Recent Statistical Developments in Considering Real World Evidence for Regulatory Decision Making

Martin Ho, MS
Associate Director
Office of Biostatistics & Epidemiology
Center for Biologics Evaluation & Research
Center for Biologics Evaluation & Research

Vaccines, Blood & Biologics

Check out new FDA home page! https://www.fda.gov/vaccines-blood-biologics
Products Regulated by CBER

- Vaccines (preventative and therapeutic)
- Gene therapies
- Human tissues, engineered cellular products
- Whole blood, plasma and blood products
- Biologics related devices
- Live biotherapeutic products
- Xenotransplantation products
- Allergenics
Scientific Evidence is Context Specific

- Specify research question
- Required level of evidence
- Design study
- Collect & analyze data
- Evidence
RWD & RWE: CDER & CBER Definitions

**Real World Data (RWD):** Data relating to patient health status and/or the delivery of health care routinely collected from a *variety of sources.*

- electronic health records (EHRs)
- claims & billing data
- data from product & disease registries
- patient-generated data including in home-use settings
- data gathered from other sources that can inform on health status e.g. mobile devices

**Real World Evidence (RWE):** Clinical evidence regarding the usage and potential benefits or risks of a medical product derived from analysis of RWD.

Generated using many different study designs, including but not limited to, randomized trials (e.g., large simple trials, pragmatic clinical trials) and observational studies.
## Spectrum of Potential Uses of RWD

<table>
<thead>
<tr>
<th>Traditional Randomized Trial Using RWD Elements</th>
<th>Trials in Clinical Practice Settings</th>
<th>Observational Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>RWD to assess enrollment criteria / trial feasibility</td>
<td>RCTs with Pragmatic designs</td>
<td>Prospective data collection</td>
</tr>
<tr>
<td>RWD to support site selection</td>
<td>RCT using eCRF +/- eHR data</td>
<td>Registry trials/study</td>
</tr>
<tr>
<td>eCRF + selected outcomes identified using EHR/ claims data</td>
<td>RCT using claims and eHR data</td>
<td>Prospective Cohort Study</td>
</tr>
<tr>
<td>Mobile technology used to capture supportive endpoints (e.g., to assess ambulation)</td>
<td>Single arm study using external control</td>
<td>Using existing databases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case – Control</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Retrospective Cohort Study (HC)</td>
</tr>
</tbody>
</table>

**Increasing reliance on RWD**

- **Traditional RCT**
- **RCTs using RWD**
- **Observational studies**

Courtesy of Peter Stein, OND
Efficacy & Substantial Evidence

“evidence consisting of adequate and well-controlled investigations, including clinical investigations, by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involve on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof.”

Federal Food, Drug, and Cosmetic Act 1962

Drug Regulation History: https://www.fda.gov/about-fda/histories-product-regulation/historical-case-studies-drug-regulation
Adequate & Well-Controlled Study

For special circumstances
1. Clear objectives, summary of methods & results
2. Design permits a valid comparison with a control (concurrent and historical controls)
3. Adequate selection of patients
4. Assigning patients to treatment and control groups minimizes bias
5. Adequate measures to minimize biases on subjects, observers, and analysts
6. Well-defined and reliable assessment of subjects’ responses
7. Adequate analysis to assess drug results

Ordinarily randomization

Blinding

Regulations 21CFR314.126
CDER & CBER
RWE Program Framework

• Intended for drug & biological products
• Outlines FDA’s plan to implement the RWE program
• Multifaceted program
  – Internal processes
  – Guidance development
  – Stakeholder engagement
  – Demonstration projects

FDA RWE Program Framework: https://go.usa.gov/xmQnf
Evaluation Framework of RWD/RWE Used in Regulatory Decisions

Are RWD/RWE fit for answering a regulatory question?

I. Data (including RWD)
II. Study Design
III. Study Conduct

Source: FDA RWE Program Framework
I. Fitness of RWD

- Criteria specific to regulatory question
- Data uses
  - Population selection
  - Outcome ascertainment
  - Safety and study monitoring
- Multiple data sources may be needed
- How to access data reliability, validity, relevance?

Source: FDA RWE Program Framework
II. Study Design

• Population selection
• Comparator groups
• Outcome ascertainment, blinding
• Exposures patterns, dropouts
• Treatment definition (estimand)
• Hypothesis (e.g., non-inferiority)

III. Reg. Considerations

• Human subject protection
• Data traceability, auditing, and record keeping
• Safety reporting
• Study integrity and responsibility

Source: FDA RWE Program Framework
Guidance Documents

2017 (CDRH & CBER)
• Use of RWE to Support Regulatory Decision-Making

2018 (CDER & CBER)
• FDA Real World Evidence Program Framework
• Use of Electronic Health Record Data in Clinical Investigations

2019 (CDER & CBER)
• Rare Diseases: Common Issues in Drug Development (Draft)
• Rare Diseases: Natural History Studies for Drug Development (Draft)
• Submitting Documents Using RWD & RWE to FDA (Draft)
ASA BIOP RWE Scientific Working Group

- 20 SMEs collaborate on RWE statistics in pre-competitive space
- Advance statistical methods that facilitate use of RWE in regulatory context
- Engage regulators on RWE guidance and/or guiding principles

<table>
<thead>
<tr>
<th>Industry member</th>
<th>AbbVie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weili He, Co-Chair</td>
<td>AbbVie</td>
</tr>
<tr>
<td>Jie Chen</td>
<td>Merck</td>
</tr>
<tr>
<td>Yixin Fang</td>
<td>AbbVie</td>
</tr>
<tr>
<td>Qi Jiang</td>
<td>Seattle Genetics</td>
</tr>
<tr>
<td>Kwan Lee, Co-Lead</td>
<td>Janssen</td>
</tr>
<tr>
<td>Xiwu Lin</td>
<td>Janssen</td>
</tr>
<tr>
<td>Yang Sung</td>
<td>Vertex Pharma. Inc.</td>
</tr>
<tr>
<td>Hongwei Wang</td>
<td>AbbVie</td>
</tr>
<tr>
<td>Roseann White</td>
<td>The Third Opinion</td>
</tr>
<tr>
<td>Richard Zink</td>
<td>Target Pharma. Solution</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Academic/FDA member †</th>
<th>CBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin Ho, Co-Chair</td>
<td>CBER</td>
</tr>
<tr>
<td>Telba Irony</td>
<td>CBER</td>
</tr>
<tr>
<td>Mark van der Laan</td>
<td>UC Berkeley</td>
</tr>
<tr>
<td>Hana Lee</td>
<td>CDER</td>
</tr>
<tr>
<td>Mark Levenson, Co-Lead</td>
<td>CDER</td>
</tr>
<tr>
<td>Zhaoling Meng</td>
<td>BMGMRI‡</td>
</tr>
<tr>
<td>Pallavi Mishra-Kalyani</td>
<td>CDER</td>
</tr>
<tr>
<td>Frank Rockhold</td>
<td>Duke</td>
</tr>
<tr>
<td>Tingting Zhou</td>
<td>CBER</td>
</tr>
<tr>
<td>Ben Goldstein</td>
<td>Duke</td>
</tr>
</tbody>
</table>

† Liz Stuart (JHU) participates as non-member
‡ Bill & Melinda Gates Medical Research Institute
ASA BIOP RWE Working Group (cont.)

• Divide into 2 Workstreams by use of RWD/RWE
  – **WS 1: Label expansion** (Weili He & Mark Levenson co-lead)
  – **WS 2: Clinical trial design** (Martin Ho & Kwan Lee co-lead)

• Apply same approach for both Workstreams

  - Literature review
  - Identify key topics
  - Gap analysis
  - Research agenda

• Deliverables (in progress)
  – Publish 2 complementary papers on findings & recommended research agenda in the same issue of a peer-reviewed journal
## WS1 Identified Key Topics

<table>
<thead>
<tr>
<th>#1: Use RWD for label expansion</th>
<th>#2: Use RWD to <strong>inform study design &amp; analysis</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Reg., scientific, &amp; ethical issues</td>
<td>Study of retrospective data only</td>
</tr>
<tr>
<td>Data sources &amp; study types</td>
<td>Prospective study with external data</td>
</tr>
<tr>
<td>Estimands (treat. effect) in RW setting</td>
<td>Causal inference issues in reg. setting</td>
</tr>
<tr>
<td>Confounding control</td>
<td></td>
</tr>
</tbody>
</table>


WS1 Label Expansion

Regulatory, scientific, ethical issues

• US regulatory threshold for drug effectiveness:
  Substantial evidence from adequate & well-controlled investigations (AWC)

• Valid control group: context-specific
  – Common: Concurrent control + randomization
  – Unmet medical needs: Historical or external controls

• 21st Century Cures Act: FDA maintains same evidentiary standards

• Potential ethical issues using patient care data sources (e.g., EHR, claims)
  – Data privacy
  – Consent
  – Data ownership
  – Transparency of use
### Data Sources & Study Types

#### Data Sources

- **Experimental:** Hybrid or pragmatic clinical trials
- **Non-experimental** (table)

#### Study Types

- Intervention vs. non-intervention
- Randomized vs. non-randomized
- Prospective vs. retrospective
- Hybrids
- Pragmatics vs. traditional RCTs

<table>
<thead>
<tr>
<th>Source</th>
<th>I: Research data sources</th>
<th>II: Transaction data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>Data collected <em>primarily</em> for research</td>
<td>Data used <em>secondarily</em> for research</td>
</tr>
</tbody>
</table>
| Example | • Data specifically for study purpose  
  o Framingham Heart Study  
  o Cardiovascular Health Study  
  • Data intended for other studies  
  o Nurses’ Health Study  
  o Some registries | • Clinical documentation  
  o Electronic health records  
  o Wearable devices  
  • Administrative  
  o Claims data  
  o Geocoding/census |
# WS1 Label Expansion: Estimands in RCT vs. Real World Setting

<table>
<thead>
<tr>
<th>Attributes</th>
<th>RCT</th>
<th>Real Word Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population</strong></td>
<td>Detailed incl/excl; more homogeneous population &amp; care following protocol</td>
<td>Broad &amp; heterogeneous pop. from routine clinical practice but:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Confounded by local reimbursement (treatment decision is typically ahead of study participation) and data sources</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Required methods to control confounding</td>
</tr>
<tr>
<td><strong>Endpoints</strong></td>
<td>Well-defined outcomes measured for the study</td>
<td>• Under-reporting or lack of disease-specific clinical outcomes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Outcome definition/algorithm with suboptimal specificity &amp; sensitivity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Info. bias can be high; uncommon use of NLP for unstructured data</td>
</tr>
<tr>
<td><strong>Intercurrent Events</strong></td>
<td>Extensive efforts devoted for ensure patients f/u &amp; data completeness</td>
<td>May be more suitable for studies with long-term f/u, but:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Treat. change common &amp; reasoning typically not well-recorded;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Treat. adherence lower as medication is not provided</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Proportion of missing data &amp; loss to f/u higher</td>
</tr>
<tr>
<td><strong>Reporting measures</strong></td>
<td>Mostly population-level group comparison for effect size</td>
<td>Very diverse measurements depending on research questions (e.g., prevalence/incidence, disease progression, treatment pattern, disease burden, population-level comparative effectiveness, patient-level treatment response.</td>
</tr>
</tbody>
</table>
WS1 Label Expansion: Some Types of Confounding

<table>
<thead>
<tr>
<th>Residual confounding</th>
<th>Confounding by Indication</th>
<th>Time-varying Confounding</th>
</tr>
</thead>
</table>
| • Distortion remains after controlling for confounding in the design and/or analysis of a study;  
  • Usually unmeasured | • A symptom or disease sign indicates (or contraindicate) for a treatment  
  • Thus, it correlates with:  
    1) use of a med. product (or its avoidance), and  
    2) outcome related to the disease for which the product is indicated (or contraindicated) | • It occurs when confounders have values that change over time |
WS1 Label Expansion: Confounding Control

• When?
  – At **design** stage
    • Ex: Randomization, matching, population restriction
  – At **analysis** stage
    • Ex: Restriction and stratification, standardization and multiple regression, matching, G-methods, disease risk scores, machine learning

• How?
  – Method for control of measured confounding
  – Methods for control of unmeasured confounding, e.g., instrumental variables
  – Hybrid methods for control of measured and/or unmeasured confounding
### WS2 Identified Key Topics

<table>
<thead>
<tr>
<th>#1: Use RWD for <strong>label expansion</strong></th>
<th>#2: Use RWD to <strong>inform study design</strong> &amp; analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reg., scientific, &amp; ethical issues</td>
<td>Study of retrospective data only</td>
</tr>
<tr>
<td>Data sources &amp; study types</td>
<td>Prospective study with external data</td>
</tr>
<tr>
<td>Estimands (treat. effect) in RW setting</td>
<td>Causal inference issues in reg. setting</td>
</tr>
<tr>
<td>Confounding control</td>
<td></td>
</tr>
</tbody>
</table>
### WS2 Inform Clinical Study Design: Issues of Using Retrospective Data

<table>
<thead>
<tr>
<th>Issues</th>
<th>Description</th>
<th>Comments</th>
</tr>
</thead>
</table>
| **Heterogeneity in data sources** | • Levels: aggregated data (AD) vs. individual patient data (IPD)  
• Purposes: RCT (masked vs not), single-arm, registry, claims, etc. | How to account for differences e.g., definitions, uncertainties, potential confounders, data collection frequencies |
| **Challenges**          | • Heterogeneous data *not* generated for research purposes  
• Unmeasured confounders | • When is “similar enough”?  
• RWD ≠ historical RCT (quality)  
• Selection bias |
| **Analysis Methods**    | Lit. review → Meta-Analysis (MA)  
→ Network MA → Multi. NMA | Evolving research areas |
| **Data quality**        | Numerous curation guidelines | Fit-for-purpose requirements |
## WS2 Inform Clinical Study Design: Prospective Study with External Data & RWD

### Roles of External Data & RWD

<table>
<thead>
<tr>
<th>External control</th>
<th>External borrowing</th>
<th>Synthetic evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serving as control in single-arm studies</td>
<td>Augmenting control arm in RCTs</td>
<td>Combining historical, con-current external data</td>
</tr>
</tbody>
</table>

### Methods

<table>
<thead>
<tr>
<th>External control</th>
<th>External borrowing</th>
<th>Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian approaches</td>
<td>Bayesian dynamic borrowing, e.g.:</td>
<td>- Suboptimal sample size estimation</td>
</tr>
<tr>
<td>Treatment modeling, e.g., propensity score (PS)</td>
<td>- Hierarchical models</td>
<td>- All outcome-blinded thus far</td>
</tr>
<tr>
<td>Various matching metrics &amp; methods</td>
<td>- Power priors</td>
<td>- Large &amp; diverse control pool</td>
</tr>
<tr>
<td></td>
<td>- Commensurate priors</td>
<td>- Prospectively specified analysis plan, including matching</td>
</tr>
<tr>
<td></td>
<td>- Robust MA priors</td>
<td>- How to leverage double-robust methods in submissions?</td>
</tr>
<tr>
<td></td>
<td>PS-based augmentation with multiple controls</td>
<td></td>
</tr>
</tbody>
</table>

- Various matching metrics & methods
### Examples

<table>
<thead>
<tr>
<th>Condition</th>
<th>Product</th>
<th>Control</th>
<th>1° Endpoint</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-cell precursor acute lymphoblastic leukemia (ALL) in 1(^{st}) or 2(^{nd}) complete remission with minimal residual disease (MRD) ≥ 0.1% ((\text{Gökbuget et al. 2018}))</td>
<td>2018 Blincyto: 1(^{st}) approved therapy for MRD positive for ALL</td>
<td>Pooled historical data set from Europe &amp; the US</td>
<td>Complete remission</td>
<td>① 189 single-arm treated subjects vs. 694 control by stabilized IPTW with covariates: age, sex, duration between initial diagnosis &amp; salvage therapy, region, prior HSCT, prior # of salvage therapies, 1° refractory &amp; in 1(^{st}) salvage, refractory to last salvage therapy.</td>
</tr>
<tr>
<td>Male with hormone receptor (HR)-positive, HER2-negative advanced or metastatic breast cancer</td>
<td>2019 Ibrance: Expanded to from F to M patients</td>
<td>Pooled historical data set from Europe &amp; US</td>
<td>Safety profile</td>
<td>② “Based on limited data from postmarketing reports and electronic health records, the safety profile for men treated with IBRANCE is consistent with the safety profile in women treated with IBRANCE.”</td>
</tr>
<tr>
<td>Symptomatic heart disease due to severe native calcific aortic stenosis at high or greater risk for open surgical therapy ((\text{Thourani 2016}))</td>
<td>2016 SAPIEN 3 transcatheter aortic valve replacement</td>
<td>938 patients in the open-heart surgery arm of PARTNER 2A trial</td>
<td>All-cause death, all stroke, aortic insufficiency ≥ moderate at 1yr</td>
<td>③</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table: Observed Event Rate</th>
<th>Propensity Score Quintile Pooled Proportion Difference (ATT Method) ([90% CI])</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAPIEN 3 ((N=1069))</td>
<td>PIIA-SAVR ((N=936))</td>
<td></td>
</tr>
<tr>
<td>13.0%</td>
<td>23.2%</td>
<td>-9.2% [-12.4%, -6.0%]</td>
</tr>
</tbody>
</table>

① Not from label; Gökbuget et al. (2018)  
② IBRANCE label (April 2019) go.usa.gov/xmpHe  
③ SAPIEN 3 label SSED (August 2016) go.usa.gov/xmp6g
Recap

• Regulatory evidence is context-specific. Thus, talk to FDA early & often.

• Generation & evaluation of RWE driven by regulatory question.

• RDF are fit-for-use to address the regulatory question.

• RWE generated by adequate study design, conduct, and analysis may help address intractable regulatory challenges and make clinical studies more patient-centric.

• Many exciting research opportunities in RWD/RWE awaiting ahead to help bridge the gaps of unmet medical needs!
Acknowledgements

Weili He for WS#1 slides
Mark Levenson for his slides on FDA RWE Framework Program
ASA BIOP RWE SWG members
ASA BIOP