Breakout Session 1:

1. What data standards and/or terminology standards are you currently using for data query and aggregation? What data fields or properties are most important to represent the kind of data you manage?
2. Which data elements / fields are most likely to be important across CRDC nodes?
3. What models are you currently using? What is working? What are the deficiencies?

Breakout Session 2:

1. What are the biggest barriers to submitting data (which would meet the discovery needs of our first breakout)? For example, you can't find the term you need; one of your variables does have the necessary field
2. How would you combine automated methods with manual processes for efficiency and efficacy? How could these combined methods be used for data QA and/or prompting the user for additional information?
3. What example tools, methods, approaches have worked best in your experience and why?

Breakout Session 3:
1. Given what we've already discussed about data standards, data models and criteria for submission and search, what queries do you want to answer across different data types / CRDC nodes? When you enter something into the search tool, what do you expect to get back?

    For example,
    When patients are treated with chemotherapy and become resistant, what parts of BRCA2 change? (i.e what BRCA mutations are responsible for resistance to chemotherapy?)

    Identify patient, cohort, dataset for later analytics

Use Case : Proteogenomic Integration

Find all the [Program Project Study] in [PDC GDC IDC] that have [Proteomic Genomic Imaging] data

Get [RNA-Seq BAM file Variant file RNA-Seq Junction file] for a [Case/s Sample/s Aliqot/s]

Get [Copy Number Data Expression Data] for a [Project Study]

I know there is genomic information available for the PDC proteomic study I am interested in. How can I seamlessly integrate the somatic/germline variant, RNA-seq predicted junctions and fusion, etc to create a custom sample specific database to search against?

Curated Knowledge Bases & Data Repositories

Clinical
- Clinical Data Collaboratives
- Clinical Trial Registries
- Clinical Information Systems Data
- Clinical Guidelines
- Clinical Research Information Systems Data
- Biomarker Knowledge bases
- Biomedical Literature

Preclinical
- Research Data Collaboratives
- Drug Knowledge Bases
- Genotype Phenotype Knowledge Bases

Research Tools & Services

Systems
- Basic Research Information Systems
- Data Quality Assurance
- Knowledge/data-driven scientific interpretation
- Data Integration
- Data Visualization

Methods
- Data Extraction & Transformation
- Data Standards, APIs, Vocabularies, Schemas
- Data Collection

Clinical Decision Support Tools & Services

Systems
- Clinical Information Systems
- Data Quality Assurance
- Knowledge/data-driven Clinical Interpretation
- Data Integration
- Data Visualization

Methods
- Data Standards, APIs, Vocabularies, Schemas
- Data Extraction & Transformation
- Data Collection

Patient-Centered Support Tools & Services

Public
- Registries
- Patient Advocacy & Research Groups
- Patient Reported Outcomes
- Automated & accessible Consent Tools
- Patient History & Phenotyping Tools

Consented
- EHR-based patient data sharing
- Multi-Organization IRBs
- Payor Claims Data