



**U.S. Food and Drug Administration  
And Project Data Sphere  
Symposium V**

Growing Dimensions of Big Data in Cancer Research



**Tuesday, October 24, 2017 | 11:30 AM – 5:00 PM | Bethesda North Marriott, Rockville, MD**

**SYMPOSIUM V SUMMARY REPORT**

**Participants: 40**

**Organizations represented: 21**

- |   |   |
|---|---|
| Alliance for Clinical Trials in Oncology Foundation | NQ Oncology                               |
| AstraZeneca   | Pancreatic Cancer Action Network          |
| Celgene Corporation                                 | Parker Institute for Cancer Immunotherapy |
| Helsinn Therapeutics (US) Inc.                      | Pfizer                                    |
| Hogan Lovells US, LLP                               | Project Data Sphere, LLC                  |
| IBM   | RTI International                         |
| Kaggle, Inc.  | SAS Institute                             |
| National Cancer Institute, NIH                      | Tulane University School of Medicine      |
| Columbia University Medical Center                  | University of Colorado                    |
| Leidos Biomedical Research, Inc.                    | US FDA                                    |
| Novartis Pharmaceuticals Corporation                |   |

**Summary Notes:**

- ◆ ***The Potential and Reality of Data Sharing in Fulfilling the FDA Vision of an Evidence-Generation System. (Sean Khozin, FDA)***
  - The 4th industrial revolution is about data and algorithms
  - Inflection point is taking us from a reduction approach to development and approval to a holistic approach. We have had a reductionist approach to generating evidence. It's about the genomic milieu -- this transformation can only occur with data. With all this data, we need algorithmic analytics. Frequentist methods may not lend themselves to where we want to be in drug development.
  - 4v's of big data -- velocity, volume, variety, veracity. At FDA, clinical trials (variety), that is mostly structured (veracity), gigabytes(volume), and intermittent (velocity). Next level is EHR (variety), unstructured (veracity), petabytes(Volume) and real-time (velocity)
  - 21st Century Cures Act -- law is urging us to get to this holistic universe and innovate. Critical to implement "drug development tools". This used to be called "biomarkers" but that's a very reductionist point of view. Now tools include biomarkers, but any other method, material or metric. Also urges the development of novel trial designs like adaptive trials or the ability to avoid trial execution at all.



# *U.S. Food and Drug Administration And Project Data Sphere Symposium V*

Growing Dimensions of Big Data in Cancer Research



- 
- ◆ **Open Science and Crowdsourcing Panel:**
  - James Costello - Practicalities of working in a Challenge environment
  - Addison Howard- Overview of Kaggle Challenge
  - Gustavo Stolovitzky- Overview of an IBM Challenge
  - Audience Interactive Demo of challenge using tumor images.
- ◆ **Project Data Sphere Initiative Update**
  - Mace Rothenberg - Facilitator
  - Dave Handelsman - NCTN / NCORP Data Archive Integration.
  - Paul Kluetz - Patient Reported Outcomes Investigations
  - Steven Cohen - Socio-Determinant Factors Investigations
  - John Dornan - Global Oncology Big Data Alliance
  - Dave Handelsman - Metrics and Milestones
  - Mace Rothenberg - Announce that PDS has reached our 100K patient lives milestone and thank Monica Bertagnolli and Alliance for providing the datasets for us to reach this goal.
- ◆ Presentation: Gideon Blumenthal
- ◆ Presentation: Laleh Amiri-Kordestani,
- ◆ Mace Rothenberg: Symposium summary
  - Discussion & Next Steps