Real World Data and Evidence for Regulatory Decision Making

Jacqueline Corrigan-Curay
Office of Medical Policy
Center for Drug Evaluation and Research
FDA
Disclaimer

The views and opinions expressed in the following slides are those of the individual presenter and should not be attributed to the FDA.

No relevant financial relationship exists
Laying the Groundwork

A Framework for Regulatory Use of Real-World Evidence

September 13, 2017
What will be Sufficient?
## Real World Data vs Evidence

### Real-World Data Sources

- Claims Data
- EMRs/EHRs
- Prospective Observational Data
- Patient Pathways
- Surveillance
- Mortality Database
- Primary and Secondary Care Data
- Administrative Data
- Disease and Device Registries
- Pharmacy Data
- Cost Studies
- Mobile Devices
- Consumer Data
- Social Media

### Real-World Evidence

#### Identifying Unmet Needs

- Natural History
- Co-morbidities
- Burden of Illness
- Incidence and Prevalence
- Disease Mechanisms
- Clinical Practice Patterns

#### Informing Clinical and Policy Decisions

- Usage Patterns
- Outcome Predictors
- Pharmacovigilance
- Population-Level Impact
- New Indications
- Benefit/Risk in Subgroups

---

**Postmarketing Evaluation (Phase IV)**

---

<table>
<thead>
<tr>
<th>Drug</th>
<th>Indication</th>
<th>Status</th>
<th>Data</th>
</tr>
</thead>
</table>
| **Lutathera**              | GEP-NET                                        | Approved 2017 | - Open label clinical trial  
- Analysis of 360 patients in an investigator sponsored, expanded access protocol of 1214 patients* |
| *(lutetium 177 dotate)*    | Gastropanc. Neuroendo tumors                    |            |                                                                     |
| **Voraxaze**               | Treatment of MTX toxicity                       | Approved 2012 | - Approval based on open-label, NIH compassionate Use Protocol  
- Retrospective, non-random, un-blinded case series of 23 patients compared to historical control group |
| *(glucarpidase)*           |                                                 |            |                                                                     |
| **Uridine Triacetate**     | Treatment of 5 FU overdose                      | Approved 2015 | - Two single-arm, open label expanded access trial of 135 patients compared to case history control |
| **Blincyto**               | Treatment of Acute Lymphoblastic Leukemia       | Approved 2014 | - Single arm trial  
- Reference for effect weighted analysis of patient level data on chart review of 694 patients at EU and US study sites* |
| *(Blinăţumomab)*           |                                                 |            |                                                                     |
| **Carbaglu**               | Treatment of NAGS deficiency                    | Approved 2010 | - Retrospective, non-random, un-blinded case series of 23 patients compared to historical control group |
| *(Ganglioside E6 Tablets)* |                                                 |            |                                                                     |
| **Myozyme**                | Treatment of Pompe disease                      | Approved 2004 | - Open-label, non-randomized study of 18 patients compared to historical control group of 62 untreated patients |
| *(Glycozyme)*              |                                                 |            |                                                                     |
| **Refludan**               | Anti-coagulation in heparin-induced thrombocytopenia | Approved 1998 | - Two non-randomized, open-label multicenter trials using historical control comparator group from HIT Registry |
| *(Glycozyme)*              |                                                 |            |                                                                     |

*Bold = RWE

*https://www.nature.com/bcj/journal/v6/n9/full/bcj201684a.html
Targeting genomic subtypes for non-oncologic disease, cont…

2017 - Indication expanded from 10 mutations to 33 mutations based in part on in-vitro data

Highlighted mutations had clinical data as well

Patients 2 years or older with one mutation in CFTR gene that is responsive to ivacaftor based on clinical and/or in vitro assay data
Kalydeco Post Marketing Commitment

- 3-year, single arm, observational study
- Various subgroups of CF patients with CFTR mutations deemed responsive to ivacaftor based on \textit{in vitro} evidence
- Include all patients registered in the U.S. Cystic Fibrosis Foundation Patient Registry who have a newly designated CFTR mutation shown to be responsive to ivacaftor who initiate ivacaftor therapy following the date of approval of this supplement.
- Patients will be followed for at least 3 years on ivacaftor after ivacaftor initiation.
- The key outcomes of interest will include lung function measurements (FEV1), nutritional parameters (e.g., BMI), pulmonary exacerbations, hospitalizations, select CF complications (e.g., symptomatic sinus disease, CFRD, distal intestinal obstruction), and the presence of select pulmonary microorganisms (e.g., \textit{P} aeruginosa).
FDA Guidance

- Purpose is to ensure that patients whose records have the code-based operational outcome definition actually experienced that event
- Basic approach:
  - Select all or a sample of cases with the codes of interest
  - Review the medical charts to determine if the patient experience the event of interest
  - Calculate the positive predictive value of the code
- If the code or algorithm has been previously validated:
  - Cite the specific literature reference
  - Describe the validation algorithm in detail, including the population studied and the database used, time frame, and performance characteristics
- Need also to describe the sensitivity of the outcome definition:
  - Within the database
  - Within the population
- Other considerations:
  - Primary versus secondary positions
  - Inpatient versus outpatient
  - Diagnostic codes and procedure codes
Transition occurred in October 2015. The ICD-9, which was in place for nearly 4 decades in the United States, included unique codes for 14,000 diagnoses and 4,000 procedures. The ICD-10 expanded to include nearly 70,000 diagnoses and 72,000 procedures.

A surveillance of hospitalizations with a diagnosis of opioid use disorder across the transition from the ICD-9 to the ICD-10 found an abrupt 14% increase in the ICD-10-coded period relative to the preceding ICD-9-coded period.

An assessment of a 20% sample of all patients in the Veterans Affairs system found that ICD-10-coded data had a 2-fold higher odds of identifying Alzheimer disease and less than half the odds of accurately identifying patients with HIV/AIDS and those with alcohol and tobacco dependence.
The case definition utilized information from an in-progress retrospective cohort study of warfarin-related bleeding in Tennessee Medicaid enrollees 30 years of age or older. It identified inpatient stays during the study period of January 1990 through December 2005 with diagnoses and/or procedures that indicated a current episode of bleeding.

Of the 186 hospitalizations adjudicated, there were 165 (88.7% [95% CI, 83.4%-92.5%]) clinically confirmed bleeding-related hospitalizations, of which 133 were definite (71.5% [64.6%-77.5%]) or and 32 were probable (17.2% [12.5%-23.3%]) (Table 2). An additional 19 hospitalizations (10.2% [6.6%-15.4%]) were adjudicated as possibly bleeding-related, with a clinical history consistent with bleeding, but no objective evidence noted in the hospital record.

A case definition for bleeding-related hospitalizations suitable for automated databases had a positive predictive value of between 89% and 99% and could distinguish specific bleeding sites.
Longitudinal Assessment of Treatment Patterns and Events After Acute Coronary Syndrome (TRANSLATE-ACS) was a multicenter, longitudinal study of 12,365 patients with acute myocardial infarction (MI) enrolled at 233 US hospitals. Medical claims forms for all rehospitalizations of TRANSLATE-ACS participants during the study follow-up period (April 1, 2010, to May 13, 2014) were collected. Medical records were collected to perform independent physician adjudication of MI, stroke, and bleeding events. Our objectives were to (1) compare medical claims–identified vs physician-adjudicated cumulative incidence of recurrent MI, stroke, and bleeding events within 1 year after MI and (2) assess the accuracy of claims identified events using physician adjudication as the criterion standard.

Agreement between medical claims–identified and physician- adjudicated events was modest, with a $\kappa$ of 0.76 (95% CI, 0.73 to 0.79) for MI and 0.55 (95% CI, 0.41 to 0.68) for stroke events. In contrast, agreement between medical claims–identified and physician -adjudicated bleeding events was poor, with a $\kappa$ of 0.24 (95% CI, 0.19 to 0.30) for any hospitalized bleeding event and 0.15 (95% CI, 0.11 to 0.20) for moderate or severe bleeding on the GUSTO scale.
RWD and Endpoints

- Review of 138 new indications added to FDA labeling found that 108 (78.3%) of the pivotal clinical trials had a primary outcome that was not identifiable in US longitudinal databases (e.g. pathology results, changes in clinical scores and radiologic tumor responses).

EHR data have advantages of:

- A more complete and granular clinical picture
- Include labs/imaging/pathology reports

Since the Health Information Technology for Economic and Clinical Health (HITECH) Act was enacted, U.S. clinical notes have doubled in length (Epic Systems. Unpublished data.). Meaningful use incentives have unintentionally created requirements for substantial, low-value documentation.
Many systems/configurations: Fragmentation

Certified Health IT Vendors and Editions Reported by Ambulatory Health Care Professionals Participating in the Medicare EHR Incentive Program, July 2017

<table>
<thead>
<tr>
<th>Vendor</th>
<th>2015 certified technology</th>
<th>2014 certified technology</th>
<th>2011 certified technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epic Systems Corporation</td>
<td></td>
<td></td>
<td>92,241</td>
</tr>
<tr>
<td>Allscripts</td>
<td></td>
<td></td>
<td>33,044</td>
</tr>
<tr>
<td>eClinicalWorks, LLC</td>
<td></td>
<td></td>
<td>26,498</td>
</tr>
<tr>
<td>NextGen Healthcare</td>
<td></td>
<td></td>
<td>19,414</td>
</tr>
<tr>
<td>GE Healthcare</td>
<td></td>
<td></td>
<td>18,410</td>
</tr>
<tr>
<td>athenahealth, Inc.</td>
<td></td>
<td></td>
<td>18,018</td>
</tr>
<tr>
<td>Corner Corporation</td>
<td></td>
<td></td>
<td>17,132</td>
</tr>
<tr>
<td>Greenway Health, LLC</td>
<td></td>
<td></td>
<td>10,605</td>
</tr>
<tr>
<td>Practice Fusion</td>
<td></td>
<td></td>
<td>8,440</td>
</tr>
<tr>
<td>eMDs</td>
<td></td>
<td></td>
<td>6,333</td>
</tr>
<tr>
<td>Integrated Practice Solutions, Inc.</td>
<td></td>
<td></td>
<td>4,919</td>
</tr>
<tr>
<td>Eyefinity, Inc.</td>
<td></td>
<td></td>
<td>4,287</td>
</tr>
<tr>
<td>MEDENT - Community Computer Service, Inc.</td>
<td></td>
<td></td>
<td>3,835</td>
</tr>
<tr>
<td>McKesson</td>
<td></td>
<td></td>
<td>3,504</td>
</tr>
<tr>
<td>MedSeek, Inc.</td>
<td></td>
<td></td>
<td>3,477</td>
</tr>
<tr>
<td>Sunquest Information Systems, Inc.</td>
<td></td>
<td></td>
<td>3,208</td>
</tr>
<tr>
<td>Modernizing Medicine, Inc.</td>
<td></td>
<td></td>
<td>3,202</td>
</tr>
<tr>
<td>Compulink Business Systems, Inc.</td>
<td></td>
<td></td>
<td>2,957</td>
</tr>
<tr>
<td>SRSsoft</td>
<td></td>
<td></td>
<td>2,672</td>
</tr>
<tr>
<td>FairWarning Technologies, Inc.</td>
<td></td>
<td></td>
<td>2,304</td>
</tr>
<tr>
<td>All other commercial vendors (n=818)</td>
<td></td>
<td></td>
<td>93,737</td>
</tr>
<tr>
<td>Self-developers (n=46)</td>
<td></td>
<td></td>
<td>13,960</td>
</tr>
</tbody>
</table>
Demonstration Projects - Assessing Data Fitness / Standards

- OneSource: “enter the right clinical data once, use many times”
- FDA collaboration with Dr. Laura Esserman, UCSF
- Integration of standards based tools into the EHR to bring together health care and research
- Demonstration in breast cancer clinical trials

Courtesy of Dr. Laura Esserman and Susan Dubman
66.9 million members currently accruing new data
292.5 million cumulative patient identifiers between 2000 and 2017
14.4 billion pharmacy dispensings
13.3 billion unique medical encounters
45.6 million members with at least one laboratory test result

Networks of Data Exist

Largely Claims data

Largely EHR data

Data at a Glance
The Sentinel Distributed Database is comprised of quality-checked electronic data held by 18 partner organizations.

Traits and types of health data repositories

Ted D Wade

Abstract
We review traits of usable clinical data and offer a typology of clinical repositories with a range of known examples. Sources of clinical data suitable for research can be classified into types reflecting the data’s institutional origin, original purpose, level of integration and governance. Primary data nearly always come from research studies and electronic medical records. Registries collect data on focused populations primarily to track outcomes, often using observational research methods. Warehouses are institutional information utilities repackaging clinical care data. Collections organize data from more organizations than a data warehouse, and more original data sources than a registry. Therefore even if they are heavily curated, their level of internal integration, and thus ease of use, can be less than other types. Federations are like collections except that physical control over data is distributed among donor organizations. Federations sometimes federate, giving a second level of organization. While the size, in number of patients, varies widely within each type of data source, populations over 10 K are relatively numerous, and much larger populations can be seen in warehouses and federations. One imagined ideal structure for research progress has been called an “Information Commons.” It would have longitudinal, multi-geared (environmental through molecular) data on a large population of identified, consenting individuals. These are qualities whose achievement would require long-term commitment on the part of many data donors, including a willingness to make their data public.

Keywords: Registry, Observational research, Big data, Information commons, Data warehouse, Federated database

Learn about your health care options

CMS.gov
Centers for Medicare & Medicaid Services

Medicare
Medicaid/CHIP
Medicare-Medicaid Coordination
Private insurance
Innovation Center
Regulations & Guidance
Research, Statistics, Data & Systems
Outreach & Education

Home > Research, Statistics, Data and Systems > Integrated Data Repository (IDR) > Integrated Data Repository (IDR)

Integrated Data Repository (IDR)
The Integrated Data Repository (IDR) is a high-volume data warehouse integrating Parts A, B, C, D, and DMIC claims, beneficiary and provider data sources, along with ancillary data such as contract information, risk scores, and many others. Access to this robust integrated data supports much needed analytics across CMS.
Data Standards Demonstration

**FUTURE State**

- **Sentinel**
  - CDM
  - 19 Data Partners*

- **PCORNET CDM**
  - CDM
  - 13 CDRN + 21 PPRN*

- **FDA, PCOR, and other Researchers**
  - Portal
  - Tools
  - Mechanism to crosswalk the models

- **OHDSI/OMOP**
  - CDM
  - 14 Data Partners*

- **i2b2**
  - CDM
  - > 60*
To these ends, as part of the President’s Fiscal Year 2019 Budget, we’ve put forward a $100M medical data enterprise proposal to build a modern system that would rely on the electronic health records from about 10 million lives. This system would expand the data enterprise that we already maintain by incorporating new information from electronic health records, and other sources that would allow us to more fully evaluate medical products in the post-market setting.

This is the next evolution in the Agency’s development of a comprehensive data enterprise to improve medical product regulation and better inform us on the safety and benefits of new innovations.
Patient Centric RWE may require more than Health Care Records
# Wide Spectrum of Potential Uses of RWD / RWE in Clinical Studies

Different Challenges and Opportunities for Each Approach

<table>
<thead>
<tr>
<th>Randomized Interventional</th>
<th>Interventional non-randomized</th>
<th>Non-randomized / non-interventional</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traditional Randomized Trial Using RWD Elements</strong></td>
<td><strong>Trials in Clinical Practice Settings</strong></td>
<td><strong>Observational Studies</strong></td>
</tr>
<tr>
<td>RWE to assess enrollment criteria / trial feasibility</td>
<td>eCRF + selected outcomes identified using EHR/claims data</td>
<td>Propective data collection</td>
</tr>
<tr>
<td>RWE to support site selection</td>
<td>Mobile technology used to capture supportive endpoints (e.g., to assess ambulation)</td>
<td>Registry trials/study</td>
</tr>
</tbody>
</table>

- Pragmatic RCTs
  - Pragmatic RCT using eCRF (+/- eHR data)
  - Pragmatic RCT using claims and eHR data

- Single arm study using external control

**Increasing reliance on RWD**

- Traditional RCT
- RWE / pragmatic RCTs
- Observational cohort

Courtesy of Peter Stein, FDA
Real-World Evidence — What Is It and What Can It Tell Us?

Rachel E. Sherman, M.D., M.P.H., Steven A. Anderson, Ph.D., M.P.P.,
Gerald J. Dal Pan, M.D., M.H.S., Gerry W. Gray, Ph.D., Thomas Gross, M.D., M.P.H.,
Nina L. Hunter, Ph.D., Lisa LaVange, Ph.D., Danica Marinac-Dabic, M.D., Ph.D.,
Peter W. Marks, M.D., Ph.D., Melissa A. Robb, B.S.N., M.S., Jeffrey Shuren, M.D., J.D.,
Robert Temple, M.D., Janet Woodcock, M.D., Lilly Q. Yue, Ph.D., and Robert M. Califf, M.D.

As we adapt the tools and methods of traditional trials to real-world settings, we must consider the components of such trials that are critical to obtaining valid results and minimizing bias.
We All Need Confidence and Experience in Using New Data Streams, Technologies, and Analytical Methodologies for RWE

Evaluating authentication options for mobile health applications in younger and older adults

Kelly Grindrod, Hassan Khan, Urs Hengartner, Stephanie Ong, Alexander G. Logan, Daniel Vogel, Robert Geburtys, Jian Yang

Published: January 31, 2017 • https://doi.org/10.1371/journal.pone.0180618

Methodological approaches in analysing observational data: A practical example on how to address clustering and selection bias.

Trutschel D1, Palm R2, Holle D3, Simon H4

Risk-Based Source Data Verification Approaches: Pros and Cons

Machine Learning Healthcare Applications 2018 and Beyond
What questions can be answered using available data and current data sources?

How can we improve and establish new systems and data sources to better answer key questions?

ALL EXPERTISE NEEDED

- Data Analytics
- Data linkage
- Quality control and validations
- Database design, maintenance, and quality assurance
- Data security and confidentiality
- Study designs
- Statistics

Future

- Research fully embedded in care settings (no data is wasted).
- Integrated/connected systems throughout the entire health care continuum with feedback loops.
- Seamless and integrated auditing and quality control mechanisms
- Flexible and linkable on-demand data aggregation from databases/registries.
- All stakeholders engaged (including patients)
- Secured and traceable access and management of data (blockchain)
- RWE continuously utilized to support decision making processes.

FIT FOR PURPOSE

Convergence of all relevant evidence
Acknowledgments

• Khair ElZarrad
• Dianne Paraoan
• David Martin
• Peter Stein
Questions/ Comments

CDERMedicalPolicy-RealWorldEvidence@fda.hhs.gov