ALLIANCE 1-SLIDE BREAKOUT SUMMARY SLIDE: **TRUTHING** SESSION: **1**

Key elements, next steps, timeline

1. Data access is the biggest burden of the scientific studies needed to evaluate truthing
2. Slide truth can be collected on multiple levels (e.g. orthogonal testing, pathologist diagnosis, patient outcomes)
3. Data agreements are the biggest pain point in this step, which we will address.

Pros for Patient, Clinical, R&D, and regulatory

1. This is the most critical factor necessary for any action to actually take place
2. Existing repositories could drive access to some company's existing therapeutics
3. Clarify pipelines from data in/data out

Standard legal template for data sharing

Concerns for patients, clinical, R&D, and regulatory

1. Timeline, risk, ability to recall in the future for subsequent studies
2. HIPAA & De-Identification
3. Metadata is variable

Implications and efforts

1. Developing the expansiveness of the data archive (inclusion & exclusion criteria) will be difficult
2. Validation of data creation: collection to treatment, to prep, all the way to algorithm. This is all one tailored "soup to nuts" workflow, and each company's complete flow is unique

ALLIANCE 1-SLIDE BREAKOUT SUMMARY SLIDE: **TRUTHING** SESSION: **2**

Key elements, next steps, timeline

1. Dataset from multiple sites of both glass and digital
2. Try to use a dataset already created, then add truth by pathologist
3. Address generalizability:
   - Test generalizability a single algorithm to multiple use sites or datasets across time
   - Test generalizability data to multiple use cases
4. Creating environments where people can explore these questions

Pros for Patient, Clinical, R&D, and regulatory

1. Can use this project to build incentives for building feedback infrastructure and correct data collection methods
2. Efficient, scalable testing & method development of algorithms
3. Providing examples of what FDA considers to be a good ground truth

MDDT Validation Dataset

Concerns for patients, clinical, R&D, and regulatory

1. "Generalizability" and other terms are still convoluted in their definitions
2. No good plan for statistical analysis: could be solved by a statistical challenge competition
3. Explainability of algorithms is still imperfect

Implications and efforts

1. Continuous addition to the dataset could be good for longevity
2. Need to plan for "future-proofing"
3. Considerations of algorithm ethics may play into the outcomes of this project
4. Sources of bias within data collection