



## Data Sharing and Transparency Roundtable Meeting

November 3, 2015

- I. Upcoming MRCT Center Conferences and Events
- II. Review of Overarching Project Goals
- III. Recap of Progress to Date
- IV. New Progress: Use Case Documents
  - Detail of 2 use cases
- V. New Progress: IT Policy Issues
- VI. Discussion and Questions

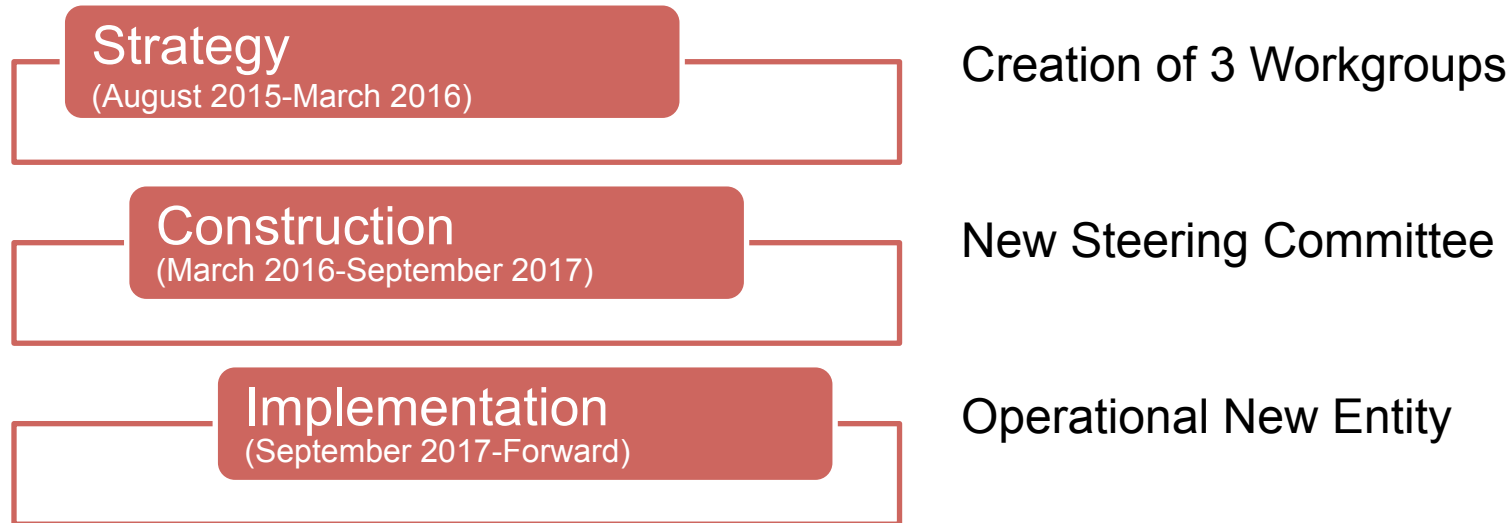
## I. Annual Conference: December 14<sup>th</sup>-16<sup>th</sup>, 2015

- A. December 14<sup>th</sup>: Post-Trial Responsibilities- Open Workgroup Meeting
- B. December 15<sup>th</sup> 2015: MRCT Center Annual Meeting – Featuring Data Sharing and Transparency with Keynote Speaker, Bernie Lo**
- C. December 16<sup>th</sup> 2015: Individual Participant Return of Results Open Workgroup Launch Meeting

## II. MRCT Center Data Sharing Conference- March 21-22, 2016 at the Wellcome Trust in London, UK

**To create a blueprint for the creation of a new, not-for-profit organization whose goal is to create, direct, implement and oversee a sustainable data-sharing platform**

## **3 Phases of Realization:**



Partnering with Wellcome Trust, IOM, Laura and John Arnold Foundation and Deloitte Consulting to ensure collaborative, sustainable, unified approach to common platform and portal

## MRCT Governance Working Group

- Created Charter with vision + mission statement
- Developed steering committee specifications

## Information Technology (IT) Working Group

- Designated commissioned report specifications
- Developed platform use cases
- Identified data sharing policy issues

## Business Models Working Group

- Examined existing sustainable business models through an environmental scan

## Update on Progress of Data Sharing Project

### Recent Progress:

- **Use Case Definition** – a list of action and event steps that account for the interactions among human and software components to achieve a desired goal
- **Use Case Documents** – Developed use case documents to define the scope, utility and feature set of new entity IT platform including:
  - Data Analysis Use Cases
  - Data Submission Use Cases
- **Policy Questions** – Developed related data sharing IT policy questions for inter-workgroup discussions

## Data Analysis

### Use Case

- Analyzing Published and Unpublished Data From More Studies

## Data Submission Use Case

- Submitting a Data Package to be Hosted by the Platform

## Related Data Sharing Policy Issues

- Policy questions generated for inter-workgroup discussions

- 1. Analyzing published and unpublished data from more studies**
  - A data requester wishes to analyze both published and unpublished data from several related studies, with a specific analytic plan in mind.
  
- 2. Research Portfolio Analyses**
  - Having both study designs and individual participant-level data (IPD) results in computable form will allow descriptive analyses of clinical trials, e.g., describing the evidence base for a topic, the research portfolio of a sponsor, what queries and data requests are made on which studies, etc.
  
- 3. Secure Access to Pre-Publication Data Set for Peer Review**
  
- 4. Linkage to Administrative and/or Clinical Data to Extend Study Follow Up**
  
- 5. Submitting a Data Package to be hosted by the Platform**
  - The responsible party for a study wishes to submit a Post-Publication or Full Data Package to the Platform, for hosting by the Platform



**Summary:** A data requester wishes to analyze both published and unpublished data from several studies, with a specific analytic plan in mind

## Data Analysis Use Case

- Analyzing Published and Unpublished Data From More Studies

- **Actors** –
  - The Data Requester
  - The PI(s) of the Relevant Studies
- **Trigger** – Data Requester submits a request to the Platform for access to the Full Data Package and any Post Publication Data Packages for the studies
- **Preconditions** –
  - Studies have been completed, or the studies have published a results paper
  - Platform stores requested data packages, or can somehow provide access
  - Requester has an account on the Platform
  - Requester has specific analytic plan in mind
  - Data sets are annotated with required metadata
  - All downloadable data is hosted by the Platform

**Summary:** A data requester wishes to analyze both published and unpublished data from several studies, with a specific analytic plan in mind

## Data Analysis Use Case

- Analyzing Published and Unpublished Data From More Studies

- **Post-Conditions** – Requester has sufficient access to the data to fit his or her analytic needs, and any results from their analysis are indexed to the original data set

### High Level Bullets of Normal Flow:

- **Initial Data Exploration** – Researcher can:
  - Search key features of trials to see if they would provide data to answer the specific analytic question
  - Define and save searches
  - Share saved searches with collaborators
  - Mark studies of interest for data requests
- **Requesting the Data** – Researcher can:
  - Attach studies of interest to a data request
  - Submit a data request as specified by Governance Working Group

**Summary:** A data requester wishes to analyze both published and unpublished data from several studies, with a specific analytic plan in mind

## Data Analysis

## Use Case

- Analyzing Published and Unpublished Data From More Studies

## High Level Bullets of Normal Flow Continued:

- **Adjudication of Request** – Should support review process as specified by Governance Working Group
- **Authenticating Access and Receiving Authorization** –
  - Need authorization to access specific data sets
  - Need authorization to download data set for approved data requests only
- **Use and Analysis of non-IPD** –
  - Examples include SAP, analytic code, CRFs
  - To be determined – full or partial access to non-IPD data

**Summary:** A data requester wishes to analyze both published and unpublished data from several studies, with a specific analytic plan in mind

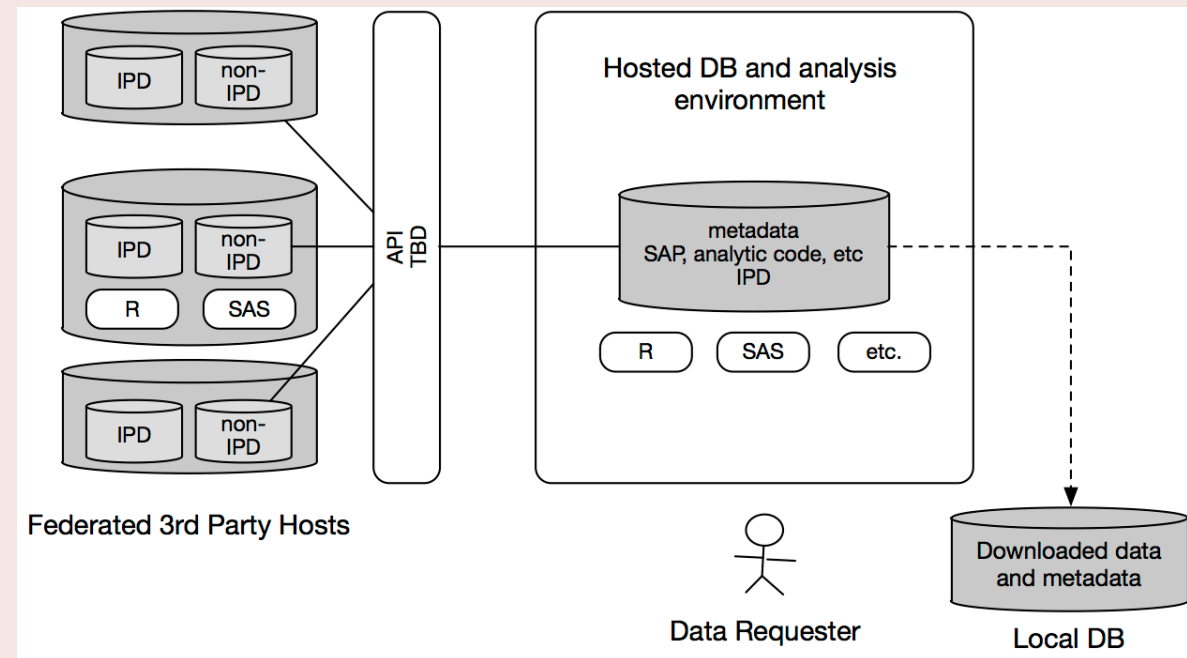
## Data Analysis

## Use Case

- Analyzing Published and Unpublished Data From More Studies

## High Level Bullets of Normal Flow Continued:

- Analysis of IPD** – Alternate Flow 1 and/or 2:



## Data Analysis

### Use Case

- Analyzing Published and Unpublished Data From More Studies

## Data Submission Use Case

- Submitting a Data Package to be Hosted by the Platform

## Related Data Sharing Policy Issues

- Policy questions generated for Governance Working Group

**Summary:** The responsible party for a study wishes to submit a Post-Publication or Full Data Package for hosting by the Platform

## Data Submission Use Case

- Submitting a Data Package to be Hosted by the Platform

- **Actors** –
  - The PI of the study
  - Delegated data submitter(s)
- **Trigger** – Submitter requests ability to upload one of the data packages to be hosted by the Platform (specific process will be specified by the Governance Working Group)
- **Preconditions** –
  - Submitter has an established Platform account
  - Study has already been registered in an official trial registry
  - Submitter has access to all the required components of the Data Package
  - Submitter has reviewed the instructions on how to prepare the Data Package on the Platform website

**Summary:** The responsible party for a study wishes to submit a Post-Publication or Full Data Package for hosting by the Platform

## Data Submission Use Case

- Submitting a Data Package to be Hosted by the Platform

- **Post-Conditions** – Data Package is fully indexed to the other contents of the Platform, and is searchable and accessible to data researchers

### High-Level Bullets of Normal Flow:

- **Initial Submission Process** –
  - Study PI sends request to Platform for submission authorization
  - Must verify identity before an account is opened
- **Help Submitters Prepare Data Package to Meet Specifications** – Data set metadata, annotations for data alignment, and de-identification processes
- **Still To Be Determined** – Data set uploading, ensuring that data sets meet specifications, indexing data sets, making data sets available

## Data Analysis

### Use Case

- Analyzing Published and Unpublished Data From More Studies

## Data Submission Use Case

- Submitting a Data Package to be Hosted by the Platform

## Related Data Sharing Policy Issues

- Policy questions generated for Governance Working Group



## Related Data Sharing Policy Issues

- Policy questions generated for Governance Working Group

- **Authorization and Authentication** –
  - Who can establish an account on the Platform?
  - Do data requesters need to have a specific analytic question, or can the request be completely exploratory?
- **Quality of Data and Timelines** –
  - Should platform users be held to identical data sharing standards as the original data generator (i.e. should platform users have to share their analysis results and analytic code back with the platform)?
- **Access & Protection** –
  - What are the policies on which datasets can be downloaded?
  - Must all data depositors use the same de-identification approach?



## Questions and Discussion