

Real World Data Session

- 1) What is the validity of real-world data – how do you measure what is representable and reproducible
 - Spark example – to enhance the quality of data and access of data, setup ability to recruit participants into the cohort – designed by researchers – question on how to maximize the usefulness of this data
 - Can you use this data for hypothesis testing
 - Psychiatric diagnosis – not a standard in terms of how a clinician gets to diagnosis

EHR

- Alternative – from health care (electronic medical data records)
- Would like to see standards around EHR
- Would like to see validation around key variables when using new devices, apparatus, tests
- Evidence care based practices – data quality in EMR data is messy
- **Example:** Build a visualization chart with measurements that are key to clinician – creates an opportunity to get high quality data – can go into medical records and de-identified to also move into a research database
- Can also use the unique identifier to also layer in additional research studies
- Have measurements in context of care pathway

How to work in real-world data environment – incentivize buy-in and participation

- Understand what brings clinicians value – emphasize the new opportunities in data sharing
- (concept of carrot flavor stick – works well with a research institution)
- How do you move to community provider – may be able to work predictions into this system
- Regular sessions with privacy and ethics office – worked to do privacy impact
- Working with clinical performance perspective, maintain firewall between clinical and research data

Concepts:

Robust, accurate and specific data is the wholly grail in terms of EHR data

Structured data only goes so far – example – no SV10 code for seizure frequency in epilepsy or change in diagnosis in structured data. Need to be able to extract this information from text notes to understand naturally history – many neurological diseases are misdiagnosed – for instance many individuals that meet criteria for dementia that are underdiagnosed.

Use of EHR data to study symptomatology – important about how we define disease and disease progression

Patient reported information not found in EHR – how is the patient feeling – becoming increasingly important for post-marketing studies, but also to inform on how they relate to patient's perception of disease. Real world data has a place in regulatory decisions

- Not just health care data that is faced with these issues
- Does the physician have time to use tools and report additional data

Critical points for data capture

- Capturing some provenance of how that data was captured and what other data was collected
- Need to build provenance of how data was collected -ie. need to understand when the data was collected
- How do you incentivize physicians – some solution is to have a scribe program to improve the population of the HER – probably more problematic in larger health care systems – how many data points are you responsible for – how do you prevent burnout – what about using voice recognition technology to populate a form structure

Cloud Structure

- Important as new modes of data collection come on line that capture large scale data
- Data Policy, privacy, security, breach management
- Data storage, data sharing and access
- Analytical tools – Jupyter notebooks - workflows
- Encourage data use in the cloud, rather than data download to local environments