**1 Quarter of FY 23 Summary of eMERGE/PRIMED Joint Meeting: February 2nd, 2023**

Zoom & In-Person

|  |  |
| --- | --- |
| **Time** | **Event** |
| 8:30-8:50 AM | [NHGRI Program Introduction | Iman Martin (NHGRI, PRIMED) & Robb Rowley (NHGRI, eMERGE)](#tv79skyqso5j) |
| 8:50-9:20 AM | [Network Overview | Eimear Kenny (PRIMED; Mt. Sinai) & Rex Chisholm (eMERGE; Northwestern)](#5iurummsrhhg) |
| 9:20-10:50 AM | [***Panel****:* ***Conceptualizing and assessing PRS performance in diverse groups.* Moderators | Anna Lewis (MGB) & Josep Mercader (Broad)**](#c3cyov2c9fz2)   * [Continuous and robust conceptualizations of genetic ancestry in PRS development | Eimear Kenny (Mt. Sinai)](#p3tjapv8sy0w) * [Population descriptors, population-specific methods and PRS development | Gen Wojcik (JHU)](#a606etu88sph) * [Approaches to improve the performance of type 2 diabetes polygenic scores in diverse and admixed populations by increasing diversity, sample size of genetic discovery, and coverage of linkage disequilibrium reference panels | Alicia Huerta (Broad)](#sy9igcz2vy28) * [An approach to calibration of population variance in implementation of PRS for diverse populations | Niall Lennon (Broad)](#34y3mshhom7) * [Discussion & Q/A](#53ue9wcxf78b) |
| 11:05-12:20 PM | ***Workgroup breakout session***   * [Phenotyping fidelity & impact](#ahm3azpnjf4h) * [PRS Evaluation](#6iapbofiv7ed) * [Cloud resources and data sharing](#pybxcjrop2ta) * [Social, ethical, and policy implications](#ad7ianz9l0p3) |
| 1:20-2:35 PM | ***Panel***: ***Translational aspects of PRS.* Moderators | Josh Peterson (VUMC) & John Witte (Stanford)**   * [Social and ethical considerations for PRS development and application | Megan Shuey (VUMC)](#1df60aqaajz) * [Integrating and contextualizing genomic risk with other risk factors | Iftikhar Kullo (Mayo)](#q4cs60moyjup) * [Threshold for evidence for clinical implementation | Beth Karlson (MGB)](#ocwczu5bos2g) * [Discussion & Q/A](#v0shs5wz17iy) |
| 2:35-2:55 PM | [Joint projects from breakout sessions. Moderators | Clement Adebamowo & Rex Chisholm](#fhly7z54hqbo) |
| 2:55-3:00 PM | [Closing remarks | Clement Adebamowo (PRIMED; U of Maryland) & Rex Chisholm (Northwestern)](#jtlwa4jydsxg) |

\*\*Action Items are called out in session specific notes

|  |
| --- |
| **Other Links** |
| [Recordings](#izk9dbq8mg6) |
| [Background Packet](https://drive.google.com/file/d/13siphWgbTQ0FXL0q6UNFudgNDH1Xtc7F/view) |

1. **NHGRI program introduction | Iman Martin (NHGRI, PRIMED) & Robb Rowley (NHGRI, eMERGE) |** 
   1. This is an opportunity for cross consortium collaboration, as both consortiums focus on polygenic risk scores (PRS). There are guests from eMERGE, PRIMED, ClinGen, and the NHGRI. The joint day focuses on performance of PRS in diverse groups and translational aspects of PRS development.
   2. Attendees filled out an introspection exercise, highlighting the most important themes in PRS research. It is important to ensure PRSes can be implemented across service contexts, and that providers, patients, and members of allied health fields can use this tool.
   3. There are multiple future opportunities that can emerge from this joint day. The goals of this meeting is to identify ways to think about PRS in the long term, to share lessons learned, and to contribute to future endeavors to help understand the translational aspects of PRS.
2. **Network Overview | Eimear Kenny (PRIMED; Mt. Sinai) & Rex Chisholm (eMERGE; Northwestern) |** 
   1. Dr. Eimear Kenny gave a summary of the PRIMED Consortium, including two major goals: (1) Develop methods and improve applicability of PRS across diverse populations and (2) Optimize integration of large-scale, harmonized genomic and phenotype data to facilitate collaborative analysis, dissemination, and development of resources. PRIMED investigators are located across 48 institutions in 9 countries. PRIMED has a focus on diversity in the datasets being gathered, defining gaps in the field to assess where new tools are needed for improving accuracy and generalizability of PRS, sharing resources with the scientific community to enable collaboration, and refining PRS to improve health outcomes. This map shows the countries of participant recruitment across 129 studies (prospective cohorts and biobanks) proposed for use in PRIMED, showing the need for PRIMED to harmonize data across these heterogeneous cohorts. PRIMED consists of 7 Study Sites with varied proposals but who all aim to address the problem of lack of generalizability of PRS from largely Northwest European training populations to other populations. PRIMED is working with AnVIL, NHGRI's cloud-based ecosystem that acts as both the biorepository and compute environment and aligns with FAIR principles and data access control. PRIMED is putting analytic workflows on Dockstore and data models on github. PRIMED is now in Year 2 and to date has been focused on harmonization, setting-up infrastructure, and needs of the community. Over Years 2-4, PRIMED will spend more time seeking collaborative opportunities and results dissemination. At the end of this session, meeting attendees were asked to [brainstorm](https://bit.ly/prs_poll) opportunities for PRIMED and eMERGE to collaborate and the important questions we should be addressing.
   2. Dr. Rex Chisohom summarized the eMERGE network. eMERGE has been active since 2007 and currently there are 10 clinical sites with the coordinating center being Vanderbilt University Medical Center. NHGRI funds the eMERGE network guided by an expert scientific panel. In addition to the CC at VUMC, the Broad completes genotyping for the PRS scores, Duke provides family history through the MeTree tool, and Invitae does monogenic sequencing. The goals of eMERGE include calculating validated PRS for several complex diseases, communicating genomic risk profiles, and recruiting and genotyping 25,000 individuals of diverse ancestry (returning their integrative risk score within their EHRs).
      1. eMERGE agreed that multiple aspects of risk should be returned and developed the Genome Informed Risk Assessment (GIRA). The report displays an individual’s relevant risk factors from 11 common complex diseases using polygenic, monogenic, family history, and clinical risks. Recruitment has started with 7,000 enrolled so far and 20 results returned. The study design includes enrolling participants, completing polygenic risk score assessments and monogenic sequencing combined into a polygenic risk score. The risk score is included within the GIRA which is returned to participants and their providers. Follow up occurs to determine if it made a difference. eMERGE is trying to cover the entire life course including pediatric in addition to adult participants.
      2. The R4 portal was created to take data from all sources and bring them together in a way they can automatically generate the GIRA report. There have been many hours of debate on what to return and what the reporting threshold should be.
      3. The goal is to identify people at risk estimating 6,200 out of the 25,000 with at least one high risk and after getting the return of results, asking if they have adopted the recommendations provided. Accomplishments and progress so far include the rate of recruitment, data moving through the different portals, and the creation of the GIRA. Challenges include applying a PRS to populations with little validation data, how to select PRS that are ready for implementation, how to integrate multiple risk types, and how to address missing data.
      4. Opportunities with PRIMED include sharing lesions for validation and implementation in diverse populations, the potential for joint development of new and improved PRS, entering large scale genomics interaction into cloud repositories, and leveraging eMERGE data to conduct discovery research.
3. **Panel: Conceptualizing and assessing PRS performance in diverse groups. Moderators | Anna Lewis (MGB) & Josep Mercader (Broad) | :** This session focused on the conceptualization, description, and use of ancestry and diversity in the development and implementation of PRS. The panel featured a presentation from each of four panelists, followed by open discussion.
   1. **Continuous and robust conceptualizations of genetic ancestry in PRS development | Eimear Kenny (Mt. Sinai)|** 
      1. Eimear Kenny presented on the complexity of race and ancestry using the BioMe biobank as an example, which collected information about grandparent origin, self-identified “heritage,” and genetic ancestry. She also discussed a method to derive a relatedness matrix using IBD, which can capture additional information about communities compared to PCA and can provide an improvement in PRS for some traits.
   2. **Population descriptors, population-specific methods and PRS development | Gen Wojcik (JHU)|** 
      1. Gen Wojcik discussed the importance of defining “population” and what it means in a PRS context. In Hispanic/Latino groups, adjusting for global ancestry percentage provides better PRS than PCs, but the effect of the global ancestry is not the same across different groups.
   3. **Approaches to improve the performance of type 2 diabetes polygenic scores in diverse and admixed populations by increasing diversity, sample size of genetic discovery, and coverage of linkage disequilibrium reference panels | Alicia Huerta (Broad)|** 
      1. Alicia Huerta presented on improvements for T2D PRS. In PRIMED Study Site D-PRISM analyses, using an in-house LD reference panel created using TagIt for PRS performed better than using HapMap reference populations. In addition, PRS calculated using multiple ancestries produced better predictions.
   4. **An approach to calibration of population variance in implementation of PRS for diverse populations | Niall Lennon (Broad) |** 
      1. Niall Lennon described the eMERGE approach to calibration of PRS with the goal of returning results to patients within a fixed time frame. Because PRS often have poorer performance in groups with non-European ancestries, they adjusted the predicted risk distributions for both mean and variance within a given group and returned results to the participants in the extremes for their group.
   5. **Discussion & Q/A|** 
      1. Following the presentation, the panelists discussed how to move beyond continental ancestry groupings in the PRS context, such as using a genetic similarity construct instead of discrete ancestry groups or including both continuous genetic ancestry variables plus variables measuring risk due to other social constructs, and that the specific approach to a PRS is likely to be disease-specific.
4. **Workgroup breakout sessions**
   1. **Phenotyping fidelity & impact**
      1. The charge is to find areas of mutual interest and opportunities for collaboration across our consortia and what deliverables should be prioritized coming out of this meeting. Areas where two groups are most likely to find shared interest are in cardiovascular conditions, cancer, and diabetes.
      2. PRIMED has sub-working groups using EHR data and cohort data focused on diabetes and diabetes complications, CHD outcomes, cancer, population descriptors, social determinants of health variables, and quantitative traits. Multiple algorithms exist for many conditions, need to determine which ones to use. Some types of information are not captured in all models and databases.
      3. Project 1. Set up infrastructure for sharing of information and ideas between the PRIMED and eMERGE Phenotype committees (monthly phone calls; shared online document space); One goal will be to work with eMERGE to finalize OMOP models for use in PRIMED.
         1. Algorithms need to be updated as information and available medications are available. Listing the most recent/oldest algorithms in a clear way would be beneficial. PRIMED has a new metadata framework that they can share, but it hasn’t been updated on their website. Developing a phenotype catalog with a mechanism for updates would be a tremendous resource. Most eMERGE algorithms are multiple algorithms bundled together. Potential idea is to create a control terminology to more precisely define the way we categorize certain conditions. There are many instances where a single ICD code isn’t sufficient for a condition. PRIMED has tried to mitigate this issue by only using cohorts with better definitions and vetting data sources. This could be leveraged as a best-practice for defining phenotyping algorithms. Want to make sure we think about patients outside of research medical centers to ensure PRS’s are widely applicable.
         2. There is a need for a formal channel for communication between two consortia.
      4. Project 2: Identify 1-2 phenotypes with eMERGE algorithms and examine PRS distributions across strata of environmental/SDoH variables; Provide feedback to eMerge about how algorithms might be further improved to better represent populations across diverse environments. To facilitate these efforts, the group needs to identify members for both consortia to attend calls and contribute to shared resource.
         1. Alisa Manning volunteered on behalf of PRIMED. Need to establish a PRIMED/eMERGE communication group. Existing eMERGE manuscript concept sheets on CHD might be a good initial project for collaboration.
      5. **Action items**: Group will discuss both proposals on next PRIMED Phenotype Harmonization call.
         1. The CC will set up a Google doc to collect information and list issues. PRIMED will create a manuscript proposal and circulate.
   2. **PRS Evaluation**
      1. The eMERGE Network is looking for collaborative project ideas to work on with the PRIMED consortium. Three projects were discussed and plan to move forward:
         1. In Silico Contextualizing the PRS (Champions: Bogdan Pasaniuc, Eimear Kenny, Jordan Smoller, David Conti)
            1. Questions for consideration will include: What variables matter? What quantitative measurements change the risk that the PRS informs? How does accounting for specific variables change an individual’s specific risk (high risk → not high risk)? Which groups of individuals are most impacted by accounting for specific variables?
            2. Variables of interest will include ancestry (continuous), sex, geography, social determinants of health, and family history. Family history, African ancestry, and monogenic variants all trigger the exact same NCCN recommendations for prostate cancer screening even though they each convey very different absolute risks. It is important to keep actual clinical utility in mind when chasing accuracy and precision in risk.
            3. PRS wil have different performances, biases, or precision according to different contexts. PRIMED and eMERGE could come up with context specific measures of precision and use the eMERGE data to see if context specific measures will change individuals going above or below the thresholds.
         2. Refinement of Absolute Risk (Champions: Pradeep and Leah Kottyan)
            1. This project will focus on leveraging clinical, demographic, and geographic data across eMERGE clinical sites to refine baseline absolute risk for conditions relevant to preventive health. Absolute risk across the country can possibly be captured with eMERGE sites being distributed geographically.
         3. PRS Development for Phenotypes in the Experimental Phase (Champions: many - eMERGE to present to PRIMED)
            1. Colon Cancer (Elisabeth Rosenthal, Riki Peters) This is a very well developed PRS that needs more data.
            2. Lupus (Leah Kottyan, Adam Gordon, Bahram Namjou) Northwestern has an ongoing project using eMERGE I-III data developing and testing SLE scores.
            3. Ischemic Stroke (Sally Adebamowo) The University of Maryland (PRIMED) has data from the International Stroke Genetics Consortium (SiGN).
            4. Primary Open Angle Glaucoma (Jibril).
            5. Prostate specific antigen (Adam Gordon) Northwestern has already started gathering some prostate specific antigen data in eMERGE for prostate cancer score analysis.
            6. Monoclonal gammopathy of undetermined significance.
   3. **Cloud resources and data sharing**
      1. The meeting consisted of participants from the eMERGE EHRI and CDS WGs and from the PRIMED Genotype Harmonization and Data Sharing WGs. The eMERGE participants were focused primarily on returning PRS results to patients and clinicians, while the PRIMED participants were focused on preparing data for upload and sharing in AnVIL. The group decided that a good common project was to develop the components of a PRS and the accompanying metadata that can be generated by PRIMED researchers and applied by eMERGE (see [report out slide](https://docs.google.com/presentation/d/1PavkQGxNsh6awW9mcSKowQSZHcuydUMQzqLYyBk3HN0/edit#slide=id.g1edbd4b8485_5_179)).
      2. **Action item**: Small group to draft short term goals and responsibilities and solicit feedback from larger WGs.
   4. **Social, ethical, and policy implications**
      1. This session brought together eMERGE and PRIMED participants to discuss the social, ethical, and policy implications of polygenic risk scores (PRS) across the translational spectrum, from methods development to clinical practice/implementation. The participants brainstormed collaborative project ideas, settling a proposal to outline challenges and lessons learned about defining and measuring population descriptors - race, ethnicity, genetic ancestry, transgender, disability, etc - and how to establish, justify, define, and label population descriptors across phases of discovery, translation, and clinical implementation. They hope to incorporate an evaluation of previous decisions made in the course of work of both consortia. They also suggested identifying what eMERGE shares with PRIMED regarding clinical utility and patient impact of returning PRS to inform upstream development and corresponding data collection.
      2. The group also discussed the importance of being disciplined and thoughtful in language and methodology related to ancestry, race, and ethnicity, with an awareness that the choices made regarding these concepts and terms in methods development later need to be able to be communicated clearly in a clinical setting. They also talked about the need to think about the ethical implications of different results in different populations. They identified the need for additional discussion in order to refine these ideas, such as identifying what population descriptors are currently in use, in what phase, and why they were chosen.
      3. **Action items**:
         1. Coordinate continued discussion between PRIMED and eMERGE in order to refine lessons learned paper ideas - see [joint project proposal](https://docs.google.com/presentation/d/1PavkQGxNsh6awW9mcSKowQSZHcuydUMQzqLYyBk3HN0/edit#slide=id.g1edbd4b8485_5_121).
         2. Identify what population descriptors are currently in use across PRIMED and eMERGE, in what phase, and why they were chosen.
5. **Panel: Translational aspects of PRS. Moderators | Josh Peterson (VUMC) & John Witte (Stanford)**
   1. **Social and ethical considerations for PRS development and application | Megan Shuey (VUMC) |** 
      1. The goal of genetic research is to understand the impact of genetic variation on human health. Polygenic Risk Scores (PRSs) are used for estimating risk prediction. It is important to consider how these scores impact the healthcare system, patients and providers.
      2. There are multiple considerations when thinking about PRS development and applications. These can include considering diversity, equity, and inclusion, support systems for healthcare providers and patients, and understanding how the regulatory environment affects the work.
      3. There is a need for representation from diverse groups. We can increase representation by recruiting and engaging diverse populations. There are research limitations due to the GWAS base study populations. These limitations impact performance evaluation and application/testing. To increase diversity, it's important to encourage participation and engender community trust. Through engagement studios, we can learn what phenotypes/outcomes are of the greatest interest to marginalized and underrepresented patient groups. Information should be shared in plain language with community and thought leaders as well as media reports. Diversifying the workplace by hiring, training, and engaging researchers and providers from underserved communities can be a strategic approach to increase engagement.
      4. It is vital for healthcare providers and patients to have support in understanding results. By creating opportunities for engagement, training, and support for providers, they will be able to provide guidance to patients to help them interpret the results. Patients need support structure to learn what the results mean and understand how risks impact their life.
      5. Regulations affect all aspects of PRS research. As results are returned, there should be considerations of how the results are standardized, reported and shared. Based on the reports, there should be guidelines for uniform uses that provide information on what can be done for the patients based on the information. There are starting to be direct-to-consumer companies and the impacts of these can be detrimental due to the loss of provider/patient relationships. Patients get information about risk without context and this can result in stress. There is a lot of FDA and governmental oversight of results, genetic data return, and marginalization of individuals. There should be additional recommendations and guidance for lawmakers on how to protect patients.
      6. The return of PRS results to consumers is not just on the horizon but is currently happening. We need to prioritize community engagement by seeking, respecting, listening to and returning results. Additionally, we need to prepare and support practitioners and work together to protect, educate, and inform patients.
   2. **Integrating and contextualizing genomic risk with other risk factors | Iftikhar Kullo (Mayo) |** 
      1. Currently, there is a PRS reporting system with various metrics to best assess performance. In breast cancer and CHD, there are existing risk frameworks in which PRSs can be included to create a refined risk estimate. In routine clinical practice for CHD, individuals are assessed for ten year and lifetime risk, and they are separated into low, intermediate, and high risk bins. In integrating the PRS, individuals in the intermediate risk bin can be reclassified. This has actionability. To include PRS, the 10-year risk of CHD is multiplied by their PRS, which provides an updated 10-year risk of CHD. Absolute risk equations can be obtained from large cohorts, preferably ancestry-specific.
      2. A problem that arises is that knowing the absolute risk does not necessarily mean there can be clinical action. Based on results of prior clinical trials, clinical action thresholds can be established. Integrating PRS is most useful for diseases when algorithms are already available to estimate absolute risk, which informs preventative strategies. The breast cancer algorithm incorporates monogenic risk and family history.
      3. In the CHD studies, they have found individuals in both the African American and European ancestry groups reclassified in their risk. However, it is unknown if this is true reclassification, as these individuals would need to be followed for risk events. Family history in CHD is not integrated with PRS, although it is known that family history increases the risk for CHD.
      4. The Mayo group examined the risks of family history, monogenic findings, and PRS independent or additive by looking at the UK biobank data set (200,000 individuals). The risks seem independent and additive. This allows researchers to multiply the risks to assess a crude estimate of the total risk.
      5. There is not currently a good measure of social determinants of health. eMERGE has discussed using surrogates such as zip codes to have an assessment of environmental factors. A solution for measuring social determinants of health could be in multimodality assessment and big data approaches in large cohorts. The All of Us study focuses on collecting environmental data.
   3. **Threshold for evidence for clinical implementation | Beth Karlson (MGB) |** 
      1. The eMERGE study began with 23 phenotypes that were initially nominated to be included. The steering committee conducted votes, based on available literature and data sets, and moved some phenotypes to the developmental pathway, meaning they would not be returned to participants. These scores are still being worked on by sites. This resulted in ten phenotypes being selected for return to participants in the production pathway.
      2. The PRS analysis group decided on two main criteria for including a PRS in the production pathway. The PRS must be validated in at least two ancestry groups out of the four (African American, Asian, European, Hispanic/Latino). The PRS must have significant discrimination with an AUC > 0.5.
      3. Many of the condition groups began with the largest available GWAS data sets, developed PRSs, or took already published PRS and applied it to other groups, and published a validation study. In order for a PRS to be CLIA validated, all results must be published. eMERGE developed metrics to evaluate PRSs. Phenotype groups have followed these reporting standards and created graphs, including AUC and stratifying PRS in different ancestry groups. eMERGE carefully examined the adjusted PPV, which is adjusted for population prevalence. All the evaluation metrics impacted the threshold decided for each PRS. Each condition has slightly different thresholds.
      4. The condition leads wanted to think of analytic viability and feasibility to implement the PRS in clinical practice, how actionable the PRS is, and how translatable it is into clinical practice.
      5. Beth reviewed the Type 2 Diabetes PRS development as an example.Three PCPs at MGB provided their perspective on how type 2 diabetes risk and risk factors are viewed. Condition leads examined the known predictors of risk. The goal was to develop a PRS that would exceed the already known factor of obesity having a 2-fold increased risk for T2D. The T2D group started with three large data sets to input into the PRS-CSX model to develop a trans-ancestry PRS. The trans-ancestry PRS was applied in several validation sets. The score met the criteria of having an AUC > 0.5. An issue the group faced was deciding whether to include the smaller Hispanic group because there was a large confidence interval (3.11, 15.15), even though the adjusted PPV is 0.43. The group decided on transparency and reported the result with the large confidence interval. From the meta-analysis, the stricter the cutoff, the higher the odds ratio. When the cutoff is relaxed, the odds ratios are lower. T2D chose the top 2% as their cutoff.
      6. Other groups considered slightly different factors in selecting their thresholds, including family history, aiming for a specific odds ratio, and balancing sensitivity with specificity. The prostate cancer group has a larger threshold of 10%.
      7. It would be beneficial for PRIMED and eMERGE, as experts in the field, to determine the criteria necessary for PRSes to be applied in clinical practice. The clinical utility and analytic validity must be considered.
      8. A professional society typically develops and establishes guidelines, such as ACMG. The US Preventive Services Task Force could be the deciding group.
   4. **Discussion & Q/A** 
      1. The GIRA has a lot of uncertainty. Are there suggestions on how the uncertainty can be accurately communicated with patients? What level of detail should be provided to patients?
         1. Shared decision making and explanatory, straightforward figures are used. Uncertainty is present in almost every clinical measure. There is uncertainty present in ancestry as well. The majority of PCPs surveyed reported not discussing numbers with patients, and focused on risk factors such as diet. Clinicians are often not quantitative with patients.
         2. An example of how a PRS could be used is for statins. Most patients are in the intermediate zone of whether to start a statin medication. Having a high risk PRS could move the patient into a higher risk zone to start taking a statin.
         3. eMERGE has taken the approach to be transparent, and address participants who are not represented within the four ancestry groups or who belong to multiple ancestry groups. The language used includes ‘we do not have enough data’ and ‘researchers are still studying this.
      2. Not all the PRSs in eMERGE are applicable across ancestries. The limitations in developing them were not methodological, but based on the data limitations encountered. The efforts that PRIMED is making with a broad scope of representation will be beneficial.
      3. In the CHD integrated score, the CHD group has integrated the Pooled Cohort Equation and the PRS. In heart disease, the regression estimates are being assumed to have no interaction. Adding the monogenic and family history results appears to be independent and additive.
      4. In the GIRA report, result categories are presented in sections, and the study is asking clinicians to integrate them in practice. Some clinicians are uncomfortable with the odds ratios and not having an integrated score to use in clinical practice. The main issue with developing an integrated score is the lack of data. It requires a large and diverse cohort, followed for many years to gain enough outcomes and endpoints, with measures in place for all the clinical risk factors and social determinants of health.
6. **Joint projects from breakout sessions. Moderators | Clement Adebamowo & Rex Chisholm |** 
   1. Phenotyping fidelity and impact
      1. This breakout session identified as a task to set up a pipeline or infrastructure for the members leading the phenotype harmonization in both consortia can communicate on a regular basis. There can be a monthly call and an online shared document space for collecting information.
      2. The goal from this is for the consortia to work together to finalize the OMOP models for use in PRIMED.
      3. A specific project discussed was to identify one to two phenotypes with existing eMERGE algorithms and examine the PRS distributions across strata for environmental and/or social determinants of health variables. The group can compare the distributions of PRS across the strata, and provide feedback to eMERGE on how the algorithms may be further improved to better reflect individuals across diverse environments.
      4. The PRIMED Phenotype Harmonization work group will discuss the tasks and goals on their next call, and the PRIMED CC will begin working on setting up collaborative materials. PRIMED will begin an MCS for the project discussed.
      5. Alisa Manning and Leslie Lange (PRIMED) will work with eMERGE to set up a monthly call and shared documents. Iftikhar Kullo (PRIMED & eMERGE) will share his lab’s MCS with the PRIMED work group. Leslie Lange will begin the proposal.
      6. Wei-Qi Wei (eMERGE), Joon (PRIMED), Jennifer (PRIMED), and Chunhua Weng (eMERGE) will support the PheKB user interface and eMERGE phenotype algorithm exchange efforts. There was a lot of feedback on improving the PheKB user interface, and eMERGE is working on improving the meta data use. PheKB does not have a cloud version. There is only an OMOP version. It is possible for the Phenotyping workgroup to start using some common algorithms, such as CHD, and moving them into cloud use.
   2. Social, ethical, policy implications
      1. This breakout session held an in depth conversation about population descriptions and diversity. The project identified was to come together and publish on the challenges and lessons learned through defining and labeling population descriptors.
      2. The work is on how to establish, justify, define, and label population descriptors across phases of discovery, translation, and clinical implementation.
      3. The group would like to work with ClinGen, including their Ancestry and Diversity workgroup.
      4. The plan is for the relevant members of each consortia to meet regularly to organize and develop the lessons learned.
      5. The leaders for this will be Ingrid Holm (eMERGE), Gen (PRIMED), and Stephanie (PRIMED). Discussions on how to communicate uncertainty in PRS can be also held in this group.
   3. PRS evaluation
      1. The group focused on coalescing some ideas for projects the consortia can work on together by identifying gaps in the field that can be addressed. The group used the introspective exercise for ideas.They identified three project ideas.
      2. Project 1: To use an in silico design to think about how PRSes are being contextualized in clinic. This will address questions about if PRSes change and why, and what impact changing PRSes would have on patients. This would also examine if this can be quantified. eMERGE focuses more on the clinical side of PRS, and this can inform both groups on what clinical factors can impact or bias PRS. The champions for this project will be Eimear Kenny, Bogdan, Jordan, and Dave Conti.
      3. Project 2: Refinement of absolute risk. There is an opportunity through eMERGE to work with clinician experts on the conditions of interest, and how to collect different clinical information across different geographic sites.This can help refine understanding of absolute risk for disease.
         1. The champions for this project will be Pradeep and Leah Kottyan.
      4. Project 3: Identification of PRS scores that were worked on by eMERGE that were not ultimately selected by eMERGE for implementation. eMERGE investigators spent a lot of time and effort on these scores, and it would be beneficial to bring them to PRIMED. Individuals in the breakout session volunteered to move the PRSs forward. Beyond method development, it would be worthwhile for the PRIMED network to learn the knowledge from eMERGE, and see if there is traction for PRIMED investigators to work with eMERGE investigators to continue to develop the scores. Some conditions listed were beyond the original eMERGE developmental conditions.
   4. Cloud resources and data management
      1. There is an opportunity in bridging between the discovery domain of PRSs that PRIMED focuses on, and looking at the eMERGE side of implementation of the PRSes.
      2. Their project focuses on the representation of PRSs as computable knowledge that can be used and generated by PRIMED. This downstream would be leveraged by eMERGE for use in the EHR and CDS. There is a lot of work regarding metadata and contextualizing the scores clinically.
      3. The group identified points of contact as Matt, Ben, Whitney, TBN/methods (PRIMED), Bob Freimuth, Luke Rasmussen (eMERGE), and Robb Rowley (NHGRI).
      4. Another common bridge is AnVIL. PRIMED is representing scores in AnVIL for new knowledge, and Robb Rowley is leading the efforts on how AnVIL can be used clinically for eMERGE.
7. **Closing remarks | Clement Adebamowo (PRIMED; U of Maryland) & Rex Chisholm (Northwestern)**
   1. Rex and Clement concluded the joint day with recognition of the accomplishments and productivity of the meeting. It will be important to identify champions to move forward the collaborative project proposals and to set up additional meetings across PRIMED and eMERGE to discuss these concrete projects. We also need to collectively pay attention to identified gaps, including ELSI considerations and what will be needed to engage care providers and professional societies so they can leverage our outputs.

**Recordings:**

|  |  |  |
| --- | --- | --- |
| **Recording Name** | **Links** | **Passwords** |
| Main room: | <https://zoom.us/rec/share/ylsSzS7tTKVmltfpz0A2eJ-UZ9DyJwOehK3FIQUre3NMHHAYJae5fayZUHApt9oe.FCQ4cQ_X98RoIf2n?iet=qCJind--9s4h8tBo2JoCioGBtdJHrUz24H_cAMDNcrM.AG.mTpRD16-0bxa1c-C64wX5oVHujKwmg3T17uyMoLl6p8eVmy2-aeZjCKK-hXRJIV7OeNRQoWTEKNMoW8Y4dL50YNMbepa0_wZCZb5nPho5_SoUiqo7A9VxwyLXTIWbpAC.pc59vZso7aGZTZaJzEaQ2g.hGY6jXlasQ8NxLLw>  <https://vimeo.com/799604258/83c3512018> | BKd\*Uvu7 |
| Phenotyping fidelity & impact breakout session: | <https://zoom.us/rec/share/ylsSzS7tTKVmltfpz0A2eJ-UZ9DyJwOehK3FIQUre3NMHHAYJae5fayZUHApt9oe.FCQ4cQ_X98RoIf2n?iet=qCJind--9s4h8tBo2JoCioGBtdJHrUz24H_cAMDNcrM.AG.mTpRD16-0bxa1c-C64wX5oVHujKwmg3T17uyMoLl6p8eVmy2-aeZjCKK-hXRJIV7OeNRQoWTEKNMoW8Y4dL50YNMbepa0_wZCZb5nPho5_SoUiqo7A9VxwyLXTIWbpAC.pc59vZso7aGZTZaJzEaQ2g.hGY6jXlasQ8NxLLw>  <https://vimeo.com/799604258/83c3512018> | BKd\*Uvu7 |
| PRS Evaluation breakout session: | <https://zoom.us/rec/share/3DZA8hiCPR7dxcA-ZeeCv04vRl88A0ypP_MxekVPkmpRBVRsq7cSBdZBHniJIFuW.41h3e_5luHx2RDQ2?iet=bTSusqU31SRPFS-t9EeiuYS3HaDWIRFpahv27koID2U.AG.ap1BI7D_8aSCy_d9ZuCH0Fbxfkb5PHR3qeLtR_uy-ECVPm-8TdMKy-hFjT3jfoHZS5VtQfJw1cdzRqEI7BhGoG9phbzwkedp0u3xQes1eD6EzZ00hrqxF10CyD8YDzH6.S4lVcLRquaSOG9vuqX2AHQ.YQtZVhl2toTUQgHP>  <https://vimeo.com/799594541/6d36f9f3c6> | Qy3z8k^9 |
| Cloud resources and data sharing breakout session: | [https://zoom.us/rec/share/C5Pvk0YiH7C1ZyGZk4qbZRuAvFEFv5GUmyB20TWnRip52V7\_ye6GjZWuREXHLOu2.BuyEvjlmqP-0Nlx-](https://nam12.safelinks.protection.outlook.com/?url=https%3A%2F%2Fzoom.us%2Frec%2Fshare%2FC5Pvk0YiH7C1ZyGZk4qbZRuAvFEFv5GUmyB20TWnRip52V7_ye6GjZWuREXHLOu2.BuyEvjlmqP-0Nlx-&data=05%7C01%7Csophie.forman%40vumc.org%7C8b14e622855f472c695808db08576014%7Cef57503014244ed8b83c12c533d879ab%7C0%7C0%7C638112943582672365%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=H7%2BigzoPbtluKD0jWCbpQxb%2FP1Llp9HMlnaRlQwXW9s%3D&reserved=0)  <https://vimeo.com/manage/videos/799576170/070eef4d31> | fHj^a1am |
| Social, ethical, and policy implications breakout session: | <https://zoom.us/rec/play/8vkOrzPFnm5I4oR93e4JlQaUMR-TeaLHXlYiaui-hS1n0o1kx1kWCXBCNsDV1C8xzNABePjs_ZXQKT0.Ov40LApnoFX8JaB0?continueMode=true&iet=-7yeiLU9AZt8T2QjyyUgwnSKHfjeRes8tyz54WVlIpE.AG.BYQoGjuPhkVlzzIt2cyJ4lvtyjrGHET8t-F0JQjOAUMvgFvNpD0QA45lFFvxZOhuRiqCZt_mgQ8exQHPGWonVBVqvNBnFmsjHNqopBeaEPY8V61BO0g5Qrj3-WEdNmQz.lSFgohey1_26Ad_JwmhaQw.JH8VvW6JB5BZ6huz>  <https://vimeo.com/799580334/58afeff83d> | %ykyc69= |