



National Human Genome  
Research Institute

## Multi-omics in Health and Disease Current Applications, Challenges and Future Directions

### Workshop Summary Report

# Multi-omics in Health and Disease

Current Applications, Challenges  
and Future Directions

**June 17-18, 2021**



The **Forefront**  
of **Genomics**



## Table of Contents

|  |           |
|--|-----------|
| <b>Welcome and Introductions</b> .....   | <b>1</b>  |
| <b>Presentation of the 2020 NHGRI Strategic Vision</b> .....                                       | <b>2</b>  |
| <b>Purpose of the Workshop</b> .....   | <b>2</b>  |
| <b>Summary of NIH Investments in the Field of Multi-omics</b> .....                                | <b>2</b>  |
| <b>Session 1: Setting the Stage - Application of Multi-omics to Study Health and Disease</b> ..... | <b>3</b>  |
| <b>Session 2: Technology, Data Integration and Study Design</b> .....                              | <b>4</b>  |
| <b>Session 2: Guided Discussion - Technology, Data Integration and Study Design</b> .....          | <b>5</b>  |
| <b>Session 3: Application of Multi-omics to Observational Studies</b> .....                        | <b>6</b>  |
| <b>Session 3: Guided Discussion - Application of Multi-omics to Observational Studies</b> .....    | <b>8</b>  |
| <b>Session 4: Future Clinical Implementation: Roadblocks and Opportunities</b> .....               | <b>10</b> |
| <b>Session 5: Recommendations to NHGRI</b> .....   | <b>11</b> |
| <b>Summary and Next Steps</b> .....  | <b>12</b> |

Rapporteurs: Marie-Luise Brennan,  
Laurie Findley, Joannella Morales

**Multi-omics in Health and Disease:  
Current Applications, Challenges and Future Directions  
June 17-18, 2021  
Virtual Meeting Summary**

**Welcome and Introductions**

On June 17 and 18, 2021, the NHGRI convened leaders in multi-omics technologies (e.g., genomics, epigenomics, transcriptomics, proteomics, single cell ‘omics and data integration) and observational studies to provide guidance to the National Human Genome Research Institute (NHGRI) in developing a research agenda based on multi-omics.

The goal of the meeting, titled “[Multi-omics in Health and Disease: Current Applications, Challenges and Future Directions](#)” was to understand the state of the field and to gather recommendations on a research strategy using multi-omics that is in line with the objectives in the NHGRI 2020 Strategic Vision.

The meeting had the following objectives:

- To gain insight on how multi-omics data can improve our understanding of health and disease
- To identify study design, data integration and technological gaps and challenges to the use of multi-omics technology and its application to observational studies
- To consider steps required for future clinical implementation
- To define opportunities to overcome these challenges that are relevant to NHGRI’s mission

June 17<sup>th</sup> (Day 1 of the workshop) coincided with President Biden signing the Juneteenth National Independence Day Act, establishing Juneteenth as a federal holiday. We recognize this as a milestone in our country’s history. Discussions were held within NHGRI on whether to reschedule Day 2 of the workshop given it would occur the following afternoon during the new federal holiday. These discussions were not taken lightly and ultimately a decision was made to continue the workshop. In recognizing and reflecting on Juneteenth, we also appreciate the multi-omics concept has implications for diversity, equity, and inclusion on multiple levels including for participant representation, access to biomedical research studies, and utility in clinical interpretation, and clinical care recommendations.

## **Presentation of the 2020 NHGRI Strategic Vision**

NHGRI published “[Strategic vision for improving human health at The Forefront of Genomics](#)” in Nature in October of last year. Dr. Eric Green, Director of NHGRI, noted that genomics is widely disseminated across the entire biomedical research enterprise and thus, the Strategic Vision focuses on where the NHGRI wants to lead and facilitate research at “The Forefront of Genomics.” This Strategic Vision describes the most compelling research priorities and opportunities in human genomics for the coming decade. It also recognizes responsibilities that come with this leadership role and stewardship. The vision identifies four focus areas: 1) guiding principles and values for human genomics, 2) sustaining and improving a robust foundation for genomics research, 3) breaking down barriers that impede progress in genomics, and 4) compelling genomics research projects in biomedicine. The Strategic Vision culminates with ten bold predictions for human genomics by 2030 that are intended to promote discussion and provide inspiration about the future of genomics research. Dr. Green noted the value of workshops such as this in collecting input that is needed for a robust discussion on deciding the right time to “germinate” and fund concepts such as multi-omics research in health and disease. Furthermore, he asked the meeting attendees to consider whether the timing is right for NHGRI to bring a program forward in this area and, if so, to provide recommendations on the appropriate focus for the Institute.

## **Purpose of the Workshop**

The 2020 Strategic Vision recognizes the potential of multi-omics applications and articulates an ambitious aim to extend multi-omic studies of human disease and health into clinical settings. The strategic vision expands on areas of focus that can lead towards achieving this goal in both research and clinical areas. Research goals include: 1) to extend genomics research beyond DNA sequence; 2) to include other multi-omics data and combine those data with clinical variables and outcomes (which will require more work at the tissue and cell-specific level, new tools and technologies, data integration and sample diversification); 3) to increase understanding of biological processes and disease onset and progression, and 4) to facilitate drug discovery efforts. The goal of clinical implementation will involve: 1) integration of multi-omics data with electronic health records and clinical decision support tools, and 2) optimally, facilitating a shift in the clinical use of biomarker testing from focusing primarily on diagnosing and treating disease to instead focusing on comprehensive health and wellness.

## **Summary of NIH Investments in the Field of Multi-omics**

Based on a recent portfolio analysis, the number of awarded NIH grant applications including the concept of multi-omics (as assessed by relevant keyword searches) has increased almost 8-fold over the past five years. A review of awarded grants and the literature on multi-omics reveal major themes of 1) deep phenotyping to generate disease signatures, 2) use of new technology (e.g., RNA-seq, single cell analysis) to generate multi-omics data, 3) data integration and 4)

longitudinal ‘omics. These grant trends are driven by the significant investments across the NIH by the NIA (IALSA); NHLBI (TOPMed); NCI, NIAID, NIMH (Brain Initiative), NIDDK and Common Fund (HuBMAP). While most of these initiatives focus on specific diseases or conditions, the NHGRI is disease agnostic and instead has historically focused on foundational issues and generalizable approaches. The existing NHGRI portfolio on multi-omics has been relatively limited and focused on approaches for data integration funded via small business innovation research and investigator initiated grants.

Assessing grants and literature reveals gaps in a) technology, both experimental and computational, and with data integration, which is challenging given the multi-dimensional and often longitudinal nature of multi-omics, b) study design considerations such as sample size, data source, diversity, and data harmonization standards, and c) the application of multi-omics in the clinical context.

Discussion centered around the approach to multi-omics funding by NHGRI, and the payoff of past biomarker investments by the NIH. These topics were directly aligned with the purpose (e.g. solicit input on best approaches and existing trends/gaps) and design (e.g. session topics, invited speakers, guided discussion questions) of the workshop. Emphasis was placed on input being needed to move the answers to these questions forward. Significant funding on multi-omics across the NIH was recognized as was the realization that despite the significant investment in and added value of biomarkers, there has not been comparable progress in implementation into clinical practice.

### **Session 1: Setting the Stage - Application of Multi-omics to Study Health and Disease**

This session consisted of talks on the application of multi-omics to study health and disease followed by discussions.

Nancy Cox hypothesized how the explosion of medical data from electronic medical records and wearable devices will make hospitals de facto phenome centers and that if, as a community, we wish to utilize this rich repository of data then we need to build an understanding of genetics within the context of the health care system. Addressing structural elements (governance, education, security) that support responsible data management and utilization will be critical in manifesting this vision. Education was highlighted as a system-wide need for: 1) hospital leaders to understand the benefits of multi-omics, 2) insurance companies on genetic testing utility, 3) healthcare providers on complex diagnoses, rare variants, and atypical patients, and 4) patient control of their personal health information. An example of where this system-wide education could have impact is with rare genetic variation and clarifying the contribution of rare pathogenic variants to atypical laboratory results and/or disease presentations. With the amount of information in electronic health records (EHR) and minimal funding, this aim as well as an overall understanding of the test characteristics of patients for whom genetic testing is

recommended is feasible. Biomarkers provide the opportunity to simultaneously consider the genetic and non-genetic components to biomarkers in their response to homeostasis and disease. Of importance, samples utilized in these studies should reflect population diversity so laboratory reference ranges and subsequent clinical interpretation and decision making are appropriate.

Mike Snyder presented on the power of deep ‘omics profiling and longitudinal measurements in monitoring a person’s health. ‘Omes (and shifts in ‘omes) are relatively stable within an individual and it is easy to identify shifts from a personal baseline when examining ‘omics data. Power calculations for sample size do not apply in the traditional sense as longitudinal data per person, as opposed to a population reference, are the most informative in showing biological shifts. Multiple examples including cardiovascular, metabolic and endocrine cases, were provided showing longitudinal monitoring leading to disease identification in pre-symptomatic persons that were not otherwise clinically suspected or identified. Of note, multiple types of ‘omes were required. Another example was provided of ‘omics analyses separating out a subtype of diabetes and selecting therapy when diet, exercise and first line medication was not therapeutically effective. At-risk groups could be good targets for multi-omics monitoring such as those for which a watchful-waiting approach is taken. However, moving to the entire population instead of just at-risk groups was articulated as important in shifting to a wellness and prevention approach. Lastly, EHR and wearable data are incredibly useful, however, the state of reliable phenotyping in the computer records and moving wearables into the clinic are challenging.

Discussion on the application of multi-omics in disease centered on the abundance of complex data available, including wearables; what ‘omes can actually capture (aging, stress, seasonality); study design including getting into the right clinical spaces and the pro/con of whether studies should be targeted to at-risk groups and the hope that medical care will include more of an emphasis on wellness and prevention.

## **Session 2: Technology, Data Integration and Study Design**

This session on technology, data integration and study design consisted of three talks and discussions.

Xihong Lin spoke on experimental and computational technology used in big data initiatives such as GSP, TOPMed, IGVF, ENCODE and GTEx. The integration of whole genome data with multi-omics to annotate functional variants is a powerful tool in studying the mechanism of variant-to-function-to-disease and discovering rare variant associations. In doing so, cell type specificity, sample diversity and collaboration between experimental, genetic, epidemiologic and clinical consortia are critical. Infrastructure includes cloud-based data access, analytic platforms, data sharing, and harmonization. Scalable analytic tools and resources and cost issues are also important.

Tuuli Lappalainen spoke on study design and sample collection. Current key multi-omics projects differ in their study designs with some (GTEx, TOPMed, Plasma Proteome studies) focusing on population variance and clinical phenotypes with others (Human Cell Atlas, Human Protein Atlas, ENCODE) focusing on multiple tissues, and cell type resolution. The latter group has deeper 'omics profiling but the entire spectrum is needed for precision medicine. Cell type composition is a key driver of molecular variation that correlates with other factors and gene-by-environment eQTLs are often due to variation in cell type. Blood is not a good proxy and screening a more diverse set of tissues is still a practical challenge. Understanding cost and target application/value of analyses is needed to intelligently design studies that will provide the integration and information needed.

Marjorie Brand spoke on data integration using hematopoietic stem cell differentiation as a model system. In general, a 30% correlation between gene and protein level is observed but no one type of 'ome is a surrogate for all processes that are occurring. Multi-omics studies are complex and need to account for cellular changes occurring in different directions, at different speeds and over time.

Discussion focused on appreciating several needs. There is no single study design or surrogate tissue/'ome that will address all the needs of the field. There is need for both consortium-level and independent-investigation level endeavors. There is also a need for both disease-enriched cohorts and population level data as one study design does not replace the other. Lastly, deconvolution and single cell analysis, while helpful, do not alone provide the answer to the challenge of being able to collect relevant cell types and tissues (other than blood) at scale.

## **Session 2: Guided Discussion - Technology, Data Integration and Study Design**

This session included three presentations followed by a guided discussion focused on addressing the questions: 1) Where do we want to be? and 2) What are the barriers and opportunities?

Sarah Teichmann articulated the future of medical functional genomics involving single cell multi-omics with longitudinal electronic health records, longitudinal cohorts with whole genome scanning, and case-control studies. The Human Cell Atlas and COVID-19 Cell Atlas were provided as examples of each of these concepts that are currently in place and generating informative data.

Lana Garmire spoke on integration of multi-omics data. Most data integration methods are not supervised by phenotypes but are instead built from combinations of genetic data. DeepProg is an example of a program that uses deep learning to integrate phenotypes such as survival in predicting cancer prognosis. Issues that remain to be addressed in the field include complex confounders (cell type; tissue type specificity); study design to address confounders and lack of

phenotyping, too many methods with not enough active benchmarking (for example RNA-seq and DNA methylation give different proportions of cells) and integration with other data modalities

Neil Hanchard spoke on ‘omics as two-dimensional entities that give rise to a three-dimensional perspective that will ideally provide a multi-omics paradigm for health and disease. This clinical disease research perspective includes the need to address four components: cohort data repositories of differing disease states, ages and ancestries; technology; data integration; and study design. Data integration, and sample considerations including understanding cost, timepoint and maximizing each aliquot were raised.

The three discussants came from different perspectives but all homed in on similar needs. Additional emphasis on benchmarking is needed, particularly for combined single cell multi-omics which have not really been established for double, triple or more data types from single cells. Methods for dealing with gaps and missing data (e.g., individuals missing data for one time point) are needed. While there are some imputation methods that are being used, a better understanding of how to address these issues should be included in study design. Understanding what information can be obtained from current datasets, as opposed to new studies, is important in maximizing resources. Continuing to add new collections is not necessarily the solution. Education is also needed in thinking about international studies regarding utility, privacy, and ownership.

### **Session 3: Application of Multi-omics to Observational Studies**

Day 2 started with four talks on the application of multi-omics to observational studies.

Nathan Price spoke on polygenic risk scores (PRSs) and how they interface with multi-omics. The Pioneer Project was a longitudinal multi-omics cohort collection focused on wellness. Biomarker reference ranges and PRSs were developed based on review of existing scientific literature. Genome-wide association studies for 54 diseases and complex traits coupled with multi-omics profiling were performed and PRSs were associated with 766 detectable alterations in proteomic, metabolomic, and standard clinical laboratory measurements from blood plasma across several thousand mostly healthy individuals. Previously known biomarker-disease relationships were confirmed. Of the laboratory variables included, clinical lab values were most strongly correlated with polygenic risk scores. Searching for biomarker-disease associations amenable to therapeutic targeting yielded previously well characterized (cardiovascular disease and phosphatidylcholine; PCSK9 inhibition therapeutic target) and unexpected (amyotrophic lateral sclerosis and omega fatty acids) associations of biomarkers with disease state. Analytes altered in high-genetic-risk individuals also showed concordant changes in disease cases. Thus, a person’s individual genetic profile and dynamic measures may provide a prioritization of health related choices suggesting it will be possible to map out the most genetically at risk people for



disease. This also emphasizes that genetics are not destiny, but the outcomes of lifestyle interventions are quantitatively affected by them. In designing health strategies for people, it is therefore important to define the areas where the most progress is likely - working with their genes rather than against them.

Kari Nadeau spoke on how environmental exposures interact with immune and genetic factors and the importance of measuring and understanding this relationship. Chronic exposures, repeat exposures, and exposures over a lifetime each have effects. In the current context of climate change, increasing air pollution and allergens, extreme heat, environmental degradation, living conditions and other social issues all come into play. These exposures are particularly damaging to children during critical development periods. There is significant air pollution in the Central Valley in California and pollution exposures were measured on an individual basis. Immune marker and cell type fingerprints with different pollution exposures as well as altered methylation patterns and alterations in gene pathways affecting inflammation and asthma susceptibility were observed. Twin studies were provided as a second example of gene/environment interactions with monozygotic twins having identical DNA and thus providing a means to separate out nongenetic from genetic factors. Analysis of metabolomics showed age-related variability in metabolic profiles in twins as they age with twins becoming more dissimilar over time. Lastly, next steps were presented including exposure analysis as a feasible target of study, and needs that include better technologies, composite exposures, and longitudinal measures.

Corrine Engelman spoke about observations, opportunities, and barriers in longitudinal cohort analyses using Alzheimer's disease as an example. Given this condition has pre-symptomatic, mild cognitive impairment, and Alzheimer's disease stages, clinical and imaging measurements can be combined with genomics and longitudinal multi-omics (such as metabolomics and proteomics in blood and cerebrospinal fluid) to generate networks for understanding disease. When examining gene-metabolite relationships, they observed no genes associated with Alzheimer's disease risk factors directly but instead indirectly (genes associated with plasma metabolites that in turn were associated with risk factors). This finding emphasized a key barrier - there is a dearth of studies with longitudinal 'omics over many timepoints to tease apart real observations from random variation. To date, most studies use a case/control study design, making it challenging to conclude whether the changes in the 'omes are the cause or result of the disease process. Opportunities include a) establishing the timing and trajectory of pathologic changes in preclinical individuals, b) using heritability estimates in large studies to determine whether the 'ome is influenced by genetics versus behavior/environment; c) using genomic data and Mendelian randomization to establish causality (i.e., that 'omics data are predictors of the outcome versus influenced by the outcome) for 'omics with moderate to high heritability; and d) exploring whether 'omes with lower heritability may be mediators of the relationship between behavioral and environmental factors?

Tes Mersha spoke on multi-omics synergism (i.e., the combined effect is greater than the sum of separate effects) and the importance of data integration. Reducing patient heterogeneity and providing improved risk prediction will involve a shift from clinical phenotype to endotype, computational endotyping, and multi-omics-based risk prediction and patient classification. Effective multi-omics based diagnosis requires accurate correlation of ‘omes with detailed phenotypic information. Asthma was provided as an example of a heterogeneous disease with multiple endotypes where integration of clinical, ancestry, and multi-tissue transcriptomics analysis provided insight. Multi-omics based risk prediction (polygenic risk score, gene expression risk score, methylation risk score) will be important and while there are high-throughput multi-omics technologies, there are not high-throughput phenotyping. Deep phenotyping will be limited by context (i.e., longitudinal; spatial; exposome; gene-level vs. pathway level); access to tissue and alternative tissue surrogates; comparable clinical informatics (standardized, harmonized); and machine learning (e.g. deep learning) approaches. Effective multi-omics based diagnosis requires accurate correlation with detailed phenotypic information.

Discussion included the opportunities to use deep multi-omics to learn consequences of exposures that would otherwise be hard to gauge, contributing to a better understanding of magnitude and impact of gene-environment interactions. Environmental health translation studies are important and feasible, providing an opportunity to decrease health disparities. More work is needed to deploy ‘omes in larger populations to validate and identify best standards. Tissue types, timepoints, study design, and technological and phenotyping needs remain areas of need and agree with earlier discussions.

### **Session 3: Guided Discussion - Application of Multi-omics to Observational Studies**

This session included four presentations, followed by a moderated discussion. Each presenter focused on addressing the questions: 1) Where do we want to be? What is our aspirational goal? and 2) What are the barriers and opportunities? After the speakers completed their presentations, a discussion open to all participants was led by the moderator.

Myriam Fornage emphasized the need for a comprehensive set of large-scale ‘omics data that would include a large number of people, with multiple measures taken at multiple time points at the single cell level and in various tissues. She also expressed the need for a comprehensive set of standardized and validated biomarkers that could be integrated into clinical settings. Finally, Dr. Fornage highlighted the need for advanced computational methods, such as machine and deep learning, to model risk prediction, diagnosis and therapeutic response in diverse populations, noting that their application and adoption in clinical settings with links to EHR is a key goal. The infrastructure and multidisciplinary expertise required to support these complex methods will also be important. Challenges related to the harmonization and integration of

heterogeneous and high-dimensional data were highlighted, including the importance of identifying unwanted sources of variation and biases and the need for interoperable data resources and ontologies. Dr. Fornage emphasized that large consortia and collaborative efforts will be critical to the application of multi-omics. These groups will need to share data, workflows and infrastructure in order to curate, harmonize, and integrate multi-omics data. The value of multi-omics increases when integrated with environmental, social and lifestyle exposures over the lifespan of an individual. Longitudinal epidemiological cohorts with exposure-driven data collection over multiple decades and prior to overt disease will be invaluable. Combining this data with disease-specific or tissue-specific data collection will offer additional insights.

Inspired by NHGRI's Bold predictions, Adam Butterworth suggested that by 2030 we should aim to have formed multiple large consortia of diverse patient and population cohorts with coordinated serial measurements of several multi-omics layers from multiple tissues and cells anchored in genomic data and linked with EHRs. Also, he noted that a goal should be to have widely accessible (but safely stored and managed) data and novel methods to better explore the complex networks from genomic variation through multi-omics to health and disease. This will require novel ways of thinking and analyzing data to maximize discovery. Dr. Butterworth noted as a barrier the fact that costs are still high for some types of 'omes and suggested that to drive technology and reduce costs it will be essential to partner with industry. He also emphasized the necessity of moving beyond the current approach of mainly assessing blood samples to one that includes more tissues, samples, time points, types of 'omes and population backgrounds (diversity).

Greg Gibson noted that biomedical research needs to increase its focus on predicting therapeutic outcomes, including aiming to understand how disease progresses and how patients respond to treatment. Dr. Gibson also emphasized the need to support consortium-based multi-omic data acquisition, especially longitudinal, multi-tissue and multi-cell data. He noted that multi-omics integration must be in the context of the patient's environment, including lifestyle choices, socio-economic context, exposures and ancestral background. Integration of 'omes improves prediction and precision because it is more pathology-proximal, especially compared to genetics alone. Single cell multi-omics can help identify personalized pathology and can provide insight into a patient's response. Single cell profiling should be an area of focus as it provides the required level of resolution. The end goal should be better forecasting a patient's outcome and improving a model's predictive value.

Using Alzheimer's Disease as an example, Alison Goate emphasized the need for increased access to relevant tissues. It is clear that different cell types have different characteristics. Therefore, single cell data sets are also important, as are methods to harmonize and integrate data across cells and 'omics layers. Dr. Goate also highlighted the need for more biomarkers for

diagnostic, prognostic and therapeutic purposes and larger SNP-array and sequencing datasets from diverse populations. She noted that it will be important to both collect new data (from blood, specific tissues and single cells) from existing longitudinal studies that have good phenotypic and genetic data as well as from new observational cohorts. These new cohorts should include families with diseased individuals but also healthy populations to define at-risk groups and collect large scale ‘omics data longitudinally. It will also be important to bank cell lines from diseased individuals for future functional studies.

The open discussion started with agreement about the importance of going beyond blood to study tissues and single cells. However, it was noted that it will be important to focus studies on disease-causing cells as opposed to symptom-exhibiting cells. The group agreed that this is why efforts to map comprehensive single cell data (such as HCA and HubMap) are critical. It was noted that for this data to be maximally useful, it will need to take into account environmental effects.

The value and use of smartwatches and other wearable tools to capture data was discussed. There was general agreement that wearable data integrated with multi-omics and phenotypic data would be a powerful way to understand health and disease. It was noted that wearables were especially helpful for patients that are unable to manually input data or describe events and environmental cues (for example, Alzheimer’s disease patients or other patients with memory loss).

Finally, there was a discussion about the importance of optimal study design. Since multi-omics can be used to both understand disease mechanisms as well as to identify biomarkers, it will be critical to carefully consider methods (technological and computational), cohorts and populations (existing cohorts vs. newly established cohorts from diverse populations) and source (blood vs. tissue vs. cells) prior to commencing studies.

#### **Session 4: Future Clinical Implementation: Roadblocks and Opportunities**

Session 4 began with presentations by Judy Cho and David Craig. The presentations were followed by a discussion.

The focus of Dr. Cho’s presentation was on genetics, multi-omics, and therapeutic targeting in Crohn’s Disease. She presented research on how loss of function alleles in *NOD2* result in a higher risk for Crohn’s Disease. Using single cell transcriptomics, her lab found that *NOD2* is expressed in activated fibroblasts and macrophages. Pre- and post-treatment cohort data was utilized. The lab is now looking into new therapeutic targets. Dr. Cho concluded her presentation by discussing the need for multi-omics integration and the scaling of novel treatments.

The focus of Dr. Craig's presentation was on lessons learned regarding moving genomics to the clinic. He noted that EHR integration and access are critical. Also, data sharing requirements are needed, though the importance of respecting sovereignty and privacy of data, particularly for underrepresented populations, should also be kept in mind. Dr. Craig also discussed the usefulness of public-private partnerships for data sharing, the federation of EHRs, projects such as 1000 Genomes and their impact on standards and open data, and the prioritization of diversity. When discussing the 1000 Genomes project, he mentioned that open data is key for the training of students.

The first topic of discussion was on federated data sharing and the integration of EHR data. The need for access to data for the validation of prediction models was emphasized. It was also noted that it would be useful to have a federated model that allows for the querying of variants across health systems, though questions arose about the feasibility of this. Steps to accomplish this include the creation of straightforward “variant by diagnosis” pre-specified tables or “look-ups,” discussions on governance, and buy-in from various stakeholders.

Other topics of discussion included the need for criteria regarding clinical utility of tests for a given disease and the need to establish reference ranges at the population level and individual level. Multi-omics cohort data could be useful for this, though finding standards for clinical utility that are universally useful will be difficult (more so than that for risk stratification etc.). Furthermore, there is a bottleneck for turning multi-omics measures into clinical tests. The development and validation of multi-analyte tests using absolute quantification will be useful. Additionally, while it will be even more difficult to validate RNA-seq tests, there should be efforts to develop and validate the multi-analyte tests and tests for richer ‘omics data (like RNA-seq data). Clinical utility of the assays will also need to be shown.

Finally, there was a discussion about equity and the need for the engagement of groups outside of the major health centers, such as county hospitals and native populations. There will need to be some flexibility here, and these groups must be allowed to have some sovereignty over their data. Additionally, there is a trust issue affecting recruitment. Finally, in regard to increasing the diversity of GWAS data, there is a need to work to increase diversity with regard to functional interpretation and reference data too.

### **Session 5: Recommendations to NHGRI**

The concluding session was a brainstorming session focused on generating a list of recommendations. A consolidated list representing discussions from throughout the meeting and from the brainstorming session was compiled by NHGRI staff and reviewed by the workshop planning committee for completeness.

## Summary and Next Steps

The [Executive Summary](#) details the lessons learned and recommendations from the meeting. All of the presentations and video recordings from the meeting can be accessed on the [Multi-omics in Health and Disease: Current Applications, Challenges and Future Directions](#) page of the NHGRI website. In addition to this Meeting Summary, co-chairs Howard Chang and Judy Cho will be working to develop a manuscript for publication that is based on the outcomes from this meeting. Speakers and moderators are encouraged to contribute as co-authors.