**Summary of External Scientific Panel/Steering Committee Meeting: September 2023**

September 20-21, 2023, Zoom & In-Person

|  |  |
| --- | --- |
| **eMERGE Day 1: Wednesday, September 20th, 2023** | |
| **Time** | **Event** |
| 9:00-9:15 AM | NHGRI Program Official Report | Robb Rowley (NIH/NHGRI) |
| 9:15-9:30 AM | Announcements, opening remarks | Rex Chisholm (SC Chair, Northwestern) |
| 9:30-10:50 AM  9:30-9:50 AM  9:50-10:10 AM  10:10-10:30 AM  10:30-10:50 AM | Panel: GIRA Across The Network  Confidence in accuracy of the GIRA | Jackie Ogdis (Mt. Sinai)  Data missingness assessment for GIRA generation | Jennifer Pacheco (NU)  Return experiences to participants | Shannon Terek (CHOP)  Q&A |
| 11:10-11:30 AM | Investigator Presentation: Differential performance of polygenic risk scores across groups: real-world experience of the eMERGE Network | Anna Lewis (MGB) |
| 11:30 AM-12:40 PM | Workgroup breakout session one  Recruitment & Return (CARE, R2/ELSI/sIRB, & QA/QC Taskforce)  Data Utilization (GRID) |
| 1:25-1:45 PM | Investigator Presentation: Low pass sequencing of children with African ancestry and electronic medical records to Cincinnati Children’s Hospital | Leah Kottyan (CCHMC) |
| 1:45-2:30 PM | Return challenges & successes | Josh Cortopassi (UAB) & Brenna Boyd (Columbia) |
| 2:45-3:05 PM | Clinical Operations Update | Katie Larkin (Broad) & Sienna Aguilar (Invitae) |
| 3:05-4:15 PM | Workgroup breakout session two  Outcomes (Provider Uptake & Outcomes, Phenotyping, & QA/QC Taskforce)  Genomic Data Integration (CIRT) |
| 4:15-4:20 PM | Closing remarks | Rex Chisholm (SC Chair, Northwestern) |

|  |  |
| --- | --- |
| **eMERGE Day 2: Thursday, September 21st, 2023** | |
| **Time** | **Event** |
| 9:00-9:10 AM | Opening remarks & comments from ESP chair | Robb Rowley (NIH/NHGRI) & Dan Rader (University of Pennsylvania) |
| 9:10-9:30 AM | Network overview: Priorities, goals, progress and ESP recommendations | Rex Chisholm (SC Chair, Northwestern) |
| 9:30-9:50 AM | Recruitment | Digna Velez Edwards (VUMC), Ingrid Holm (BCH), & Wendy Chung (Columbia) |
| 9:50-10:15 AM | Comprehensive Risk Assessment & Return (CARE) | Gail Jarvik (UW), Iftikhar Kullo (Mayo), & Cindy Prows (CCHMC) |
| 10:35-11:00 AM | Provider Uptake & Outcomes | Nita Limdi (UAB) & Dave Veenstra (UW) |
| 11:00-11:20 AM | Phenotyping and Outcomes data collection | Chunhua Weng (Columbia) & Wei-Qi Wei (VUMC) |
| 11:20-11:45 AM | QA/QC Taskforce | Jennifer Pacheco (NU) & Lisa Martin (CCHMC) |
| 12:25-12:45 PM | Genomic Risk Innovation and Discovery (GRID) | Adam Gordon (NU) & Matt Lebo (MGB) |
| 12:45-1:05 PM | CIRT | Eta Berner (UAB), Emma Perez (MGB), Bob Freimuth (Mayo) |
| 1:30-1:55 PM | Input/Feedback from the ESP, general discussion |
| 1:55-2:00 PM | Closing remarks | Rex Chisholm (SC Chair, Northwestern) |

|  |
| --- |
| **Other Links** |
| Action Items |
| Recordings |
| Official ESP Recommendations |
| ESP Packet |
| Background Packet |

**eMERGE Day 1: Wednesday, September 20th, 2023 |**

1. **NHGRI Program Official Report | Robb Rowley (NIH/NHGRI)|** 
   1. There has been a lot of impactful progress in the eMERGE network over the past year.
   2. The eMERGE-AnVIL collaboration will provide cloud computing capabilities for genomic medicine research.
      1. The AnVIL collaboration is intended to simplify collaboration and data sharing between sites.
   3. Upcoming NHGRI milestones:
      1. July 2024: The NHGRI request for budget needed for extension due.
      2. October 2024: Submit the request for extension with cost.
      3. The extension request will be due March 27, 2025.
   4. The NHGRI will host the Genomic Medicine XV: Genomics and Population Screening meeting on November 8-9, 2023.
   5. The NIH is inviting comments and feedback on a proposed update to the NIH mission statement.
   6. Jessica Chong, P.A.-C. And Karyn Roberts, Ph.D., R.N. are the new NIH-ACMG fellows.
   7. There are multiple NHGRI funding opportunities:
      1. Notice of Special Interest: Clinical and Translational Science Award Program: Collaborative and Innovative Acceleration Award for Advancing Recruitment through Trial Innovation Network applications are due on October 17.
      2. Genomics enabled learning health systems (gLHS) applications are due on November 7, 2023.
      3. Supporting Talented Early Career Researchers in Genomics applications are due on February 27, 2024.
      4. Advancing Genomic Medicine Research will be an upcoming funding opportunity.
2. **Announcements, opening remarks | Rex Chisholm (SC Chair, Northwestern) |** 
   1. Goals: GIRA return and ESP
      1. Day 1 will focus on GIRA across the network. This will include identifying and agreeing upon what GIRA concerns remain in order to increase confidence in the GIRA and increase return rate of GIRA.
         1. Increasing the rate of return will need to happen in order to meet milestone targets.
      2. Day 2 will focus on the ESP report out, including rate and network confidence in the GIRA, recruitment and return progress, and outcomes plans, workgroup progress/evolution and discovery initiatives.
   2. GIRA generation workflow
      1. Steps/elements of the GIRA generation
         1. Participant data (EHR data, participant survey, family history)
         2. Genetic data (PRS (Broad), monogenic (Invitae)
         3. GIRA generation (implementation of specifications, presentation of output, edge cases)
         4. GIRA return (patient and provider comprehension, confidence in GIRA prediction, edge cases)
      2. Confidence in GIRA requires consideration of components:
         1. Quality and accuracy of processes to obtain participant and genetic data.
         2. Algorithms and logic selected by the network.
         3. Implementation of specifications and GIRA presentation.
         4. Detection and management of edge cases.
         5. Comprehension of participants and providers.
   3. Enrollment progress: 17,000 enrolled.
      1. We do expect that enrollment will slow down a bit as sites have maxed out on recruitment of white participants and need to focus on recruitment of non-white participants.
   4. GIRA generation & return progress
      1. Nine of ten sites are returning GIRA.
      2. Total GIRA returned to EHR is 1,745.
         1. This is a low number relative to the 25,000 target.
         2. Need to think about what we need to do to ramp it up.
   5. Counts of GIRA triggers by condition
      1. The total number of high risk GIRA generated by condition was broken down into the number triggered by the PRS.
      2. We are seeing many conditions with high contributions of family history triggering a high risk designation for a participant.
         1. These numbers are higher than anticipated. Because of this, many sites are seeing higher percentages of participants with a high risk.
         2. We want to spend some time thinking about the role of family history in high risk designation.
         3. Family history reflects your environment and your genetics. For example, it makes sense that family history of the lifestyle behavioral factors that provide you coronary heart disease contributes to risk independent of the genetics.
      3. Difference in # GIRA observed and expected
         1. There may be things that are environmentally driven that are affecting the differences.
      4. Challenges ahead
         1. Collating and cleaning interim data freeze.
         2. Ensuring the data we collect is useful and sufficient for outcomes analysis.
            1. The interim data freeze will help see how we are doing and if there is a lot of missing data.
         3. Increasing the rate of return.
         4. Goal: Recruitment completed by June 2024 (Current deficit: 7,882).
         5. Goal: Return completed by October 2024 (current deficit: 23,255).
3. **Panel: GIRA Across The Network**
   1. **Confidence in accuracy of the GIRA | Jackie Odis (Mt. Sinai) |** 
      1. Enrollment at Mount Sinai began in May 2022 and 1,929 participants have been enrolled with a majority female sex assigned at birth with a large Hispanic, Latino, or Spanish percentage.
      2. Return of results at Mount Sinai began in February 2023 and 227 results have been returned so far.
      3. During Alpha and Beta testing, an in-depth review was performed on GIRAs and issues were reported in the GIRA Live error log resulting in a few amendments. Mount Sinai paused RoR to focus on developing a thorough QA/QC process.
      4. A QA/QC process was developed using the GIRA logic knowledge base and a detailed condition specific info tab was added to provide additional information regarding variables that feed into the GIRA.
      5. Score Verification was performed based on participants variables in Canrisk. In July 2023, the CC modified the module variables instrument to allow for easier comparisons. Mount Sinai also kept an internal tracker to calculate differences between R4 calculated scores and manual verification.
      6. MeTree verification was performed to review pedigrees in detail to confirm that triggers were calculated correctly.
      7. In addition, surveys were reviewed to confirm that Clinical Risk Factors were populating correctly
      8. For breast cancer verification, the average difference between R4 and Canrisk was 2-3% with all Boadicea estimates being higher than Canrisk manual calculation.
      9. For CHD verification, an average difference between the R4 calculation and a manual calculator was found to be <1%.
      10. Example 1 described participants who had inaccurate language for individuals with personal history of cancers. This was updated in a recent amendment.
      11. Example 2 described a participant who had a lifetime breast cancer risk and with the CC determined that R4 calculates total lifetime risk and not remaining risk.
      12. Alpha/Beta testing has improved GIRA confidence. Processes for manual review may be changed after beta testing to change who gets
      13. There was a question regarding how long it takes to complete a review. The answer was 21 minutes and getting faster.
      14. There was a question regarding the imputation of age of relative. This was confirmed to be a year prior to current age or age of death.
      15. A question regarding lifetime risk versus residual risk was tabled for later discussion.
   2. **Data missingness assessment for GIRA generation | Jennifer Pacheco (NU) |** 
      1. It is assumed that if participants spent enough time on a provided survey or MeTree, they added all they knew and it is impossible to separate missing versus condition nonexistent.
      2. There is a disclaimer stating that risk can only be provided based on what is provided by the participant.
      3. Some possible reasons the GIRA is not complete and there is possible passive withdrawal include participant information missing/incorrect, MeTree/family history not completed, sample not collected, or DNA failed genotyping (at Broad or Invitae), among others.
      4. The R4 dashboard displays around 90% completion of baseline and pre-RoR surveys and around 70% completion of family history data from all current eMERGE participants.
         1. R4 is not tracking sample information until the sample is sent to the Broad.
      5. Only 20% of participants have PRS data and 34% have monogenic data.
      6. Missing/incomplete MeTrees are a big contribution to missing GIRAs.
      7. Some baseline survey errors include errors with BMI like height entered without weight and heights and weights impossible values.
      8. UW completed a statistical analysis on individual response missingness from R4 broken down by race and ethnicity. BMI again was missing due to missing data.
         1. This script is available on GitHub here.
         2. There is also an All of Us paper on survey missingness and how skipping questions like demographics led to more missingness overall which can be found here.
      9. UAB’s average completion time for MeTree is 22 minutes.
      10. Across the network, only 79% of participants that completed the other surveys also completed MeTree.
      11. Even with applied limits to data, like blood pressure for example, there could still be errors.
      12. The QA/QC taskforce is working on creating tools to assess missingness and developing an implementation guide in addition to shareable code to run ar sites to complete assessments.
      13. Next steps include tracking passive withdrawals and their reasons, creating and disseminating reports in R4 where possible, and discussing better ways to keep minority subjects in the study.
      14. UAB checks participant entries before data is locked so errors can be fixed.
   3. **Return experiences to participants | Shannon Terek (CHOP) |** 
      1. Current CHOP numbers: 1,295 participants consented, 979 samples sent to the Broad, 652 PRS results received, 22 Invitae samples submitted (all negative), 483 GIRAs placed into Epic (62 High Risk GIRA, 421 Not High Risk GIRA).
      2. While CHOP’s experience is pediatric heavy, many experiences can be applied to eMERGE participants.
      3. RoR process at CHOP includes GIRA component review, GIRA approval and generation, Result components entered into EHR, PDF loaded into EHR, MyChop message sent to participant and provider (based on Network cover letter).
         1. CHOP has a Clinical Decision Support system in place for the 4 High Risk conditions.
            1. When an individual is identified as high risk their PCP will get a Best Practice Alert showing care recommendations, order sets, and resources.
      4. EHR data tracking at CHOP includes overall numbers of results and breakdown by condition, reports on if/when parent and PCP have viewed results, access logs to track if PDF was opened, message logs for all patients in the study, time to well visit before or after RoR session, capture of CPT codes in time-frame before/after RoR.
      5. Not High Risk RoR Experiences (483 in EHR).
         1. 40% of the participants have viewed the results/MyChop message within one month, with an average time of ~5 days.
            1. Those who have not viewed the results/message have their results mailed.
         2. While all participants have the ability to reach out for any questions/concerns, at this point only a handful have.
            1. Feedback has been both positive and negative, with one participant expressing excitement regarding results and ease in viewing in MyChop and one participant indicating they were disappointed in the results and felt that the report needed more graph colors and that the results were downplayed.
         3. Not High Risk results example (Family 1)
            1. Two siblings enrolled (17 year old male diagnosed with asthma at age 10, 11 year old female diagnosed with asthma at age 6).
            2. Both came back ‘not high risk’ for all 4 conditions.
            3. The mom called the same day as GIRAs were uploaded with concerns (including the accuracy of the results and whether there is something that they could have done to prevent the condition). The genetic counselor spent some time explaining multifactorial conditions, addressing guilt, and answering questions/concerns.
      6. High Risk RoR Experiences (62 in EHR).
         1. 13.3% of PRS results are High Risk. This aligns with expectations.
         2. Majority of CHOP’s High Risk is obesity, followed by asthma and then Type 1 Diabetes and Type 2 Diabetes.
         3. Disease prevalence in participants identified as High Risk: many participants had already been diagnosed with the condition. This may be something to take into consideration when returning results.
         4. Return of Results visits: 34 in person (telehealth or phone call visits), 14 could not schedule, 4 declined visits, 3 have been scheduled, 7 are pending scheduling.
            1. CHOP is monitoring race related to High Risk and also RoR.
         5. High Risk RoR challenges:
            1. There has been difficulty in scheduling RoR visits.
            2. The children typically do not join the RoR visits.
            3. There is a high prevalence of conditions and parents do not necessarily see the need to discuss results related to a condition their child has already been diagnosed with.
            4. Many participants have “Bigger issues” going on medically (cancer, cardiac issues, intellectual disabilities).
         6. High Risk RoR successes:
            1. Parents have generally been responsive and happy with the results.
            2. It has been easy to share notes with PCPs and parents following RoR visits.
            3. Telehealth allows flexibility in schedules for appointments.
         7. High Risk results example (Family 2)
            1. Family of seven with three enrolled children. Two out of the three children enrolled tested High Risk for Type 1 Diabetes.
            2. One of the children not enrolled has a diagnosis of Type 1 Diabetes (identified during chart review prior to RoR visit).
            3. The genetic counselor was able to discuss results for all three children during RoR and was able to explain to the parent that all of her children were at increased risk due to the sibling’s diagnosis.
            4. The parent had concerns about the 5 year old due to excessive thirst and urination. These concerns had not previously been brought up to PCP and it was recommended that the parent talk to the PCP and the genetic counselor messaged the PCP following the visit to relay the parent concern.
            5. The child was tested further and testing was negative.
         8. High Risk results example (Family 3)
            1. Eight year old female tested High Risk for obesity.
            2. Participant’s parents have concerns regarding weight and the participant has visited the healthy weight clinic in the past with no improvement.
            3. The participant also has severe immunodeficiency and previous genetic testing including negative microarray and several VUS on Whole Exome Sequencing.
            4. The focus of the RoR was to help alleviate feelings of defeat by discussing the genetic side to obesity. While it was difficult to not be directive, the genetic counselor did discuss ways to manage weight knowing there is a genetic component involved.
            5. There is a high prevalence of obesity in the pediatric population and those who are High Risk tend to be very off their growth chart.

Factors such as autism, medication, behavior, and living situations have all been brought up as reasons why there is difficulty in maintaining a healthy weight.

* + - 1. High Risk results example (Family 4)
         1. Eight year old female tested High Risk for Type 2 Diabetes (T2D).
         2. The participant was not diagnosed with T2D and was currently at the 97% for weight. The participant does have a diagnosis of sickle cell disease.
         3. The mother has a history of gestational diabetes and is prediabetic.
         4. Challenges to the RoR included figuring out how these results fit into an already health-complex child.

The mother expressed that her main concern is the participant’s sickle cell disease. She was excited about how these results would provide her more reason to continue to find ways to promote a healthy lifestyle, which had already been discussed by the PCP as impactful to sickle cell disease outcomes.

* + - 1. High Risk results example (Family 5)
         1. 12 year old female with a medical history of autism, anxiety and atopic dermatitis and a family history of T2D and asthma.
         2. The GIRA came back as High PRS for asthma and T1D and Not High for T2D and obesity.
         3. The GIRA was uploaded in the EHR and multiple scheduling attempts were unsuccessful.
         4. 10 days after the GIRA upload the participant presented to the ED with a syncopal episode, chest pain, facing heart and shortness of breath. ECG and labs were normal and the patient followed up in cardiology 5 days later. ECG and echo were normal. Chest pain was reproducible on exam. There was consideration for musculoskeletal pain or exercise induced asthma.
         5. The PCP followed up with the family by phone and prescribed albuterol as needed.
         6. The GIRA RoR visit was completed 4 weeks after the EHR upload and risks were reviewed. The parent relayed that they had read the results a few days before the syncopal episode and the results and possibility of asthma helped them stay calm during the episode.
    1. Chart review performed on participants who had received High Risk results AND had a Well Child Visit since receiving the results (n=8 as of 8/29/23).
       1. Five of eight were able to meet with a genetic counselor before seeing the PCP.
       2. All PCPs reviewed the results in the EHR but only four of eight utilized Smart Tools created for eMERGE.
       3. Three of eight discussed eMERGE results and Care Recommendations with the patient, but no next steps were made. Reasons included parent declined, child already going to the healthy weight clinic.
    2. Two new genetic counselors at CHOP are doing role playing for returning results. If any sites need practice returning results to pediatric patients, please reach out to CHOP.
    3. Results should be communicated to participants as “below study threshold”, since “not high risk” is an internal phrase.
       1. The cover letter which is sent to “not high risk” participants uses the “below study threshold” phrasing.
    4. Access logs can be used to determine when participants opened the GIRA report. However, there is no way to tell how long was spent reviewing the GIRA or which pages of the report were viewed.
       1. Demographics of participants who meet in person for return of results are being studied. In the future, It will be important to also study the demographics of participants who are not opening the GIRA.
  1. Q&A
     1. Lifetime risk of breast cancer, rather than residual risk, should be communicated to participants.
        1. There is no maximum age of participants for return of breast cancer risk, except the study limit age.
        2. The NCCN guidelines recommend the use of lifetime risk rather than residual risk.
        3. There is an option to view lifetime risk on the Canrisk website.
        4. There is an SOP for return of results, which will need to be updated to include how lifetime risk is communicated to participants.
           1. ACTION ITEM:The Breast cancer group will generate talking points on explaining lifetime and residual breast cancer risk during return.
        5. Sites should consider manually reviewing participants who have a breast cancer score >20, rather than using >15 as a threshold.
        6. Family history alone can trigger a high risk breast cancer score. However, this will require a very complete family history, including accurate ages.
           1. This is a challenge for participants who are not flagged as high risk based on study criteria, but have concerning family histories.

The nature of the study design means that risks will be reported as binary: high risk, or not high risk.

Sites can encourage follow up visits with physicians for risks that are outside the scope of the eMERGE study.

1. **Investigator Presentation: Differential performance of polygenic risk scores across groups: real-world experience of the eMERGE Network | Anna Lewis (MGB) |** 
   1. The differential performance of PRS by group is one of the major ethical barriers to their clinical use.
   2. The danger that we inadvertently promote the idea that there are significant genetic differences between socially-defined groups.
   3. Performance is the strength of the association of a PRS with the outcomes it is designed to predict.
   4. How to define the groups used for validating PRS performance:
      1. Validation cohorts relied on different ways of measuring group membership and used different types of population labels.
      2. The network discussed best practices and methods to genetically define groups but left it up with condition leads to choose how to define groups.
      3. There are limitations to this approach. The groups chosen aggregated individuals perceived to be similar based on determinations arising from different approaches to measuring. The 4 buckets fail to account for the extensive human genetic variation known to exist.
   5. Use group-specific scores or the same score for everyone:
      1. It was decided at an early stage to not use group-specific scores but instead use the same PRS calculation for every individual.
   6. Include PRS for which is was not possible to validate in all the defined validation groups:
      1. Prospective patients were not concerned about the lower predictive power of PRS outside European descent populations and were comfortable receiving scores even if inaccurate for their ancestry.
   7. Communicate the differential performance between validation groups:
      1. Possible strategies included reporting performance by validation groups and reporting an aggregate performance - giving a range encompassing the lowest lower range and the highest upper range of the confidence intervals across the groups.
      2. A layered approach was adopted. On the GIRA, the most prominent information is whether an individual was identified as high risk for any condition.
      3. The PRS report first indicates in which conditions the individual was determined to be high risk.
      4. To aid in conversation between patient and provider for the case where participants identify with none of multiple validation groups, FAQs were included during recruitment.
   8. What terminology to use to describe the different validation groups:
      1. The established validation groups included different ways of measuring human difference.
      2. Using the word population was decided against due to ambiguity.
      3. MGB consulted with a group who recommended use of the term “descent” to strike a balance between scientific accuracy and lay public understanding.
   9. How to explain the differential performance of the PRS reported:
      1. Without an explanation as to why there are differential performance of PRS, inaccurate conclusions could be drawn about (for example, different predispositions of groups to the same condition in question).
   10. Early decisions around validation groups have large downstream consequences and decisions all have ELSI dimensions.
   11. Other PRS implementation efforts will have these same challenges and hopefully these experiences will be useful. There is currently a lack of guidance for implementers.
       1. It would be useful to explicitly capture genetic similarity from between training and test data and actually model that into the numbers that are shown to patients.
2. **Workout breakout session one**
   1. **Recruitment & Return (CARE, R2/ELSI/sIRB, & QA/QC Taskforce) |** 
      1. Effect of withdrawals
         1. There is a need to make decisions as a Network about what constitutes a withdrawal.
            1. If there is no DNA sample at 1 year post recruitment (or kit sent)

As we get closer to recruitment end dates we may need to look at how long we are waiting to send in samples.

* + - * 1. ACTION ITEM: The QA/QC, R2/sIRB, CARE, and Outcomes workgroups will define active withdrawal, administrative withdrawal, and loss to follow up definitions. The CC will update terminology for withdrawal on R4.
      1. There are different combinations of missing data and guidance is needed regarding what to do with each.
         1. The first data freeze will help assess missingness.
         2. There is a racial discrepancy in those who complete the baseline survey.
         3. Pediatric conditions are not triggered by anything in the survey. A complete GIRA can be generated from family history and BMI in the absence of a baseline survey.
         4. We do not want to lose participants if we have the ability to generate an accurate GIRA without survey data.
         5. Once PRS is generated it needs to be returned to the EHR.

For positive samples with missing survey data we are able to do chart review and return.

* + - * 1. ACTION ITEM: CARE & QA/QC will complete the missing data combos matrix and provide overall guidance to the network regarding tolerance for missingness. The groups will also provide recommendations for when participant withdrawal should be considered.
      1. It will be difficult to hit exactly 25,000 participants.
         1. The Broad is limited to 25,000 samples.
         2. Sites are concerned about whether it is better to go over or under.
         3. All sites need to balance between underrepresented, pediatric and white participants.
         4. GUIDANCE: Sites can overshoot enrollment by ~100 individuals to ensure enough participants in the study to account for withdrawals and loss to follow up.
    1. Timeframe for collection of variables
       1. There is a need for standards regarding how far back in the EHR to pull missing data for variables for GIRA generation.
          1. In the case of no MeTree data for those identifying as female at birth, GIRA can still be generated, there will just be no breast phenotype assessment.
          2. GIRA can still be entirely generated for all other conditions.
       2. DECISION MADE: Sites should make every effort to obtain missing survey data. As long as the participant has provided a genetic sample the site can generate and return GIRA.
          1. Sites are encouraged to obtain missing survey data from the EHR if feasible, and this will be documented.
          2. If a participant has never provided a sample, or if they do not complete the baseline surveys and no sample is available, sites can withdraw and replace the participant.
       3. There will need to be a distinction made from an outcomes perspective.
          1. Certain missing data may exclude participants from analyses even though GIRA was returned.
    2. Phenotype specific templates
       1. CC has created a template for each condition regarding handling of missing data, imputations and GIRA generation.
          1. Missing data is more critical for certain conditions.
          2. These condition logic pages will lay out all the data points that feed into the GIRA, show the effect of missing/erroneous data, list out variables used for missingness assessments by QA/QC taskforce, and provide recommended actions for missing data.
          3. These logic pages can be used as a guideline for the sites and the study staff as they are reviewing their GIRA, what variables are needed for the GIRA and what variables are needed specifically for outcomes analysis.
          4. This is part of the confidence in GIRA.
       2. Condition leads need to go through each to determine what are critical elements.
          1. The CC will provide templates/drafts for condition leads.
          2. It should be clear where the data is coming from (survey, etc).
       3. ACTION ITEM: The CARE workgroup and the CC will work on SOP for condition return templates. These SOP will provide guidance on missing data and GIRA generation for each condition.
       4. ACTION ITEM: The CC will work with A fib, CKD, CHD, and PCa leads to update FHx risk estimates for the four conditions that use FHx as a trigger.
       5. Sites are allowed to complete and return GIRA as long as PRS is generated.
          1. Height & weight (BMI) is the critical data element sites should prioritize obtaining from the EHR (within the last 5 years) for GIRA generation.
    3. Edge cases
       1. The sex chromosome abnormalities reported by Invitae have so far been limited to Mayo (3 cases).
       2. Invitae explained that there is a comment that goes on any Invitae diagnostic test reports or proactive test reports that describes the fact that as part of the QC process various sites on the x chromosome are assessed to confirm that sample integrity is maintained throughout the testing process.
          1. Occasionally it is identified that not all of the markers that are expected to be seen are seen, or apparent amplifications are seen.
          2. The assay is not validated to report out duplications on X or absence of X in the amounts expected. Internal policy still requires Invitae to disclose the findings for appropriate clinical correlation, at the discretion of the ordering clinician.
       3. Invitae is unsure if these comments have only affected Mayo results, or if other sites have just not noticed them.
          1. These comments are not currently captured by Invitae data pulls.
          2. The data analyst team at Invitae is working on an update to fix this.
          3. This will allow Invitae to get a number for how many of the cases within eMERGE have this comment included on the report.
    4. Future conversations
       1. There is significant under enrollment in peds.
       2. This has been discussed in applicable groups.
       3. ACTION ITEM: Due to under enrollment of pediatric participants the R2 group should inform the Network regarding the plan and trajectory for pediatric enrollment.
  1. **Data Utilization (GRID) |** 
     1. This workgroup came out of the June SC meeting when the PRS group was sunsetted.
     2. AnVIL Demo sessions and champion users, eMERGE AnVIL Data & Analysis Needs
        1. Robert Carroll introduced himself as the co-PI for the AnVIL Clinical Resource (ACE-IT). He is a co-PI with Matt Lebo.
        2. This is an RFA that Ken Wiley began and Robb Rowley has taken over.
        3. The goal is to have clinical genomics users on AnVIL. eMERGE is a community that uses AnVIL, and the ACE-IT team wants to work closely with eMERGE to ensure the AnVIL Clinical Resource can meet the consortium’s needs.
           1. The next five years will be spent working with the AnVIL team to help drive science forward using AnVIL.
           2. ACE-IT has members at VUMC, the Broad, the AHA, and MGB, along with other sites.
           3. ACE-IT would like to show users what AnVIL can be used for, and how eMERGE data can be used.

ACTION ITEM: Post-docs and researchers who will be using the AnVIL Clinical Resource should reach out to Sophie Forman (sophie.forman@vumc.org) to join the GRID workgroup.

* + - 1. ACR and eMERGE are focused on genomic medicine, so it makes sense to collaborate.
      2. ACE-IT is in the process of planning a series of demo sessions for eMERGE. They are targeting to launch in October.
         1. These will include AnVIL basics and more technical demos.
         2. ACE-IT would like to have a one to two champion users from eMERGE involved in the initial discussions with the AnVIL outreach team and ACE-IT to help scope out what should be covered in the demos.
         3. The next planning meeting will hopefully be next week.

Adam Gordon and Atlas Khan have volunteered as champions.

* + - * 1. Ideas of what to include in the demo include logging on, billing account set up, workspaces, cloning workspaces, using genotyping and phenotypic data for basic analysis and summary, and potentially touching on GWAS.

PRS calculation should be added, along with ensuring SNPs are in the right orientation.

It is important to show eMERGE members how to use AnVIL including WDLs and scratch spaces.

* + - 1. WDLs are workflows on Terra that would allow individuals to run workflows more easily with simple input files.
      2. A good first goal for the group would be to create a WDL that would take a weight file and run a PRS on it.
      3. The demos will be open to the entire network, and it is important to have representation from each site.
      4. It was suggested to have principal components using reference data and projecting onto reference data as well as admixture estimates, because those are information that is used downstream from the PRS.
      5. NHGRI and ACE-IT’s goal is to have everyone across the eMERGE network comfortable using AnVIL, as this is where the eMERGE CC will be sharing the data.
      6. There must be an understanding on how the phenotypic and genetic data are stored and accessed.
      7. The CC is actively working with the data access control. It is a private space shared with the consortium. The access controls around that are enforced on the AnVIL side, but managed by the eMERGE CC.
      8. The genetic data is going to be from the microarray chip and imputed from the Broad.
      9. It is not the premise of this group to delineate data governance policy.
         1. The group can discuss how to use AnVIL in a compliant way.
         2. The eMERGE CC has already been working with AnVIL to create a secure data environment.
      10. AnVIL is learning through working with eMERGE, how AnVIL could work for other consortia.
      11. The workgroup would like to encourage the network to act in AnVIL, instead of individuals just downloading data.
      12. The data in AnVIL now is owned by the network, and will not be compliant with deposition into dbGaP or any other public repository.
      13. Each site has money in their budget set aside for AnVIL.
          1. This could include STRIDE accounts.
      14. This group should inform the network what data is already present on AnVIL and tools to use it.
      15. This group can also discuss the best strategies for freezing the data.
      16. The ACE-IT scope is strictly consortium owned space used for consortium members.
      17. Data storage and the workspaces is being paid for the CC.
          1. Once the publicly shared data is handed over to the AnVIL team in dbGaP, AnVIL takes over sharing that.
      18. There is a set of about 6,500 Broad samples that were part of the August 25th, 2023, data freeze that will go through a joint call set.
          1. ACTION ITEM: Katie Larkin will share a list of what eMERGE 4 data is available in AnVIL to the GRID workgroup (emerge-grid@googlegroups.com).
      19. The cost point has to be appropriate, especially considering site resources that may be at lower or no cost.
      20. Matt Lebo developed an initial onboarding sheet to track progress.
      21. AnVIL is migrating the publicly shared data to Azure within the next year. This opens the door to eMERGE potentially being working on AnVIL on Azure instead of GCP.
    1. Ideas of what to try on the AnVIL
       1. The workgroup made live updates to the ESP GRID report out slides.
       2. There are PRIMED/eMERGE collaboration projects of PRS contextualization, refinement of absolute risk, and developing PRSs that were chosen to remain on the development track.
       3. A suggested project was to discuss how to integrate other large data sets with the data sets in AnVIL.
          1. For example, the reference panel used to develop the PCA was from AoU.
          2. This is a reference population used to develop a PRS.
       4. It was suggested to select one of the PRSes from the development track to use as an example on how to use AnVIL to the network.
       5. In the PRIMED Method Development workgroup, there is a subgroup for genetic access, and a methods workgroup.
       6. Members of the workgroup have been discussing obtaining granular family history information out of MeTree and transforming it into something more parsable and useful. Discussions can occur on how to link that family history with the genomic data, and examining the relationship between PRS percentile and the granular FHH.
          1. This would be a great project for this group as there would be many opportunities for analytics.
          2. The FHH component is differentiating for eMERGE.
       7. The goal of this workgroup is to determine how the network can bring the most out of the data we have, and have it in a place that is usable and well governed.
       8. ACTION ITEM: Please review the GRID roadmap sheet and add any information by **October 6th.**

1. **Investigator Presentation: Low pass sequencing of children with African ancestry and electronic medical records to Cincinnati Children’s Hospital | Leah Kottyan (CCHMC) |** 
   1. Cincinnati Children’s Discover Together Biobank has 100,000+ participants with samples and EHR data. This biobank allows for the ability to access and share internally and externally and preserves the ability to recontact participants with medically actionable results.
   2. Low pass sequencing initiative sequences samples at a depth of 3.8x to detect variants to impute genome wide variants. This results in a much lower cost per sample but calls fewer variants (3.6 vs 5 million) and may reduce confidence in genotype calls.
   3. When processing data, Stitch is used to impute genotypes (Sequencing to Imputation Through Constructing Haplotypes). Using this method, for the cleaned sample, 12,034,432 variants have been called.
   4. 15,832 samples from Black/African American patients were included in an analysis dataset after dropping participants with unclear race, low coverage and call rate.
   5. A graph was shown to describe how a read depth corresponds to percent of the genome coverage with a depth of 5 showing a plateau of genome coverage around 90%.
   6. Quality assessments and tests were performed to verify accuracy and reproducibility of the data.
   7. A graph was shown comparing average read depth and variant call depth. This relationship varies based on race and concordance.
   8. Because the cohort included relatives (parents, siblings and second degree relatives), an analysis strategy was undertaken to identify cases, remove family controls and use PCA matchR to find controls.
   9. Several analysis results were shown showing chromosome locus and phenotype including sickle cell disease, anxiety, and asthma.
   10. Several collaborations with clinicians also looked at some asthma and diabetes projects.
   11. A community engagement group is partnering with Black community members to offer perspectives and focusing on the need to address disparities and structural racism.
   12. The next steps for this analysis include a flagship manuscript, institutional studies, and data sharing with AnVIL.
2. **Return challenges & successes | Josh Cortopassi (UAB) & Brenna Boyd (Columbia) |** 
   1. Brenna Boyd presented on Columbia.
   2. Challenges faced at Columbia include EHR integration, REDCap utilization, and study population.
   3. The current workflow does not include results being sent through Epic due to low usage rates of MyChart. Automated, secure emails are being sent to participants.
   4. Their modified data flow stemmed from their success in recruiting underserved and underrepresented populations.
      1. The majority (60%) are non-white with approximately 30% Spanish speaking.
      2. To assist those with lower health and/or technology literacy, Columbia has helped participants complete surveys.
      3. Columbia has also utilized alternative methods to deliver reports based on participant needs.
   5. Columbia ‘cold calls’ participants with results due to the geographic area could be prohibitive for in-person appointments.
   6. Columbia encourages all sites to have regular server back-ups due to a recent server failure.
   7. Ongoing difficulties include the regular updates to the R4 Portal and lack of an R4 testing environment.
      1. Medical provider changes also present a challenge.
      2. The GIRA being in English only is a challenge for non-English speakers.
   8. Josh Cortopassi presented on UAB.
   9. UAB started returning results in January 2023 with a pause in the summer to address local issues and then resumed returns in August.
   10. Most of the issues UAB faced in the first 4 months were on the REDCap level. Adjustments allowed for a more efficient return process.
   11. Triggering family health history risks by imputing data (age of occurrence of events) added additional steps, which prolonged time for GIRA generation. UAB discovered that data can manually be manipulated in REDCap to trigger a positive instead of manipulating the MeTree survey and creating a different pedigree.
   12. Providers that leave UAB present a challenge for returning results.
       1. Currently, UAB works to find out the replacement physician for a participant. A “point physician” was established to return positive monogenic results if a replacement physician has not been assigned to the participant.
   13. UAB has two participants eligible for medical care coverage.
       1. One participant declined a referral for advanced care. The other participant already had advanced care move forward prior to genetic results being completed.
   14. One piece of provider feedback is that they are notified of a past GIRA upload when a participant arrives for a future clinic visit.
   15. Reactions to cold calls at Columbia have been mostly welcomed and participants have generally taken the calls.
   16. Columbia has started asking participants which one of their physicians they want the GIRA returned to.
   17. CHOP is updating the problem list for high risk participants based on the request of their physicians.
3. **Clinical Operations Update | Katie Larkin (Broad) & Sienna Aguilar (Invitae) |** 
   1. Invitae Update
      1. There have been 4,187 total results reported.
      2. There have been 69 positive results (1.65%).
      3. There have been 6 carrier results (0.14%).
      4. There have been 4,110 negative results (98.16%).
      5. There have been 2 indeterminate results (0.05%).
      6. There has been one result with two findings (MLH1, APOB).
      7. Prevalence of individual gene findings aligns with expectations.
      8. ACTION ITEM: Invitae will provide numbers on sample failure rates.
   2. The Broad Update
      1. There have been 7,623 samples received across all sites.
      2. There have been 6,579 reports generated and delivered to R4.
      3. There have been 6,483 samples staged in AnVIL (there is a joint callset in progress).
      4. 98 test not performed results have been delivered.
      5. The mean TaT for RoR is 28.6 days.
      6. The average call rate for arrays is 98.95%.
      7. Metadata issues
         1. 26 out of 90 sample submissions (batches) so far have had metadata issues.
         2. This adds 18 days on average to turnaround time.
         3. Most common error modes are missing metadata, incorrect metadata, missing R4 orders and duplicate records.
         4. Sites should utilize available resources:
            1. SOP: Sample Management and Submission
            2. R4 SOP
            3. Mock-up of data elements (including mapping to Broad fields)
         5. Sites should make sure to double-check metadata prior to bulk uploads.
         6. Future ClinOps calls will talk about onboarding for new study staff.
         7. ACTION ITEM: The sample chain of custody and getting a hold of staff handling sample plating and shipment can be tedious. The ClinOps group will discuss how to improve this process.
      8. Sample failure reporting
         1. A Test Not Performed (TNP) is issued for the following reasons:
            1. Failed receipt (sample not in Broad tubes, sample uncapped or poorly capped).
            2. Failed sample QC (sample is <10ng/uL).
            3. Failed genotyping (<98% genotyping call rate).
            4. Failed concordance check (the reported sex and genotyping sex do not match and the discrepancy could not be resolved).
         2. R4 is the first line of communication for TNPs.
            1. Sites can use “Broad Results - All sites” report in R4. this can be found under General Use reports.
            2. Sites can also build their own reports.
         3. The Broad Project Managers report on sample failures in two ways.
            1. Sample failures are reported at the monthly ClinOps meetings.
            2. Raquel Jacobs sends out monthly email updates (as applicable) regarding sample failures.
            3. Raquel (rjacobs@broadinstitute.org) is available via email to communicate about sample resubmission or to answer questions.
      9. The PCA plot (looking at samples that are high risk in 1 or more conditions and samples that are high risk in no conditions) continues to look as expected.
      10. The number of samples that are high risk for each condition look as expected.
      11. The number of samples that are high risk for more than one condition do not show anything concerning from a QC perspective.
      12. For samples that are high risk in two conditions the distribution is as expected.
      13. The correlation between conditions remains stable.
      14. Distribution of z-scores for each condition continues to look good.
4. **Workgroup breakout session two**
   1. **Outcomes (Provider Uptake & Outcomes, Phenotyping, & QA/QC Taskforce) |** 
      1. The CC is building a de-identified R4 database that will mirror R4 exactly but will be deidentified. The database will include eMERGE ID and the delta of age at event.
         1. A copy of the de-identified data set in R4 will be uploaded on eMERGE AnVIL alongside genomic data.
         2. Anything with names and text boxes will be stripped. eMERGE ID will allow for sites to go back and identify participants.
      2. Email Megan He (megan.he@vumc.org) with any issues with any reports.
      3. The race/ethnicity groupings/combinations need to be considered.
      4. UAB works with the marketing department, which maps addresses to census tracts to identify medically underserved areas.
         1. The R4 dashboard has a metric for underserved communities. This includes race, ethnicity, gender, sex, low-income, high school, underinsured, and rural. Zip code can also be piped in.
         2. The Google sheet here defines what is on the R4 dashboard and shows which data elements are pulled to provide a count for the underserved categories.
      5. The data freeze and AnVIL data availability timeline can be found here.
         1. The data freeze was kicked off in September and sites have begun uploading data. QC will be done on the files and they should be released by early 2024. In October 2025, a final data freeze will be kicked off.
         2. This will be in the AnVIL\_eMERGE\_PRS\_Arrays workspace where the array raw files are, PRS raw, adjusted, and risk assessment values are along with the imputed single sample vcf files.
      6. The outcomes from EHR data pull redacted elective pregnancy termination, gender-affirming, and suspected child abuse codes. Site will make sure these codes are not included in the data files sent to the CC.
         1. The publication policy has been revised to include sensitive data information.
         2. It is being decided what will be omitted from the de-identified database.
      7. Combinations of ways data can be missing is in the Google doc found here and a Google doc with a withdrawal checklist with loss to followup, withdrawal, and potential excursion criteria can be found here.
         1. It will be important to differentiate investigator or participant withdrawal and indicate why they are withdrawn.
         2. Although the baseline and RoR surveys are marked complete in R4, they may be blank.
         3. If there is genetic information on a participant, it can be used unless they indicate they would like to withdraw.
         4. It is important to consider when to replace participants so that 25,000 GIRAs are returned.
         5. A consort diagram of the study can clarify decisions between loss to follow up and withdrawal.
            1. ACTION ITEM: Nita Limdi will create a draft diagram to begin collecting the information for a consort diagram of the study.
         6. If a sample fails within a year from the date of consent and another is trying to be collected within another 12 months, they are not considered withdrawn. If no sample is collected for a participant, they will be withdrawn.
         7. If someone has a sample and the only data missing is BMI, the site should really put an effort in to get that missing data so at the end there are 25,000 monogenic and phenotypic data (excluding pediatrics).
         8. If there is only a biobank sample and a consent but nothing else, it would be a concern. The GIRA could be generated and returned if the participant is in the EHR however the followup would be different.
            1. UW stated that these participants are consented for eMERGE IV. They are not lost to follow up. Only about 5/600 from the biobank actively withdrew and those would be replaced with other participants.
         9. If no survey data has been provided and that participant cannot be reached, it is assumed they have withdrawn without telling the network so. If they have formally signed a consent form, they are officially consented.
         10. Sites should be opening the surveys to look for missingness to confirm it has been completed before sending the clin cards.
         11. As long as there is PRS, the GIRA can be generated. BOADICEA needs family history to run for women to receive a GIRA integrated score for breast cancer.
   2. **Genomic Data Integration (CIRT) |** 
      1. CIRT is an acronym for Clinical Implementation and Return Technologies and is combined from the EHRI workgroup and CDS subgroup.
      2. The phenotyping workgroup is currently discussing how referral data will be extracted.
         1. Referral data is the data for referrals to specialty clinics.
         2. HL7 is working on how to standardize this type of data and may be useful in the future.
         3. The likelihood of this happening is low at this time.
      3. The provider survey was developed and sIRB approved to determine how the HCP perceive the high risk GIRA report as useful for clinical care.
         1. An invitation to complete the survey is emailed to the provider of record.
         2. Sites need to work out how the survey will be distributed and if the GIRA was opened by the provider.
         3. Determining who the first high risk GIRA for a provider is difficult. REDCap is not linked to a provider database.
         4. Another difficulty is relying on participants to list the provider name in a consistent manner.
         5. A participant-specific link will be created in R4 and can be sent to the provider listed in the pre-screen survey.
         6. There was some discussion on using the National Provider Identifier to determine the correct provider to send the survey and determined that using the NPI is too cumbersome.
         7. Care should be taken to only notify the provider the site identifies as appropriate that the GIRA is available in order to try to minimize specialists from opening the report.
         8. Many sites can track that an in-basket message is opened but most likely cannot tell who opened the message.
            1. Column J in this spreadsheet outlines site capabilities.
      4. For future publications, the spreadsheet has tracked several data points that can be variables within the manuscript.
         1. The topics within will need to be updated by each site when manuscript development starts.
         2. A code of Z15.89 (Genetic susceptibility to other disease) has been used in the problem list at some sites.
         3. Manuscript topics being focused on currently are for a lessons learned paper and integration of a clinical PRS into the EHR.
            1. The workgroup goal is to have a manuscript concept sheet submitted before 2024.
         4. ACTION ITEM: The CIRT co-chairs will develop a timeline to plan manuscript development for the workgroup.

**eMERGE Day 2: Thursday, September 21st, 2023**

1. **Opening remarks and comments from ESP Chair | Robb Rowley (NIH/NHGRI) & Dan Rader (University of Pennsylvania)**
   1. The ESP recognizes network progress, particularly in recruitment and GIRA return.
2. **Network overview: Priorities, goals, progress and ESP recommendations | Rex Chisholm (SC Chair, Northwestern)** 
   1. The main ESP meeting goals:
      1. Recognize progress in recruitment, sample processing, and return of results.
      2. Discuss approaches to missing data and withdrawals.
      3. Introduce new workgroups and scopes of work.
      4. Provide updates about timeline, pending decisions, and network challenges to the ESP.
   2. Participant recruitment across the network is going well.
      1. As of September 19, 17,118 participants have been recruited throughout the network.
         1. 10% of participants are pediatric participants.
         2. 49% of participants belong to racial and ethnic minorities.
         3. 66% of participants are female at birth, and 34% of participants are male at birth.
   3. The projected completion of recruitment is February 2024, in advance of the April 2024 target.
      1. Recruitment may slow as sites focus on completing enrollment for pediatric and underrepresented minority participants.
   4. The network has accomplished many milestones since April 2023:
      1. 68% of the recruitment goal has been achieved.
      2. Nine out of ten sites are returning GIRAs. Vanderbilt University is awaiting results from partner sites to generate a GIRA.
      3. 6 month post-RoR survey, provider survey, and family history surveys have been developed.
      4. The publication policy has been updated to highlight sensitive and stigmatizing language, in order to protect participants from legal recourse while maintaining scientific accuracy.
         1. Pregnancy termination, child abuse, and stigmatizing language codes from EHR data were redacted. Child abuse codes were redacted to prevent any issues for sites who are mandatory reporters. In order to allow the study of substance abuse, drug abuse codes were not redacted.
         2. The network is currently using traditional race descriptors. Acknowledging the self-reported nature of participant race in publications will help alignment with the recent report from the National Academies regarding population descriptors.
      5. Data missingness is being assessed by the QA/QC taskforce.
      6. Minimum sample size for analysis is being assessed by the outcomes workgroup.
      7. An EHR outcomes data freeze has been initiated and is projected to be released to the network in early 2024. The data freeze will be used to assess data missingness and the ability to do robust analyses.
   5. We are currently in early Year 4 within the project timeline.
      1. The EHR data freeze was initiated in September 2023, and multisample VCF should be available in AnVIL by the middle of the year.
      2. A de-identified R4 database will be released to the network in late 2023.
      3. The goals for completing recruitment and GIRA returns will allow Year 6 to be reserved for data analysis.
   6. A priority of the network is to ensure that sites are confident about providing accurate, correct GIRAs for participants.
      1. The sites and the CC have worked collaboratively to address issues with GIRA issues, such as logic and readability.
      2. To help ensure that the correct reports are returned to participants, Invitae and Broad samples are identified using name, study ID, and date of birth. Invitae and Broad results are returned with the identifying information and is then parsed to input into the R4 portal.
      3. Of 2145 GIRAs generated total, 1745 have been returned.
         1. The current proportion of pediatric vs. adult GIRAs returned is higher than the expected final proportion.
         2. 34.5% of participants are at high risk for at least one condition, which is a higher proportion than initially predicted. During yesterday’s sessions, there was a robust discussion regarding genetic vs environmental contribution to risk for all studied conditions. There is missing family history data for participants with high risk GIRAs (<30%). Even if a MeTree is marked as complete, it does not indicate if all fields were completed. Not all conditions use all three triggers (monogenic, PRS, family history).
   7. The workgroups have evolved to address changing needs of the network:
      1. The PRS workgroup was sunsetted and is working on the group’s final paper.
      2. The EHRI and CDS workgroups combined to form the Clinical Implementation and Return Technologies (CIRT) workgroup.
      3. The QA/QC task force and genomic risk & innovation discovery (GRID) workgroups were created.
   8. Multiple manuscripts have been submitted or are in preparation.
   9. The upcoming goals are:
      1. Assess data quality of the 6 month post-RoR survey.
      2. Compare rates of actual proportion of high risk GIRAs compared to the expected proportion.
      3. Ensure that recruitment numbers and data are sufficient for analysis.
      4. Complete the data freeze onto the AnVIL.
   10. The main study challenges surround:
       1. Efficient GIRA generation and return to meet network goals.
       2. Minimizing passive withdrawals (loss to follow up) and missing data.
       3. Deployment and use of the provider survey.
       4. AnVIL use for prospective and discovery analysis.
3. **Recruitment | Digna Velez Edwards (VUMC), Ingrid Holm (BCH), & Wendy Chung (Columbia) |** 
   1. The current timeline is updated includes enrollment ending in June 2024 and returns ending in October 2024.
   2. The prospective cohort timeline has participants receiving the post-RoR survey 6-12 months after enrollment.
   3. The use of automated REDCap emails throughout enrollment assists in participant retention.
   4. As of September 8, a total of 16,730 participants have been enrolled with 2,025 GIRA generated.
      1. Approximately 50% of those enrolled are from underrepresented races.
      2. The vast majority of participants are female at birth (66%).
      3. Overall, almost 70% of participants have completed family history data.
      4. There are 7,700 participants that currently have samples recorded in the R4 Portal.
   5. Current challenges include:
      1. Moving forward with sample processing through GIRA returns.
      2. Handling passive withdrawals.
      3. Participant retention.
      4. Communication challenges regarding results.
4. **Comprehensive Risk Assessment & Return (CARE) | Gail Jarvik (UW), Iftikhar Kullo (Mayo), & Cindy Prows (CCHMC)** 
   1. Return Progress
      1. There has been excellent progress in enrollment. The Network is doing well sending samples to the Broad.
   2. GIRA Generation
      1. All sites have conducted the GIRA QC process and generated GIRA reports. Nine sites have returned results.
      2. Breast cancer return required changes. These have now been finalized.
      3. The GIRA review had been initially time consuming, but is getting better. There is still a need to make the process faster. There are SOPs for each condition. Phenotype leads are available to address questions related to GIRA by email/phone.
      4. Edge cases are discussed at CARE meetings (known monogenic results, decedents, non-responders, etc have been discussed.)
      5. There have been incidental findings. Invitae is returning sex chromosome abnormalities. The CARE workgroup has discussed how to handle these and similar situations.
   3. Return Process: CDS
      1. RoR education documents and talking points sheet for study staff returning high risk results across sites are being standardized.
         1. This is not currently standardized for return of not high risk results. This might be something to develop as well.
      2. There is an EHR encounter note template with EPIC smart phrase inputs under development for each disease.
      3. The placement of the Invitae report, and in some sites the Broad report, may occur before high risk GIRA face-to-face disclosure with participants in order to be in compliance with the Cures Act. There had been previous discussion about potential issues with Mendelian reports going in before the GIRA generation. This has not been a problem because those participants are in a separate analysis group.
   4. Tests Recommended for Medical Cost Coverage (MCC)
      1. A list of phenotypes and recommended tests/procedures has been developed.
      2. The Network is reviewing experiences related to MCC at sites.
      3. Laura Rasmussen-Torvik is leading the writing of a medical costs manuscript.
   5. Pediatrics
      1. Recruitment of pediatric participants is a bit behind.
      2. There are challenges with pediatric recruitment, including the need to have a parent at consent and RoR.
      3. Disagreement between a parent and adolescent has been brought up as a potential future issue.
      4. The Rescue Survey will be used as the primary source for family health history in pediatric participants.
   6. Family History
      1. The Rescue Survey is now live and available for online data entry by participants. The Rescue Survey is still being tweaked for breast cancer.
      2. A disclaimer is in the process of being added to all GIRA regarding missing family history.
   7. Missing Data
      1. The Network is in the process of defining the categories of withdrawals. Active withdrawals are participants who no longer want to participate and request to be withdrawn from the study. If there is no sample after 1 year it is considered a loss to follow up.
      2. The Network is making every effort to obtain any missing data by following protocols for participant contact and obtaining missing data from the EHR based on Network guidance.
      3. The decision was made to generate and return a GIRA for any participants who consent and provide a sample, even if the baseline survey was not filled out.
      4. The QA/QC group was formed to review the extent of missingness, how to limit missingness, as well as potential imputation options.
   8. Discussion:
      1. The Network is no longer using the term “passive withdrawal”.
      2. Sites have seen differential rates of missingness based on race and ethnicity. MGB is working on a patient navigator model to help with this disparity. For example, Spanish speaking research assistants are being assigned to Spanish speaking participants and materials written in Spanish are being used.
      3. As of August 2023 sites are submitting monthly Recruitment and Return Projections monthly. This will be a useful tool to help the Network meet targets.
5. **Provider Uptake & Outcomes |Nita Limdi (UAB) & Dave Veenstra (UW) |** 
   1. The study framework will include 25,000 (20,000 adults and 5,000 pediatrics) GIRAs either high risk or not high risk. High risk reports will have recommendations and not high risk reports will not.
   2. The group will measure if the providers read the reports and if they ordered tests across the high risk and not high risk categories. Phenotype specific analysis will also be done on additional testing leading to new diagnosis.
   3. Time is set to 0 at the point of GIRA return and up to 12 months EHR and participant surveys will be used to assess uptake of recommendations and new diagnosis/risk reducing interventions.
   4. Over the last year, the outcomes workgroup has been working with other workgroups to identify needs from the EHR, ensure data quality, consistency, and missingness, and ensure documentation of ROR methodology to harmonize streamline approaches for GIRA return and communication. The impact of family history data is also being evaluated.
   5. It was expected that 2% of participants would have a high risk PRS for chronic kidney disease and right now it is between 3 and 4 percent.
   6. Actively withdrawn patients will be tracked and recruitment will be increased and the goal will be 25,000 DNA samples and returned GIRAs.
   7. Next steps for the outcomes group will include the interim data freeze and deploying 6-month post RoR participant and provider surveys. Additionally, the group will assess participants with multiple high risk conditions.
      1. There was no pre RoR survey for providers. In the post RoR survey for providers, the group is asking if the GIRA was informative. The plan is to have the providers fill out the survey only once after the first high risk GIRA is received.
      2. Surveying both high risk and not high risk return providers could be a recommendation however some sites are recruiting at multiple sites so the number of providers will be different and much higher.
   8. If a participant already has a diagnosis after receiving a high risk GIRA, they will be excluded from analyses.
   9. All sites are entering the GIRA into the EHR even just as a pdf. Time between GIRA generation and EHR integration/GIRA return differ between sites but some are automatic after they are entered into REDCap.
6. **Phenotyping and Outcomes data collection | Jennifer Pacheco (NU) |** 
   1. The current progress of the phenotyping workgroup includes a completed review of outcomes EHR elements, redacted codes to protect participant data, and having begun data quality assurance for GIRA and outcome data. The CC will quality check all data files from sites to ensure the redacted codes were not included in the files.
   2. The group is working on deciding how to pull referral data from the EHR which is challenging. Codes for referrals differ across sites so keywords will be used and each condition lead may need to consult with sites on clinic codes and keywords.
   3. The extraction of the GIRA clinical elements is currently ongoing and sites are sharing ebay practices for quality assurance across the network.
   4. The outcomes from EHR data pull is also currently ongoing. A larger list of meds and labs was reviewed and finalized. Visit data is a newer category being pulled as a proxy for referrals.
      1. The data dictionary will be revised as needed. For example, age to three decimal places is being captured in place of event date so if that is not sufficient it will be revisited.
      2. If sites develop SQL code that other sites can use, it will be shared to facilitate standard and easier/faster implementation.
      3. Managing the data for vulnerable groups is another discussion point within the workgroup. The phenotyping workgroup was finding the balance between protecting participants but also making sure research is able to be conducted.
      4. DUAs are also in place to allow for de-identified data sharing.
      5. Additional workgroup work will include determining if missing variables for the GIRA need to be found in the EHR. For example, cutoffs for BMI need to be determined for less necessary chart review.
      6. Next steps for the phenotyping workgroup includes continuing to coordinate with the QZ/QC task force and the outcomes workgroup to ensure data quality and completeness and revise and refine the data being extracted from the EHR.
      7. The use of care everywhere is not permitted for research at many sites so using it to pull missing data may not be an option.
7. **QA/QC Taskforce | Jennifer Pacheco (NU) & Lisa Martin (CCHMC) |** 
   1. Sites can run other sites’ R4 reports to show their own site data.
   2. The QA/QC task force has been sharing reports in order to be more efficient and to learn from other sites. Not all sites have been tracking the same things. Sharing reports allows sites to gain insight on what can be tracked and measured. For example, David Crosslin and Jose Irizarry created an R script that connected to R4 which generated statistics on missingnes. This can be run at all sites.
   3. The task force has realized that underserved groups must be defined, and would like to determine if those groups have disproportionate amounts of missing data. Defining cohort groupings for reporting will improve inclusion and diversity.
   4. The task force has been working on an Implementation Guide with intent to distribute through the network.
   5. There have been discussions of classifying different types of withdrawal, with the recent decision during this meeting to change ‘passive withdrawal’ to ‘lost to follow up’.The task force collaborated with the R2 workgroup for this task.
   6. The task force has worked with the R4 developers to expand withdrawal tracking in R4. A future goal is to allow for potential loss to follow up reports to be run in R4.
   7. The implementation guide has been used to define key metrics.
   8. It was proposed to add a handful of questions to the rescue survey to obtain critical missing variables.
   9. Sample tracking is performed at individual sites. It would be helpful to share R code or SAS code across sites. Currently, the task force has been sharing excel sheets that are used to import data into.
   10. ACTION ITEM: The Task Force will confirm with sites that all critical data is located on a database or a file server that is backed up.
   11. The task force has been focusing on items that would prevent the network from generating a GIRA. This includes name mismatch and duplicate participants.
       1. Name mismatch is defined as participants with a different name in their medical record than the one provided in the study. For example, a participant using a hyphen in their last name in the medical record but not in the study.
   12. Different sites have been presenting to the task force on how they are tracking data.
   13. The task force has at least one representative from each site and from each work group.
       1. ACTION ITEM: Data wranglers and data analysts interested in joining the QA/QC task force should reach out to Sophie Forman (sophie.forman@vumc.org).
   14. The task force has not been able to obtain analytics on participants lost to follow up due to most participants being enrolled within the year.
   15. If participants move out of the site and provide a mailing address, the study could still provide them with their results and send them a post RoR survey.
   16. It is important to track separately the percent of participants who enrolled and never provided a biospecimen, and the participants who provided a biospecimen but were lost to follow up. All participants who provided a sample will receive a GIRA.
8. **Genomic Risk Innovation and Discovery (GRID)| Adam Gordon (NU) & Matt Lebo (MGB) |** 
   1. The GRID workgroup was formed a few months ago.
   2. The overall scope and goal of the group is to be the eMERGE home for cloud-based data analysis and resources.
   3. This is centered on the AnVIL analysis environment.
   4. The GRID workgroup has been in discussions with the AnVIL Clinical Resource, which is a newly funded grant through the NHGRI to create an AnVIL space that can hold PHI and more detailed clinical information. Matt Lebo is a co-PI on ACR.
   5. There are demos and training planned, both for general use and for more specific eMERGE use.
   6. The GRID workgroup will also be the home for the PRIMED collaborations, initiated at the eMERGE/PRIMED joint meeting in February 2023.
   7. The group plans to choose a developmental PRS to work through as a group as a model example in AnVIL.
   8. The group has been in conversations with MeTree on how to use the MeTree data on a more granular level that can be de-identified and usable.
      1. ACTION ITEM: Junior researchers interested in joining the Genomic Risk Innovation & Discovery workgroup should email Sophie Forman (sophie.forman@vumc.org).
   9. The group will aim to provide guidance on pain points associated with using AnVIL.
   10. eMERGE I-III, eMERGEseq and pgrn seq data are on AnVIL.
   11. The phenotypic data are on AnVIL.
9. **CIRT | Eta Berner (UAB), Emma Perez (MGB), & Bob Freimuth (Mayo) |** 
   1. CIRT (Clinical Implementation and Return Technologies) is the combination of the EHRI workgroup and the CDS subgroup.
   2. The main focus of this workgroup is to brainstorm solutions to common return/implementation issues, harmonizing workflows where possible, and manuscript development.
   3. Manuscript considerations for lessons learned with integration implementations and integrating PRS into the EHR are being discussed.
   4. Current workflows for pulling referrals and distribution of the provider survey are current workgroup topics.
   5. The ESP started discussion regarding sites putting the report in the precision medicine module in Epic or if the PDFs being scanned into the EHR.
      1. This is site dependent based on local regulations. Some sites have to place the PDFs in the media tab while others might be placing the reports in the lab section.
      2. Invitae clinical reports are placed in the EHR independently from the GIRA. Most sites are putting the GIRA in the media tab.
      3. Caution in placing the genetic results in the problem list is needed as to not overload this list and to mitigate confusion as to what is standard practice regarding genetics with the provider.
10. **Input/Feedback from the ESP, general discussion**
    1. The ESP agrees that the eMERGE network has made a lot of progress, particularly in the areas of recruitment and return.
       1. The eMERGE project is a very exciting and difficult project, with the potential to be transformational in the field.
    2. The provider survey will provide important qualitative information about the provider experience with the GIRA return process.
       1. The network should strongly consider sending the provider survey to providers who do not receive a high risk GIRA report.
       2. Sampling providers who do not receive any high risk GIRA reports will provide a different perspective on the eMERGE study.
       3. The timing of the provider survey is crucial and should be discussed further.
          1. Currently, the provider survey will be sent after the first positive report.
          2. Sending the survey multiple times will overload the providers.
    3. Manual chart review, while time and labor intensive, is an important step of data quality control.
       1. The ESP suggests creating a structured way to manually review a sample of the data to confirm that data which is pulled automatically is correct.
          1. It will be important to share with the community that the data was confirmed by manual review.
          2. While comparing survey responses to medical charts, it will be interesting to see any discrepancies between the data.
    4. The iterations of the GIRA report resulted in a great product, but there are a few suggestions to improve participant understanding.
       1. In a negative GIRA report, the language surrounding the assessment of the breast cancer condition on page 1, 2, and 3 is confusing as a naive clinician or patient.
       2. There is an emphasis on collecting data surrounding how high risk reports impact outcomes, but data should also be collected on how not high risk reports will impact participants.
          1. Possible impacts could include participants ignoring normal care.
       3. There was previous discussion regarding clearly stating the implications on the utility and interpretability of the report when a GIRA report is generated using incomplete information.
       4. The FAQ on the negative GIRA report includes the question: “What does it mean if I am not high risk?”
    5. Standardizing the method by which providers are informed that a GIRA has been added to a participant’s EHR may increase the rate of utilization by the provider.
       1. Currently, sites are using different methods including email and direct message in the EHR.
    6. The estimated time of 3 months between the end of enrollment and end of GIRA return may be ambitious.
       1. Enrollment indicates that the participant has consented but time will be needed to obtain participant biospecimens.
       2. While 22,000 GIRAs remain to be returned, 70% of these reports will be not high risk which will require little effort by the sites to return.
       3. Samples should be submitted ASAP, to account for unexpected delays with sample processing.
11. **Closing Remarks | Rex Chisholm (SC Chair, Northwestern)** 
    1. An important follow up discussion for the network will be to ensure consistency in the use of “not high risk” versus “below study threshold” language.
    2. The guidance thatsites can overshoot enrollment by ~100 individuals to ensure enough participants in the study to account for withdrawals and loss to follow up was reiterated.
    3. The eMERGE dataset will create a very unique final dataset for research.

**Action Items:**

1. **Recruitment & Return (CARE, R2/ELSI/sIRB, & QA/QC Taskforce)**
   1. The QA/QC, R2/sIRB, CARE, and Outcomes workgroups will define active withdrawal, administrative withdrawal, and loss to follow up definitions. The CC will update terminology for withdrawal on R4.
   2. CARE & QA/QC will complete the missing data combos matrix and provide overall guidance to the network regarding tolerance for missingness. The groups will also provide recommendations for when participant withdrawal should be considered.
   3. The CARE workgroup and the CC will work on SOP for condition return templates. These SOP will provide guidance on missing data and GIRA generation for each condition.
   4. The CC will work with A fib, CKD, CHD, and PCa leads to update FHx risk estimates for the four conditions that use FHx as a trigger.
   5. Due to under enrollment of pediatric participants the R2 group should inform the Network regarding the plan and trajectory for pediatric enrollment.
2. **Data Utilization (GRID)**
   1. Post-docs and researchers who will be using the ACR should reach out to Sophie Forman (sophie.forman@vumc.org) to join the GRID workgroup.
   2. Katie Larkin will share a list of what eMERGE 4 data is available in AnVIL to the GRID workgroup (emerge-grid@googlegroups.com).
   3. Please review the GRID roadmap sheet and add any information by **October 6th.**
3. **Breast Cancer group**
   1. The breast cancer group will generate talking points on explaining lifetime and residual breast cancer risk during return.
4. **Outcomes (Provider Uptake & Outcomes, Phenotyping, & QA/QC Taskforce**
   1. Nita Limdi will create a draft diagram to begin collecting the information for a consort diagram of the study.
5. **CIRT**
   1. The CIRT co-chairs will develop a timeline to plan manuscript development for the workgroup.
6. **ClinOps**
   1. Invitae will provide numbers on sample failure rates.
   2. The sample chain of custody and getting a hold of staff handling sample plating and shipment can be tedious. The ClinOps group will discuss how to improve this process.
7. **QA/QC Task Force**
   1. The Task Force will confirm with sites that all critical data is located on a database or a file server that is backed up.
   2. Data wranglers interested in joining the QA/QC task force should reach out to Sophie Forman (sophie.forman@vumc.org).
8. **GRID**
   1. Junior researchers interested in joining the Genomic Risk Innovation & Discovery workgroup should email Sophie Forman (sophie.forman@vumc.org).

**Decisions Made:**

* Sites can overshoot enrollment by ~100 individuals to ensure enough participants in the study to account for withdrawals and loss to follow up.
* Sites should make every effort to obtain missing survey data. As long as the participant has provided a genetic sample the site can generate and return GIRA.
  + Sites are encouraged to obtain missing survey data from the EHR if feasible, and this will be documented.
  + If a participant has never provided a sample, or if they do not complete the baseline surveys and no sample is available, sites can withdraw and replace the participant
* There will need to be a distinction made from an outcomes perspective.
  + Certain missing data may exclude participants from analyses even though GIRA was returned.

**Official ESP Recommendations**

Meeting Summary

eMERGE Network- External Scientific Panel and Steering Committee

*Executive Session – 09/21/2023*

|  |  |  |  |
| --- | --- | --- | --- |
| **ESP** | **Dan Rader,** University of Pennsylvania **– Chair**  **Kimberly Doheny,** Johns Hopkins University  **Stanley Huff**, Intermountain Healthcare  **Janina Jeff**, Illumina\*  **Brendan Lee**, Baylor College of Medicine\*  **Lisa Parker**, University of Pittsburgh  **John Witte**, Stanford University | **NHGRI** | **Teri Manolio**  **Jahnavi Narula**  **Robb Rowley**  **Rene Sterling** |

\* Was not able to attend the ESP meeting

The External Scientific Panel (ESP) met with NHGRI program staff during the executive sessions of the eMERGE Steering Committee/ESP Meeting held on September 20-21, 2023. The ESP was extremely impressed with the progress the Network has made, including with the Network’s thoughtful discussion of challenging topics, recruitment, and return of Genome-Informed Risk Assessment (GIRA) reports. The ESP provided observations and recommendations about several aspects of the research to help move efforts forward. Initial feedback is elaborated upon in the summary below.

**The Network should discuss sending the post-return-of-results survey to all providers.**

Currently, providers get a post-return-of-results (post-RoR) survey after they return their first high-risk GIRA report. The ESP recommended that the post-RoR survey be sent to all providers, not just those who have returned a high-risk GIRA report, as understanding the importance of the GIRA report to providers who have returned GIRA reports not identified as high risk per study criteria is valuable information. Additionally, having a larger and broader survey group could help determine the effect of completing the survey on using the intervention. Completing the survey could give the impression that providers’ behavior is being monitored, inadvertently encouraging them to use the GIRA report in a different manner. The ESP noted that sending the post-RoR survey to a subset of providers, rather than all providers, could also work; however, the subset need not be limited to those who have seen at least one high-risk patient.

**The Network should consider adjusting the GIRA report’s language, especially in the context of incomplete or missing data.**

The ESP was impressed with the GIRA report, but they pointed out that there might be some areas of confusion for a naïve clinician or patient. For example, in a negative GIRA report, it’s possible to have the statement “Note that breast cancer cannot be assessed due to a personal history of breast, ovarian, or pancreatic cancer.” The ESP felt that this becomes confusing when paired with the Frequently Asked Questions on the next page, which states that breast cancer was assessed in the study. Additionally, it may be beneficial to put language on the first page of the GIRA report or in the cover letter if conditions were not able to be assessed due to missing data. Finally, the ESP asked that the Network take another look at language around being “high risk” and “not high risk,” and suggested that the latter be replaced with “does not meet study thresholds.”

**The Network should perform manual chart review as part of the ongoing Quality Control efforts.**

The ESP was impressed by the efforts of the QA/QC task force. They noted that it would be helpful for the Network and for the community to perform a limited set of manual chart reviews for both GIRA generation, as the Network is already doing, and for outcomes analysis to make sure what is being pulled electronically is valid. Although it is a time- and labor-intensive process, it would be beneficial for the community to know what percent of the charts were sampled and whether there is any discordance.

**The Network should monitor both escalation and de-escalation of care.**

The ESP noted that the Network should make sure to understand that providers and patients may increase or reduce care after they receive a GIRA report, depending on the GIRA report. For example, a provider who receives a “high risk” GIRA report may decide to increase the patient’s statin dose, but if they receive a GIRA report that does not meet study thresholds for high risk, they might decrease a dose of statin. Both of these outcomes should be captured and are important to record and assess.

**The Network should make sure to record heterogeneity across sites and standardize where possible.**

Given the scale of the Network, heterogeneity is inevitable. For example, different sites have different methods of letting providers know that the GIRA report has been uploaded into the EHR. The ESP suggested that standardizing how providers are informed that the result is ready would subsequently standardize awareness of the GIRA report during a patient encounter. When this is not possible, creating a document that summarizes the heterogeneity is essential for outcomes analysis.

**The Network should look at timelines to make sure there is enough time for GIRA return and make sure to get samples out as soon as possible.**

The ESP was extremely impressed with the Network’s progress in terms of recruitment. However, they raised a concern regarding the three-month timeframe between enrollment completion and return of the GIRA report, suggesting a second review of this timeline be considered. The ESP also emphasized the importance of sample collection and shipment from sites, acknowledging that delays may occur. Additionally, they agreed that sites could consider slightly overshooting their goals by collecting samples for another 100 participants, to allow for some drop-out in participants ultimately providing samples. They noted there could be some flexibility in this number as more information becomes available.

ESP Recommendations for the Network:

1. The Network should discuss sending the provider survey to all providers that receive both high risk and not high risk GIRA reports.

2. The Network should consider adjusting and moving the GIRA report’s language, especially in the context of incomplete or missing data.

3. The Network should make sure to perform manual chart review as part of the ongoing Quality Control efforts.

4. The Network should make sure to monitor both an escalation and de-escalation of care among recipients of a GIRA report.

5. The Network should make sure to record and keep a summary of heterogeneity across sites and standardize where possible.

6. The Network should look at timelines to make sure there is enough time for GIRA return and make sure to get samples out as soon as possible.