Maximizing Epilepsy Data Usage for Better Outcomes: Informatics and Common Language

Samden D. Lhatoo, MD, FRCP,
Satya S. Sahoo, PhD
Case Western Reserve University
ICARE: Interagency Collaborative to Advance Research in Epilepsy

NIH National Institutes of Health
NIH Center for SUDEP Research (CSR)

Predictive genes, clinical biomarkers and basic mechanisms of SUDEP
4000 epilepsy patients undergo pre-surgical monitoring

- 2000 patients undergo surgery
  - 1000-1500 seizure free patients at low risk of SUDEP
  - 500-1000 patients with active epilepsy & at risk of SUDEP
- 2000 patients not appropriate for surgery & at risk of SUDEP
  - Total = 2500-3000 patients at risk of SUDEP

25-30 SUDEP cases per year
The Challenges

• Multi-center – each center follows different approaches
  o Disparate terminology
  o Paper based, Electronic documents
  o Distinct form structure

• Difficulty in post collection data analysis: o Patient cohort identification across centers o Uncovering disease association
  o Correlating signal data (EEG) across centers

• Data quality – without increasing workload or cost
• Focal dyscognitive seizure
• Complex partial seizure
• Dialeptic seizure
• The Requirements
  o Consistent terminology – facilitates interoperability, data sharing
  o Real-time data curation – range check, auto-completion
  o Optimized data entry – skip patterns to dynamically change form structure
  o Structured data stored in patient database – easy to query and analyze
  o Shared resource – develop once, deploy at multiple locations
  o Easy maintenance – changes are easily propagated
Informatics Highway: Interoperability and Integration

- Common Terminology System – upfront or backend
- Common Data Elements (CDE): NINDS CDEs
- Shared Meaning for Data Integration
- Accurate and Consistent Interpretation of Data
What Can Common Terminology Achieve?

- Reconcile different categories of heterogeneity:
  - syntactic heterogeneity: Slow Spike and Wave complex vs. Slow Spike-and-Wave complex
  - structural heterogeneity: ADNFLE vs Autosomal Dominant Nocturnal Frontal Lobe Epilepsy
  - semantic heterogeneity: Focal dyscognitive seizure vs. Dialeptic seizure vs. Complex partial seizure
Managing Epilepsy Well (MEW) Network Informatics

- 8 centers across US
- Multiple approaches for self management of epilepsy
- Integrated data from different studies for cohort studies and comparative analysis
Multi-Center Epilepsy Cohort Identification

- Integrate Multiple Studies
- Common Vocabulary

Insight!
Clinical Research Platform

SELECT STUDY DATA ELEMENTS

- Race
  - Asian
  - White
- Education
  - Grades 1 through 8 (Element)
- Relationship Status
  - Married
Multi-Center Epilepsy Cohort Identification

**Insight!**
Clinical Research Platform

Data Visualization using Pie Charts for Individual Research Studies

Data Visualization using Histograms

Users can modify cohort identification queries in real time to update results
Ontology-driven Patient Information Capture (OPIC)

- Structured data entry
Ontology-driven Patient Information Capture (OPIC)

- High level of granularity
Ontology-driven Patient Information Capture (OPIC)

- Support embedding of multi-media images
- Data entered can be immediately queried
- Important for both care and clinical research
- Easier sharing across multiple collaborators
Epilepsy and Seizure Ontology

- Follows four-dimensional classification approach
  - Etiology
  - Seizures
  - Anatomical Location
  - Related Medical Conditions

- Uses Web Ontology Language (OWL2)

- Description Logic allows consistent and accurate interpretation of terms

Sahoo et al., JAMIA 2014
Epilepsy and Seizure Ontology: Multi-axial Structuring

- **OWL2 Object Properties:** Allows classification of ontology terms along multiple dimensions
  - Etiology
  - Anatomical Location
  - Genetics
  - Others…

- Objects properties link terms together: Networked terminological system
Epilepsy and Seizure Ontology: Links to Existing Terminological Systems

• RxNorm: Developed by the NLM
• Gene Ontology
• Foundational Model of Anatomy (FMA) Ontology
Epilepsy and Seizure Ontology: Formal Model and Multi-Dimensional

ADNFLE

Genes

Etiology

Signal Features

Semiology

Sahoo et al., JAMIA 2014
Community Participation: Epilepsy Ontology Consortium

1. Add Name of Syndrome
   - Add New Syndrome
   - Edit Syndrome label

2. Add Details of the Syndrome
   - List of Syndromes
     - Childhood Epilepsy with Occipital paroxysms
     - Primary Reading Epilepsy
     - Parietal Lobe Epilepsy
     - Generalized Epilepsy with Febrile Seizures Plus (GEFS+)
     - Autosomal Dominant Nocturnal Frontal Lobe Epilepsy (ADNFLE)
     - Autosomal Dominant Partial Epilepsy With Auditory Features (ADPEAF)
Community Participation: Epilepsy Ontology Consortium

- Community participation through wiki-based features
- Interactive discussions and threads
- Community agreement towards a common epilepsy terminology
Ongoing and Future Directions: Neuroscience Big Data

- High performance computational pipeline for electrophysiological data
Ongoing and Future Directions: Neuroscience Big Data

Scalability and Adaptability of NeuroPigPen with Increasing Data Size and Hadoop Data Nodes

- **Hive**
- **Pig**
- **MapReduce**
- **HBase**
- **HDFS**

Cloud Computing
Acknowledgements