and additional validation is required to demonstrate their utility in non-EA populations. Finally, there is a strong need for implementation science that investigates the best ways to integrate PRS into physician workflow and existing risk calculators. In addition to this meeting summary, GMWG will consider creating a white paper to summarize recommendations and future directions for the field of PRS and genomic risk prediction.