The use of RWE for label expansion – Progress to date from the ASA BIOP RWE SWG Work Stream 1

Weili He, AbbVie, Mark Levenson, FDA
on behalf of the WS1 members
ASA BIOP RISW Workshop
September 25, 2019

Disclaimer

The comments provided here are solely those of the presenters and are not necessarily reflective of the positions, policies or practices of presenters' employers.

Acknowledgement – SWG and WS1 Team Members (in bold)

Industry	Affiliation	Academic/FDA†	Affiliation
Weili He §,1	AbbVie	**Martin Ho §,2**	CBER
Jie Chen	Merck	Telba Irony	CBER
Yixin Fang	AbbVie	Mark van der Laan	UC Berkeley
Doug Faries	Eli Lily	Hana Lee	CDER
Qi Jiang	Seattle Genetics	**Mark Levenson 1**	CDER
Kwan Lee 2	Janssen	Zhaoling Meng	BMGMR‡
Xiwu Lin	Janssen	**Pallavi Mishra-Kalyani**	CDER
Yang Song	Vertex Pharma. Inc.	**Frank Rockhold**	Duke
Hongwei Wang	AbbVie	Tingting Zhou	CBER
Roseann White	The Third Opinion	**Ben Goldstein**	Duke
Richard Zink	Target Pharma. Solution		

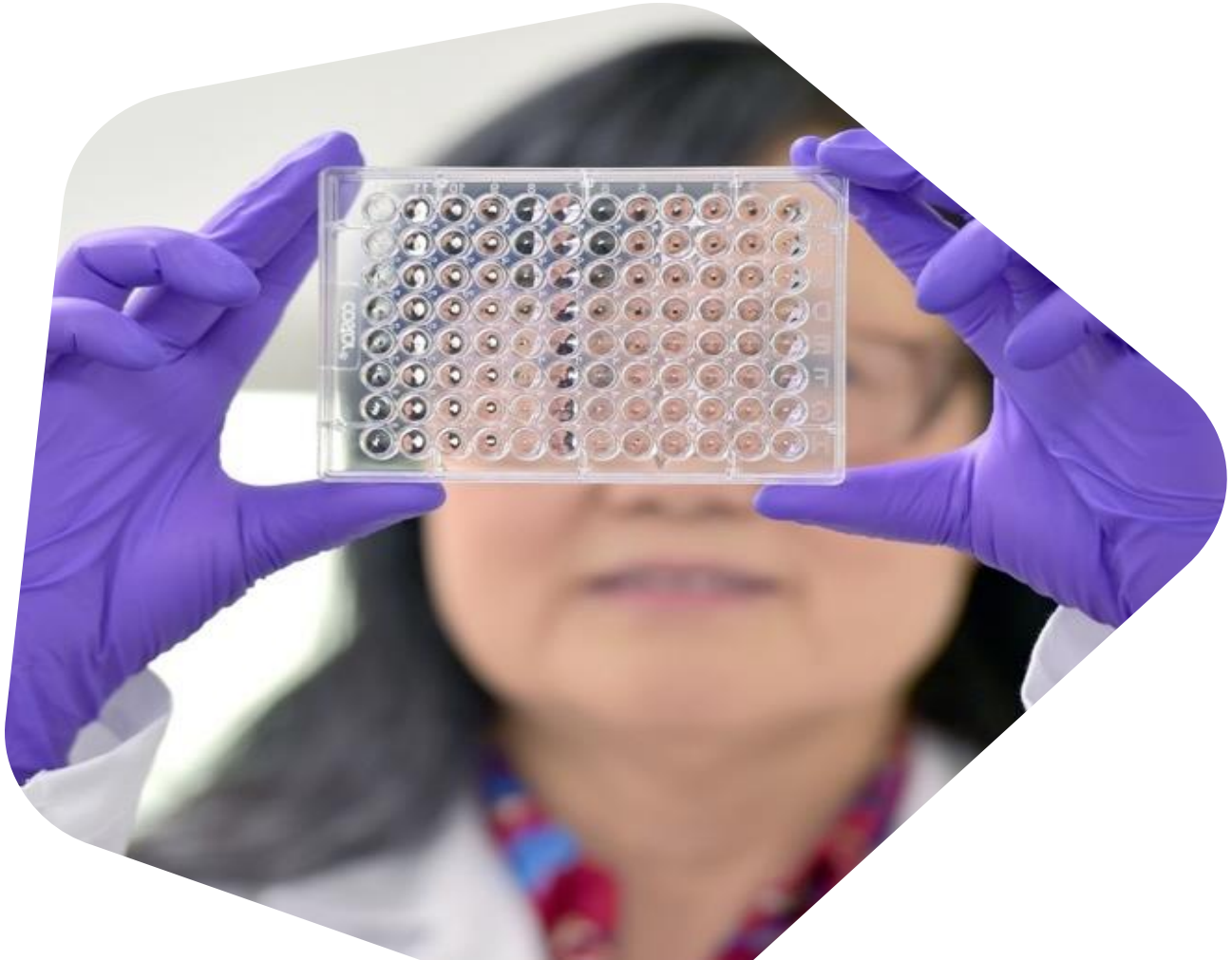
§ Co-chairs of the SWG; 1 Co-leads of Workstream #1; 2 Co-leads of Workstream #2

† Liz Stuart (Johns Hopkins University) participates as non-member; ‡ Bill & Melinda Gates Medical Research Institute

Outline

- Key topics identified in the use of RWD for label expansion
- Brief overview of each of key topics
- Future research

Key topics identified in the use of RWD for label expansion



1. Regulatory , scientific, and ethical issues
2. Treatment effects and estimands in RW setting
3. Study types
4. Data sources
5. Outcome measures
6. Confounding control

Treatment effects and estimands in RW setting

- Estimand as defined in draft ICH E9 (R1) addendum
  - Four attributes of estimand (possibly with one more attribute)
    A. the **population**, that is, the patients targeted by the scientific question
    B. the variable (or **endpoint**), **to be obtained for each patient**, that is required to address the scientific question
    C. the specification of **how to account for intercurrent events** to reflect the scientific question of interest.
    D. the **population-level summary for the variable** which provides, as required, a basis for a comparison between treatment conditions
- Much more complicated in RW studies related to adherence issue

Treatment effects and estimands in RW setting (Cont'd)

Attributes	RCT	RW Studies (sample cases below)
Population	detailed incl/excl; more homogeneous population and care following protocol	broad and heterogeneous population from routine clinical practice, which may vary related to local reimbursement (treatment decision is typically ahead of study participation), data sources, and methods to control confounding
Endpoints	well-defined and specifically collected for the study	under-reporting or lack of disease-specific clinical outcomes exist; measurement definition or algorithm may be associated with suboptimal specificity and sensitivity; information bias can be high; leverage of unstructured data via NLP is not common practice.
Intercurrent Events	extensive efforts are made to ensure patients f/u and data completeness	frequent for long-term f/u where treatment change is common and reasoning typically not well-recorded; treatment adherence tends to be low as medication is not provided; proportion of missing data and loss to f/u can be high due to non-interventional nature and data collection mechanism
Reporting measures	mostly focusing on population-level group comparison for effect size	very diverse measurements depending on research questions, for example, prevalence/incidence, disease progression, treatment pattern, disease burden, population-level comparative effectiveness, patient-level treatment response.

Consideration of Strategies for Intercurrent Events (IE) in RW Setting

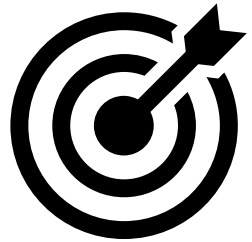
Strategy	Advantages	Disadvantages
ITT principle (treatment policy - Ignore IE strategy)	In randomized settings: upholds principles of randomization	When non-compliance/crossover is high, may not capture treatment effect
Composite strategy (Occurrence of the IE "taken to be a component" of the endpoint)	Allows for more accuracy in characterizing data	If intercurrent events are unbalanced, may introduce bias in favoring or un-favoring certain components of the composite endpoint
Hypothetical strategy (Impute as if no IE strategy)	Simple/straightforward result	Assumes that intercurrent events are not a source of bias (similar to MCAR for missing data)
Principal stratum strategy (Exclude IEers strategy)	All intercurrent events are post-randomization and all methods deal with them	Modeling assumptions subject to misspecification; Can only estimate effect in stratum without intercurrent events, likely not realistic in RW setting
While on treatment strategy	Simple/straightforward approach	Only allows for a limited treatment effect (up until the time of intercurrent events), valuable information such as informative dropout may be ignored

Treatment effects and estimands in RW setting (Cont'd)

Key challenges and opportunities

- Population – Understand the heterogeneity of the underlying population with the disease of interest, which may reflect different aspects of general demographics, disease variation, and healthcare systems
- Variables/outcomes - Choice of variable/outcome will depend on the availability of patient data and the feasibility of collecting required information in an uncontrolled clinical setting
- Intercurrent event - Maybe common and not well-defined. Patients discontinuation, switch, add-on, or adherence to treatment in real-world clinical settings are not controlled. Change of health insurance plan, treating physicians or facilities further complicate the situation
- Summary measure - Diverse measures for summary of real-world data, each corresponding to different research questions
- Most importantly, strategies on how to deal with intercurrent event in RW setting needs to be well thought out. Approaches beyond what are currently proposed in ICH E9 (R1) addendum may be explored.

Study Types



Study types

- Intervention vs. non-intervention
  - Interventional study: participants are assigned to receive one or more intervention/treatment (or no intervention) so that researchers can evaluate the effects of the interventions on biomedical or health-related outcomes
  - Non-interventional study: no form of intervention but observational; patient data is gathered during routine treatments and evaluated using outcome measures collected in routine clinical care setting.
- Randomized vs non-randomized
  - Randomized study: participants are allocated at random (by chance alone) to receive one of several clinical interventions
  - Non-randomized study: participants are not assigned by chance to different treatment groups

Study Types (Cont'd)

Study types (Cont'd)

- Prospective vs retrospective, hybrids
  - Retrospective observational: study identifies the population and determines the exposure/treatment from historic data; the variables and outcome of interest are determined at the time of the study is designed
  - Prospective observational: the population of interest is identified at the start of the study, and exposure/treatment and outcome data are collected from that point forward
  - Hybrids: an observational study that may have both retrospective and prospective data components
- Pragmatics vs. traditional RCTs, and hybrids
  - Traditional RCTs: a study that is usually supported by a research infrastructure that is largely separated from routine clinical practice and is designed to control variability and maximize data quality
  - Pragmatic trials: designed to evaluate the effectiveness of interventions in real-life routine practice conditions,
  - Hybrids: certain elements of a clinical trial could rely on collection and analysis of RWD, such as medical claims, EHRs, etc.

Study Types (Cont'd)

Key challenges and opportunities

- In the US, the standard for effectiveness of drugs is based on the substantial evidence criteria.
- Substantial evidence is established based on the findings of adequate and well-controlled investigations (AWC).
- Valid control group is generally required to minimize bias in subjects' perception and observers' assessment of responses
  - Concurrent control and randomization is typically used
  - Historical or external controls can be used in diseases with high and predictable mortality (e.g., certain malignancies) and studies in which the effect of the drug is self-evident (general anesthetics, drug metabolism)
- The 21st Century Cures legislation specifically states that the act does not change the evidentiary standards.
  - FDA exercises flexibility in practice, while adhering to statue and regulation
  - It is not clear whether a purely observational study can fit the regulatory definition of AWC.
  - Single arm interventional studies with external controls may be AWC under some circumstances.

Data Sources

There are two major sources of RWD:

- **Experimental data sources**
  - data from hybrid clinical trials
  - pragmatic trials
  - external RWD used as a control for non-randomized single arm clinical trials (Rockhold and Enas 2011)

	Hybrid clinical trial	Pragmatic trial	External control using RWD
Description	Certain elements of a clinical trial relying on collection and analysis of RWD	Clinical trials with design elements closely resemble routine clinical practice	Use of RWD as basis for external control of a non-randomized, single arm clinical trial
Example		Salford trial studying an experimental drug vs. usual care in asthma and COPD in routine clinical care setting	Blinatumomab vs. historic standard therapy of adult relapsed and refractory acute lymphoblastic leukemia

Data Sources (Cont'd)

- **Non-Experimental data sources**
  - non-experimental data primarily collected for research purposes, such as surveys, large cohort studies, and registries
  - transactional real-world data created by the routine operation of the US healthcare system, such as health insurance claims or electronic health records (EHRs), among others

	Category I: Research data sources	Category II: Transaction data sources
Description	Data collected primarily for research	Data used secondarily for research

Examples	• Data specifically for study purpose
	  ○ Framingham Heart Study
	  ○ Cardiovascular Health Study
	• Data intended for other studies
	  ○ Nurses' Health Study
	  ○ Some registries
		• Clinical documentation
		  ○ Electronic health records
		  ○ Wearable devices
		• Administrative
		  ○ Claims data
		  ○ Geocoding/census

Data Sources (Cont'd)

Key considerations for the evaluation of fit-for-purpose RWD

- Relevance
  - Exposure
  - Outcomes
  - Covariates
- Quality of data
  - RWD collected for resource purposes are generally more fit-for-research question
  - Transactional data sources, even with high quality, may lack certain key data points and not fit-for-purpose
- Precision in data ascertainment
  - Variables definitions
  - Method and consistency of ascertainment
  - Magnitude of missingness in real clinical care setting

Data Sources (Cont'd)

Key considerations for the evaluation of fit-for-purpose RWD (Cont'd)

- Measurement bias
  - It's unrealistic to assume that all variable are perfectly measured
  - Measurement bias results when the association between treatment and outcome is weakened or strengthened as a result of the process by which the study data are measured (measurement errors)
- Generalizability
  - Use experimental data source or data collected primarily for research purposes, generalizability is how we can generate the study population considered for the primary research question to other specific population or more general population.
  - Use non-experimental transactional data sources, generalizability depends on fit-for-purpose evaluation of data source and quality, leading to robust real-world evidence.

Data Sources (Cont'd)

Key challenges and opportunities

- Linking of different data sources
  - Methods to address duplication of patient information in different data sources
  - Understand the definitions of key variables
  - Link data associated with a single patient across different data sources either definitively or probabilistically
- Data standard and data system inter-operability
  - Follow a patient across sites of care, by different providers, and with different payers
  - Use of standardized data format and system inter-operability
- Use of quantitative approach to assess fit-for-purpose data sources, targeting to specific research questions

Outcome Measures

Fit-for-purpose outcome measures in RW setting

- Identified to answer the research question and relevant to patients, clinical practice, and treatment choices
- Defined based on the estimand framework with regard to
  - population,
  - intervention as appropriate (or exposure or treatment in RW setting); Consider the strategies in handling intercurrent events in the definition of outcome measures
  - outcome variables by which patients' health are expected to improve over the course of study

Outcome Measures (Cont'd)

Ascertainment of outcome measures

- Consider alternative variables or surrogate outcomes, if patient-important outcome variables cannot be ascertained
- Consider the use of composite endpoints as a way to handle intercurrent events, if no single patient-important outcome variable could be identified
- Consider the use of text analytics or natural language processing (NLP) to ascertain important outcome variables that are captured in unstructured data sources

Outcome Measures (Cont'd)

Key challenges and opportunities

- Define outcome measures that are meaningful and closely relate to clinical outcome of interests
- Ensure that available RWD sources capture the information or collect data prospectively
- Understand and make use of machine learning and text analytics to augment the information

Confounding Control

Confounding Biases

- Refers to a distortion in the estimated effect of exposure on outcome due to the presence of another variable.

Type of Confounding—Measured or unmeasured

- Residual confounding—The distortion that remains after controlling for confounding in the design and/or analysis of a study– usually unmeasured
- Confounding by indication—A type of confounding that occurs when a symptom or sign of disease is judged as an indication (or a contraindication) for a given treatment, and is therefore associated both with the use of a drug or medical procedure (or its avoidance) and with a higher probability of an outcome related to the disease for which the agent is indicated (or contraindicated)
- Reverse causality occurs when exposure is related to outcome
- Time-varying confounding occurs when confounders have values that change over time

Confounding Control (Cont'd)

Control of confounding

- Confounding can be controlled at
  - Design stage, e.g.
    - Randomization, matching, population restriction,
  - Analysis stage, e.g.
    - Restriction and stratification, standardization and multiple regression, matching, G-methods, disease risk scores, machine learning
- Control of confounding
  - 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

Confounding Control (Cont'd)

Key challenges and opportunities

- Limited experience regarding quantification of residual bias after performing bias reduction analysis
- Need more education and influence for better and more appropriate use of confounding adjustment methods
- Require better understanding of the operating characteristics of these methods under certain confounding circumstances
- There is no consensus or statistical guidance on the best practice for control of confounding in real-world studies for practitioners

- **Future research**

Future Research

- Continue to look into the challenges and opportunities in each of the topics we identified, and propose recommendations
- The SWG has identified additional research topics and will be working on them after the landscape papers, such as
  - Determine estimand in the setting of complex exposure scenarios
  - Explore principles and approaches for the evaluation of fit-for-purpose data sources utilizing quantitative approaches

References

- Hernan, M. A., and Robins, J. M. (2019), *Causal inference*, Boca Raton: Chapman & Hall/CRC.
- Rockhold, F. W., and Enas, G. G. (2011), “10 Years with ICH E10: Choice of Control Groups,” *Pharmaceutical Statistics*, 10, 407–409. https://doi.org/10.1002/pst.500.
- Schneeweiss, S., and Avorn, J. (2005), “A review of uses of health care utilization databases for epidemiologic research on therapeutics,” *Journal of Clinical Epidemiology*, 58, 323–337. https://doi.org/10.1016/j.jclinepi.2004.10.012.

Thank You!