The FDA BEST System: Leveraging EHR data and Innovative Approaches for Surveillance of Biologic Products

Steve Anderson, PhD, MPP
Director
Office of Biostatistics & Epidemiology
Center for Biologics Evaluation and Research, FDA
FDA Center for Biologics Evaluation & Research (CBER)

Regulatory responsibility for:
- Allergenics
- Vaccines
- Blood and blood products
- Tissues
- Cellular Therapies
- Gene Therapies
- Advanced therapeutics
OBE Staff & Responsibilities

Areas of Expertise
• Biostatistics
• Computer modeling
• Epidemiology
• Pharmacovigilance
• Clinical expertise
• High performance computing - HIVE

Regulatory & Research Activities
• Covers all CBER-regulated products
• Statistical review efficacy and safety data submitted by sponsors
• Prelicensure review of safety monitoring plans (PVP)
• Routine pharmacovigilance
• Passive surveillance – FAERS, VAERS
• Active surveillance-Sentinel, BEST, CMS
• High performance computing and next generation sequencing
• Benefit-Risk Modeling
Legislative mandates drive CBER regulatory activities

• FDA Amendments Act (FDAAA) of 2007
  – Sentinel, government data, Clinical Trials.gov, PMC, PMR, REMs, safety reporting, pediatric safety

• Prescription Drug User Fee Act (PDUFA) VI
  – Sentinel, Real-world evidence (RWE), innovative clinical trials, patient input, benefit-risk assessment, guidance development, public meetings, others

• 21st Century Cures (2016)
  – RWE, vaccine innovation, patient-informed drug development, and others
Guidance: Good Pharmacovigilance Practices

- Identifying and describing safety signals
- Investigating a signal through observational studies
- Interpreting safety signals
- Developing a pharmacovigilance plan

CBER Active Postmarket Surveillance Systems

- FDA Sentinel Initiative
  - Sentinel PRISM – Harvard Pilgrim system (Claims)
  - CBER BEST (Largely EHR data and some claims data)

- Center for Medicare & Medicaid data
- Veterans Administration Data
- VSD
Sentinel PRISM – Harvard Pilgrim system

- Workhorse for many years for CBER regulatory/safety studies
  - >80 queries, >20 comprehensive studies
- Largely claims data
- PRISM for vaccines – covers 170 million persons
- BloodSCAN – covers ~200 million persons
Limitations of Claims Data for Regulatory Studies

• >6 months to retrieve analyze medical charts
• Long study times
• Long data lags ~9-12 mos
• Transfusions not always captured

• Need for another EHR-based system to better address CBER regulatory surveillance needs
Biologics Effectiveness and Safety (BEST) Initiative

• New CBER Active Post-market Surveillance Program
  – Largely EHR-based

• Started in 2017 with pilot program

• Awarded two five-year IDIQ contracts in FY2019
  – IQVIA
  – IBM Watson Health
  – Acumen
  – Dovell

• BEST continues as a pilot program into 2020
Why BEST?

- Expanded EHR data sources
- Claims data
- New linked EHR-Claims data
- Reduced data lag ~3-4 mos
- On-demand analytic capabilities
- Improved operational speed & more rapid study turnaround

- Better addresses unique characteristics of biologics
  - Incorporates unique coding such as ISBT-128 for blood
  - Rapid access to medical charts
1. Fully Operational Query, study and production enterprise system

2. Leverage innovations such as AI, NLP, and semi-automation of medical chart review and automated AE reporting
CBER Surveillance Priorities

• Routine Surveillance: Conduct evaluations of safety and effectiveness of biologic products (including vaccines)
• Evaluating safety of vaccination during pregnancy
• Signal Detection – use of NLP and Artificial Intelligence
• Pandemic Preparedness – near real-time surveillance
• Emerging Infectious Disease Surveillance & Monitoring
Add New Data Sources for Millions of Patients (2)

<table>
<thead>
<tr>
<th>Data Sources</th>
<th>IBM</th>
<th># Patient Records (millions)</th>
<th>Data Sources</th>
<th>Acumen</th>
<th># Patient Records (millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MarketScan</td>
<td></td>
<td>60</td>
<td>LRxDx (Claims)</td>
<td></td>
<td>160</td>
</tr>
<tr>
<td>CED (linked EHR and claims data)</td>
<td></td>
<td>4.9</td>
<td>Regenstrief Inst. (claims, EHR)</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>MedStar (EHR)</td>
<td></td>
<td>5</td>
<td>Columbia Univ (EHR)</td>
<td></td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Univ of Colorado (EHR)</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cerner (EHR)</td>
<td></td>
<td>23</td>
</tr>
</tbody>
</table>

Blue Health Intelligence (Claims) | 23
BEST Initiative GOAL 1: Data, Tools and Infrastructure for Surveillance of Biologics

ACCOMPLISHMENTS:
DESCRIPTIVE STUDIES
## Current Activities

### Descriptive Epidemiologic Studies:

#### Vaccines
- HEPLISAV-B
- SHINGRIX
- ZOSTAVAX
- FluMist
- All influenza vaccines
- TRUMENBA and BEXSGRO
- GARDASIL9

#### Blood-derived products
- Intravenous immunoglobulin (IVIG)
- Factor VIIa
- Factor VIII
- von Willebrand factor
- FEIBA
- KCENTRA
- Fibrin sealant
- Fibrinogen concentrate
Outcomes

- Syncope
- Thromboembolic events
- Coagulation product inhibitors (Factor VIII inhibitory antibodies)
- Hemolysis
- Anaphylaxis
Replication of Vaccine Study (Test Case)

- To test the new system, reproduced components of a published study

**Measles-Mumps-Rubella-Varicella Combination Vaccine and the Risk of Febrile Seizures**

- **Study Objective:** To assess the risk of febrile seizures in children receiving first dose of Measles, Mumps, Rubella, & Varicella (MMRV) compared to that of MMR and Varicella administered separately on the same day
LINKED EHR-CLAIMS DATABASE
Study Population

Use of deterministically linked Claims-EHR Data (IBM CED) enables:

- **Algorithms** to be applied to claims data elements
- **Validations** to be performed using structured EHR data elements

Source: IBM, 2018
CED DATABASE: PREGNANCY OUTCOMES & GESTATIONAL AGE VALIDATION
Study Objectives

1. Develop algorithms using ICD10 diagnosis codes and CPT/HCPCS procedure codes to
   a) Determine gestational age
   b) Classify pregnancy episodes as one of 4 outcomes:
      i. Full-term birth
      ii. Pre-term birth
      iii. Stillbirth
      iv. Spontaneous abortion
Study Objectives

2. Using GAIA* case definitions as a reference method
   – To validate estimated gestational age and outcomes classifications
   – Stratification of outcome certainty by strength of evidence
   – By comparing to clinician-adjudicated results based on review of structured CED (EHR) data elements
   – Example outcomes of interest: stillbirth, spontaneous abortion, live birth, low birth weight, etc.

*GAIA = Global Alignment of Immunization Safety Assessment in pregnancy
Study Population

- **n= 35,842 (100%)**
  - Pregnancy episodes identified in claims

- **n= 33,698 (94%)**
  - Pregnancy episodes with GA estimates in claims

- **n= 6,122 (17%)**
  - Pregnancy episodes with GA, LMP, or IUI/ET AND outcome in SNOMED or LOINC in EHR

- **n= 2,144 (6%)**
  - Pregnancy episodes without GA estimates in claims

- **n= 27,576 (77%)**
  - Pregnancy episodes without GA, LMP, or IUI/ET AND outcome in SNOMED or LOINC in EHR

Source: IBM Watson Health, 2018

Note: Preliminary results and subject to change
Clinician Adjudication Using Semi-Automated Chart Review

- Built-in questionnaire
- Structured components of EHR

Clinician Review

- Display GAIA-related structured EHR elements

Outcome Adjudication

- Full chart of structured EHR pregnancy episode available to clinician in detailed view

Source: IBM Watson Health, 2018
BEST Goal 2: Progress

• Semi-automated data extraction from EHR of:
  – CBER Product exposures
  – Product-related adverse events that meet defined criteria
• Development and validation of descriptive case reports
• Electronic submission of reports via FDA electronic gateway (e.g., FAERS and VAERS)
Innovative approaches are needed to achieve BEST’s aim of biologic product active surveillance

<table>
<thead>
<tr>
<th>Phase</th>
<th>Traditional Approaches</th>
<th>Innovative Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td><strong>Claims-Based:</strong> Insufficient for determining biologic exposure and outcomes</td>
<td><strong>EHR-Based:</strong> Access critical data including clinical notes, product codes, vital signs, and time stamps</td>
</tr>
<tr>
<td></td>
<td><strong>EHR-Based:</strong> Access critical data including clinical notes, product codes, vital signs, and time stamps</td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td><strong>Claims-Focused Common Data Models:</strong> Reduced granularly, data loss, and data lag</td>
<td><strong>EHR-Oriented Standards:</strong> Interoperability and granularly enabled through use of HL7 FHIR and OHDSI OMOP</td>
</tr>
<tr>
<td>Detection</td>
<td><strong>Rules-Based and Expert-Curated Algorithms:</strong> Built for curated claims queries</td>
<td><strong>Artificial Intelligence:</strong> High-dimensional EHR requires natural language processing, machine learning, deep learning, and computational phenotyping</td>
</tr>
<tr>
<td>Validation</td>
<td><strong>Manual:</strong> Costly and slow</td>
<td><strong>Semi-Automated:</strong> Efficient application-based tool</td>
</tr>
<tr>
<td>Reporting</td>
<td><strong>Manual:</strong> Voluntary and underreported</td>
<td><strong>Automated:</strong> Automated population and submission</td>
</tr>
</tbody>
</table>
BEST Aim 2: Contracts

I. Pilot Year (FY2018):
   IQVIA, including: Georgia Tech Research Institute (GTRI), Columbia University, Stanford University, and Regenstrief Institute

II. Five-Year Vehicle (FY2019-FY2023):
   IBM (awarded FY2019 task order), IQVIA, Acumen, and Dovel
Accomplishments for Goal 2: Pilot year

1. **Improved sensitivity and granularity of transfusion exposures compared to claims data alone**
   - Results: Applied NLP to transfusion nursing notes, component identification for 34,000 transfusions.
   - Chart review validation of 100 cases demonstrated 100% accuracy.

2. **Development of computable phenotypes (CP) for Transfusion-Associated Circulatory Overload (TACO)**
   - NLP-based queries = ~10% improvement in PPV for identified TACO cases.
Accomplishments for Goal 2: Pilot year (2)

3. NLP-based development of computable phenotypes (CP) for Post-Transfusion Sepsis (PTS)
   - Identified two definitive and several candidate PTS cases.

4. Infrastructure to support interoperability within BEST through CLARITY NLP platform

5. Building infrastructure for scale-up of CP-based case identification and automated report generation
   - Demonstrated with MIMIC test database using case characteristics from a published TACO case.
FY2019: Semi-Automated Validation Tool

- **Chart Review Tool**: Enables semi-automated clinical assessment with an intuitive UI
- **Abstraction**: Allows for simplified visualization of patient EHR information
- **Classification**: Reviewers efficiently document information related to, including:
  - Certainty of **biologic exposure**
  - Certainty of **adverse event** (or health outcome of interest)
  - Assessment of **causality** (or imputability)
  - **Evidence** for conclusions (used for both ICSR reporting and algorithm training)
REAL WORLD DATA GENERATION AND REAL WORLD EVIDENCE (RWE)
21st Century Cures Act
Framework for the RWE Program

• The 21st Century Cures Act (Public Law 114-255), signed into law on December 13, 2016 (Cures Act)

“Food and Drug Administration (FDA) is required to develop a framework for a program that will evaluate the use of real-world evidence (RWE) to help support the approval of a new indication for an approved drug or to satisfy post approval study requirements (RWE Program)”. 
CBER RWD/RWE Generation Systems

CBER uses a number of population-based data systems to conduct RWE safety and effectiveness studies including:

1. **CBER BEST** - Covers >20 million (EHRs)

2. **Sentinel – Harvard Pilgrim (HP)** - ~200 million persons (claims)

3. **Center for Medicare & Medicaid Services** ~50 million persons (claims) >65yrs

*Goal is to build RWE generation systems for use by FDA and stakeholders*
CBER RWD/RWE Studies
CBER RWE Safety Studies: CMS and Sentinel HP

Two examples of published studies leading to label changes, regulatory actions:

1. Immune globulins and Thrombotic events (Transfusion 2012) – CMS

2. Rotavirus vaccines and Intussusception (NEJM 2014) – Sentinel HP
CBER RWE Effectiveness Studies:
CBER-CMS collaboration to study vaccine effectiveness in Medicare population (>65 years)

Two examples of published studies:
2. Herpes zoster vaccine effectiveness and duration of effectiveness (CID 2017)

One very recent study:
3. Rapid response effectiveness study of cell versus egg-based influenza vaccines, 2017-18 season

Project Leads: Rich Forshee, Hector Izurieta
1. High dose vs. Standard dose influenza vaccine effectiveness study, 2012-2013 season

• **FDA-CMS Retrospective cohort study**
  >2.6 million beneficiaries > 65 yrs

• **High-dose vaccine = 22% (95% CI 15–29)** more effective than the standard-dose vaccine for:
  – prevention of probable influenza infections and
  – influenza hospital admissions

• **Sponsor required confirmatory trial**
  (n> 30,000) showed clinical benefit of HD vaccine

• **Relative efficacy HD vs Standard = 24.2%**
  (95%CI, 9.7-36.5)

• **Conclusion:** Good agreement between the two studies – FDA-CMS RWE study much greater study size and power
MERCK Zostavax for Herpes Zoster (HZ)

Pre-approval efficacy trials:
• **Shingle Prevention Study (SPS)**
  – Double-blind, placebo-controlled (DBPC) RCT 38,000 individuals > 60
  – Median follow-up 3.1 years - reduction in HZ incidence 51%

• **ZOSTAVAX Efficacy and Safety Trial (ZEST)**
  – DBPC RCT of 22,200 individuals 50-59 years of age
  – Median follow-up 1.3 years - reduction in HZ incidence 70%
Post Marketing Commitment to study long-term efficacy in ages 50-59

• Prospective observational study run by Kaiser Permanente Northern California
• Data on 1.3 million members, with over 350,000 individuals who received Zostavax and 100,000 individuals with more than 5 years follow up post vaccination
• Study is ongoing and will continue through 2023

Clinical studies section of labeling updated:

– In assessing effectiveness adjustments made for calendar time, age, sex, race/ethnicity, healthcare resource utilization, comorbid conditions, and immunocompromise status
– Vaccine effectiveness (VE) against HZ for 50-59 over first 3 years following vaccination was 60%
– For individuals 60-69, 70-79 and 80 or older average VE against 49%, 46% and 44% respectively.
Summary

- CBER uses a number of active surveillance systems to evaluate biologic product safety and effectiveness

- Launched BEST in 2017: a new active surveillance system for biologic products
  - Incorporates multiple large sources of EHR
  - Improved access to EHR provides
  - Reduced data lag

- BEST and other active surveillance systems can be successfully used for RWE generation
Acknowledgements

- CBER Sentinel Central Team
- Office of Biostatistics and Epidemiology
- CBER product offices: OVRR, OBRR, OTAT
- CMS Colleagues
- CDC and VSD Colleagues

- IBM Global Business Services, IBM Watson Health Team
- Acumen Team
- IQVIA Team
- OHDSI Collaborators
  - Columbia University
  - Regenstrief Institute
  - University of Colorado
  - Cerner
  - University of California Los Angeles
  - Georgia Tech Research Institute
  - Stanford University
Thank You
Data Quality Assessment

- Data Completeness
- Data Conformance
- Data Plausibility

Data Quality Assessment Checks