

Using Real World Data in Support of Regulatory Submissions

Lorraine Fang, Juanyao Huang, Shuang Zhang, Jenny Han, Michelle Li,
Joy Wang

Bristol Myers Squibb, Berkeley Heights, New Jersey, USA

Abstract

Real world data is increasingly critical, particularly in supporting regulatory submissions of single arm study and rare diseases. In recent two years, we successfully created two synthetic cohorts to compare the efficacy between standard care and car-t therapies. The two packages were well accepted by FDA and EMA and they were acknowledged as a great add-on value to the entire submission packages.

This presentation will provide an overview of special features of RWE studies and our real working experience of these two RWE studies that we have accomplished. Following content will be covered: the differences between RWE study and regular clinical trial, data sources that we used, data collection, harmonization and sharing process, creation of firewall among different function groups, missing data imputation, propensity score calculation and matching, derivation of efficacy endpoints, statistical methodologies, content and data format of final submission package, the roles of statistical programming.

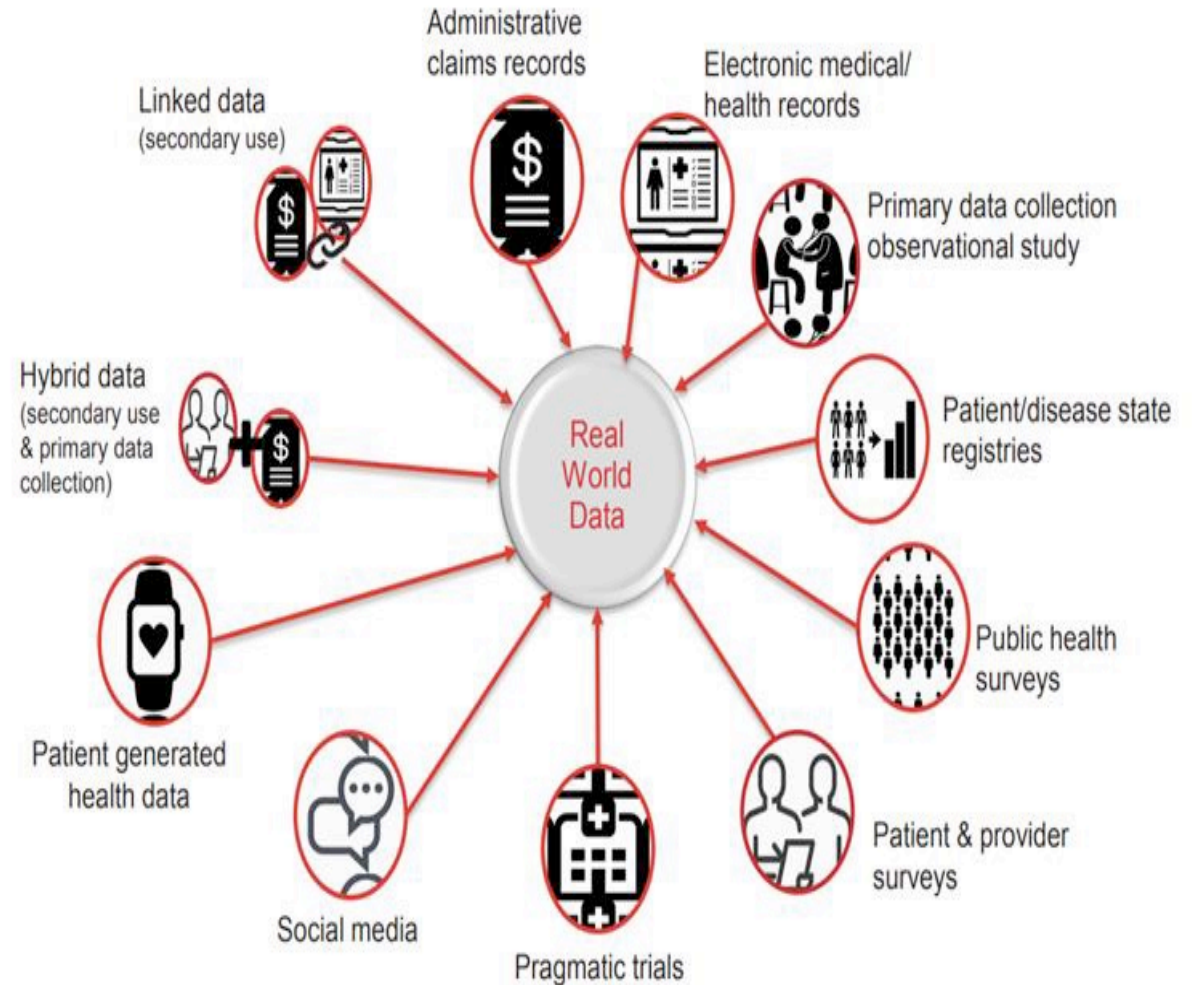
In addition, we will discuss the challenges that we encountered in the two RWE studies and the solutions that we proposed.

What is Real-World Data?

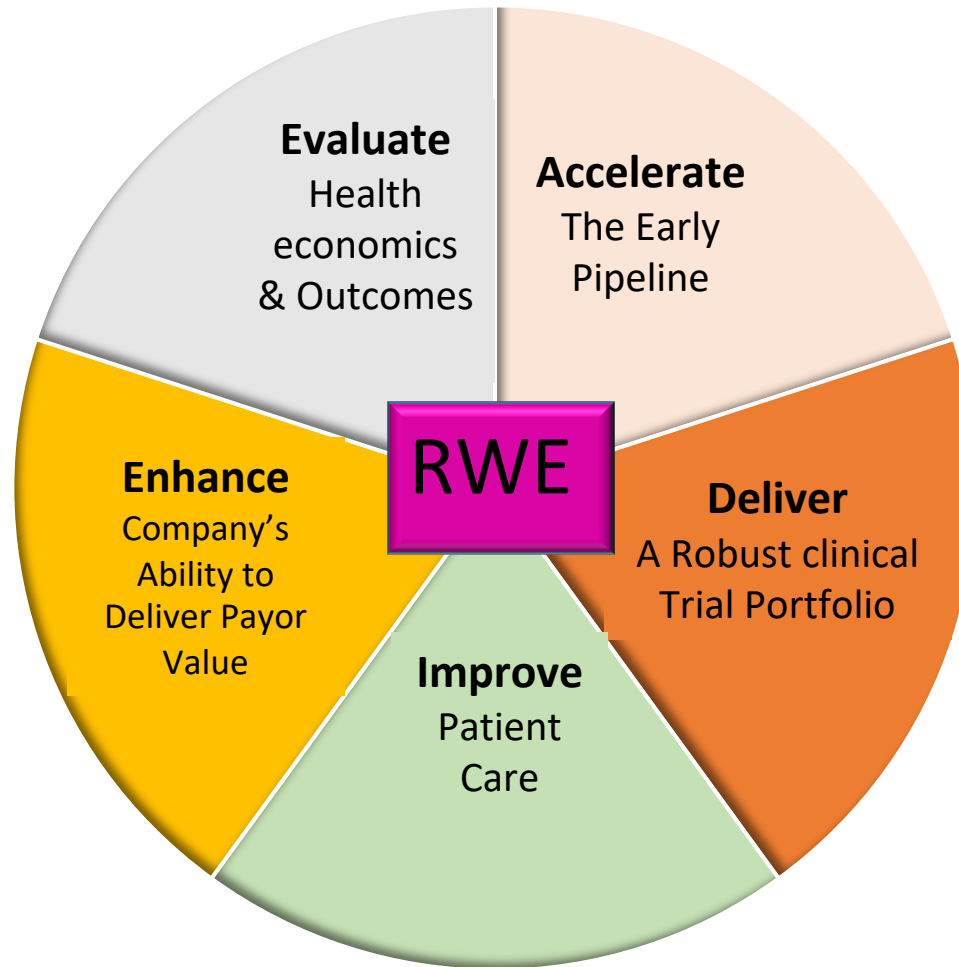
- **RWD:** The FDA defines **real-world data** as healthcare information derived from multiple sources **outside of typical clinical research settings**. This data is collected, de-identified, and stored in a variety of sources to be later analyzed.
- **RWE:** is evidence about the benefits and risks of a product, derived from analysis of RWD

Real World Data Sources

- EHRs, Registries, case studies, past clinical trial data
- RWD quality check for
 - Accuracy
 - Completeness
 - Provenance
 - Transparency of data processing



RWE is increasingly used to support a wide variety of business needs



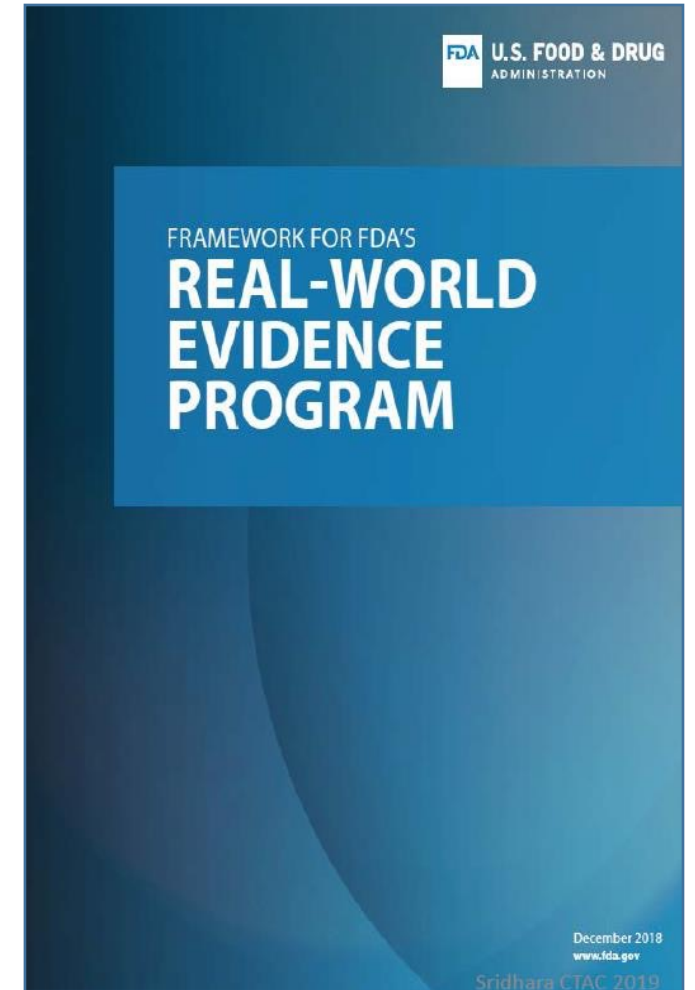
Background & Motivation of Using Synthetic Cohorts

- Randomized Controlled Trials (RCTs) are the gold standard
 - Used to evaluate efficacy and safety of treatments
 - Good randomization corrects for confounders and biases
- Barriers to randomization
 - Disease rarity
 - Scarcity of patients
 - Ethical & feasibility considerations
 - Temporal & financial costs

FDA's RWE Program

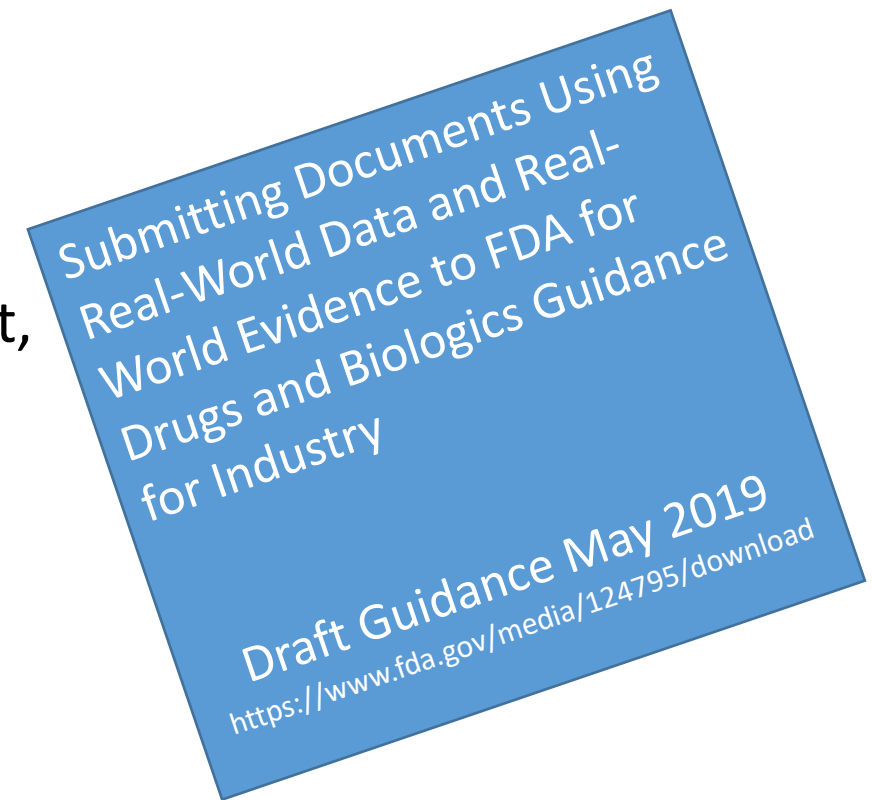
External Controls (Synthetic Control Arm)

- Not part of the original sample randomized into treatment arms
- Augment randomized control arms in RCTs (“hybrid” control arms)
- Help interpret data from single arm trials



When to Use External Controls

- **Single-arm Studies**
 - Uncontrolled studies require context to interpret outcomes
 - Clinical settings with no standard of care treatment, limited options
- **Orphan & rare diseases**
 - Difficult recruitment of patients
- **Consider for**
 - Diseases with highly predictable courses and mortality
 - Studies in which the drug effect is self-evident
 - Situation when the effect of treatment is substantial

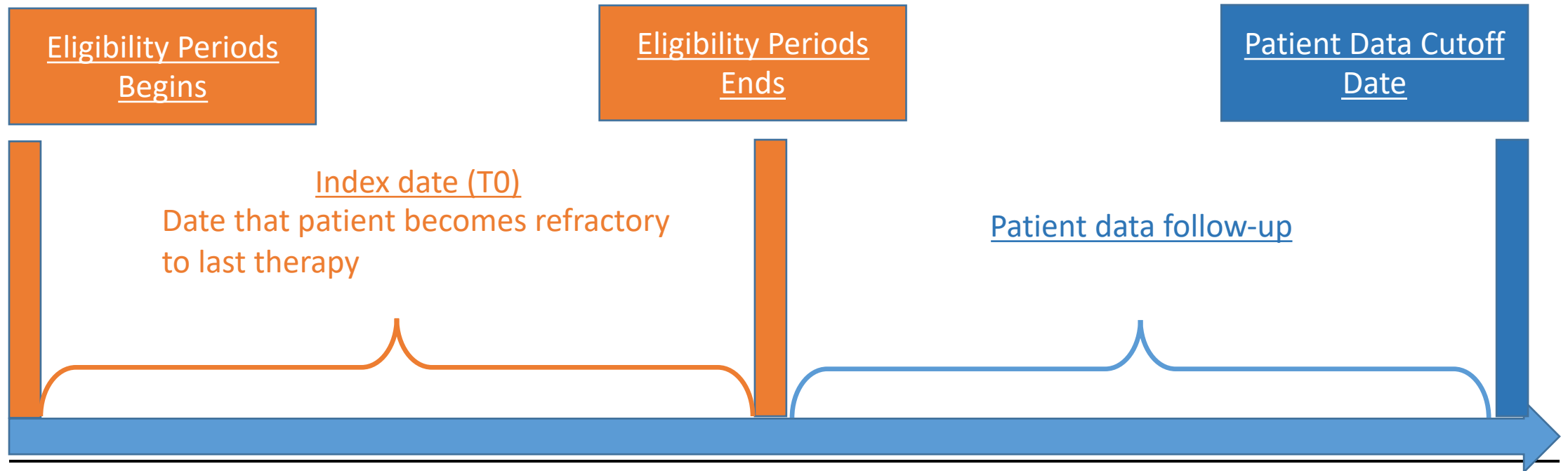


Case Study:

- A global, non-interventional, retrospective, multi-center study to generate **real world evidence** of subjects with relapsed and refractory **multiple myeloma** (RRMM) with prior exposure to an anti-CD38 antibody

Study Design

- Real world subjects were selected from a larger cohort with characteristics similar to subjects in clinical trial
- Study Timeline



Study Objectives

- To assess treatment patterns in real-world RRMM patients with characteristics similar to the treated population in clinical study
- To compare outcomes of standard-of-care therapy in a synthetic cohort vs treatment in clinical study

Study Analysis Populations

Real-World Analysis Population	Clinical Study Analysis Population	Note
Eligible RRMM to be balanced to Treated Population	Treated Population	Primary analysis
Eligible RRMM to be balanced to Enrolled Leukaphesed Population	Enrolled Leukaphesed Population	Sensitivity analysis
Eligible RRMM to be balanced to Lymphodepleting Chemotherapy Population	Lymphodepleting Chemotherapy Population	Sensitivity analysis
Eligible RRMM to be balanced to Efficacy Evaluable Population	Efficacy Evaluable Population	Sensitivity analysis

Study Efficacy Endpoints

Endpoints	Definition	Response Assessment by	Dataset Names
ORR (Primary)	Proportion of subjects with PR or better	Investigator or treated physician	ADRS
VGPR (Key Secondary)	Proportion of subjects with VGPR or better		
Other Secondary Effectiveness Endpoints			
TTR	Time to PR or better for responders	Investigator or treated physician	ADTTE
DOR	Time from first response to progressive disease or death, whichever occurs earlier		
PFS	Time from infusion* to progressive disease or death, whichever occurs earlier		
OS	Time from infusion to death due to any cause	NA	ADTTE

RWD Data Sources Used

Clinical Sites

Data collected by vendors using
eCRFs

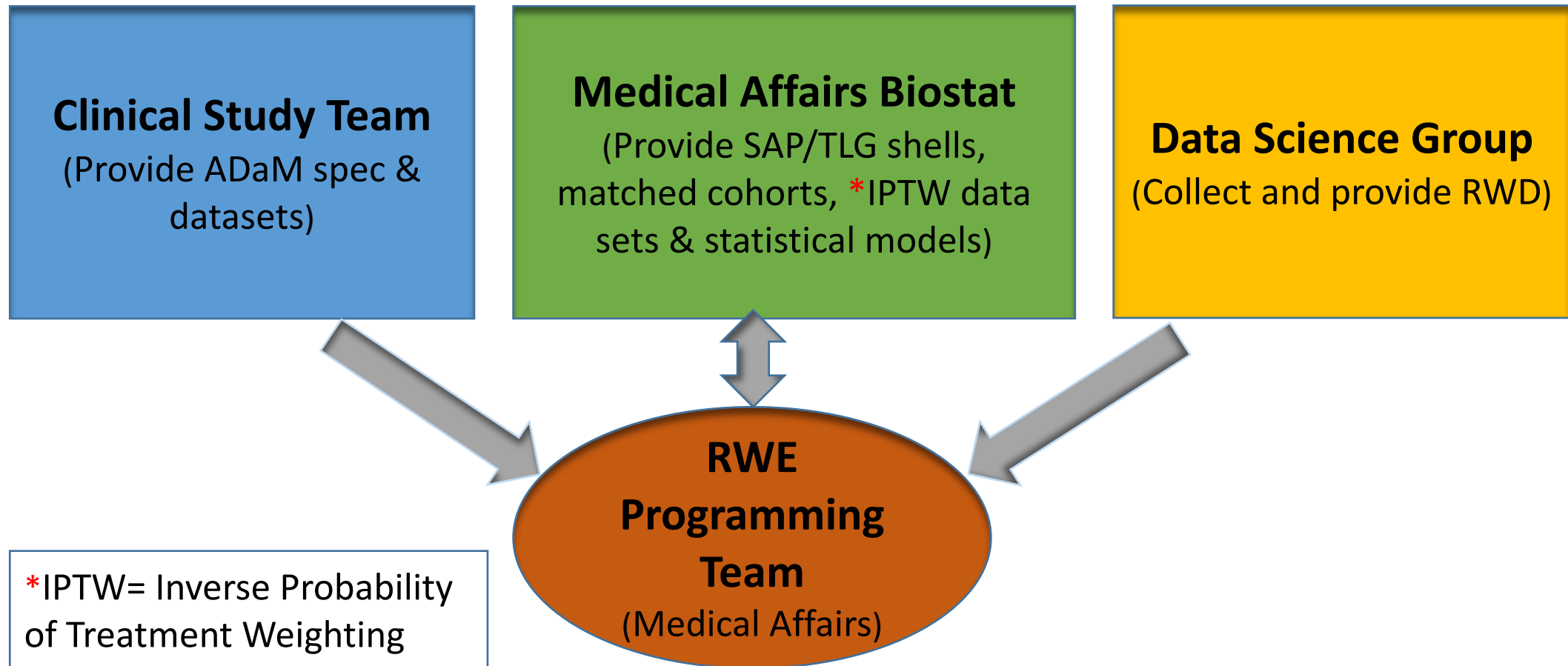
Registry Studies

Data from the CONNECT
MM Registries are
extracted

Research Databases

Flatiron
COTA
GRN

Data Processing Function Groups



RWE Programming Steps

- **Step 1: Prepare RWE ADaM Spec**

- Work with clinical team and Biostat to identify confounders
- Request ADaM Spec and test datasets (**blinded ADaM**) from clinical study with outcome variables masked
- Review RWD spec and test data files (**blinded**) from data science group
- Instruct and oversee data science group on baseline variables and efficacy endpoint derivation
- Draft RWE ADaM spec

RWE Programming Steps (con't)

- **Step 2: Generate Blinded ADSL**

- Request production **blinded** data transfer from clinical study and data science group
- Harmonize and re-structure RWD
- Assess RWD outliers together with Biostat and clinical team
- Apply baseline window for each confounder and comorbidities
- Apply additional exclusion and inclusion criteria based on protocol and SAP
- Generate **blinded** RWE ADSL and provide it to Biostat for propensity score calculation, matching and creation of IPTW data sets

RWE Programming Steps (con't)

- **Step 3: Generate un-blinded ADSL and efficacy data**
 - Request **un-blinded** data transfer from clinical study team and data science group after propensity score calculation/matching is done
 - Update RWE ADSL by adding outcome related variables, such as death date, death flag, last known alive date, discontinuation reasons, etc.
 - Derive ADTTE, ADRS

RWE Programming Steps (con't)

- Step 4: Create TLGs and Submission Package

- Understand statistical models for efficacy tables and graphs provided by Biostat
- Merge ADTTE and ADRS with IPTW and matched data sets
- Generate tables and graphs
- Prepare final submission package (TLGs, SAS xpt data sets, define.xml, ADRG)

Primary Analysis

- Baseline prognostic variables considered for stabilized IPTW were identified and ranked by a scientific steering committee
- Generated 30 datasets by using multiple imputation for missing covariates
- For each dataset of the eligible RRMM cohort versus clinical study treatment cohort:
 - Calculated propensity scores (PS)
 - Weighted subjects using stabilized IPTW methodology
- Combined estimates from each of the 30 datasets using Rubin's rules

Submission Data Package

- RWD are not in SDTM format. They are not submitted
- ADaM eCRT package:
 - xpt file, ADSL, ADRS, ADTTE
 - Define.xml
 - Analysis Data Reviewer's Guide
 - RWD Data process flow and programming specifications

Key Challenges

- Items collected in research but not typically done by the community physician will not be present
- Identifying patients similar to those in the clinical trial
- Inclusion/exclusion criteria may not be identifiable from RWD, e.g. severity of co-morbidities
- Geographic difference may exist
- Genetics, standard of care, approved therapies
- Discrepancies with how often patients are seen in route practice compared to a clinical trials
- Quality of the data



Challenge Mitigations Used

- Pre-specified study protocol
- Pre-specified SAP
- Clearly pre-defined inclusion/exclusion criteria and used contemporaneous synthetic external cohort
- Pre-specified important prognostic and confounding variables
- Clearly pre-specified propensity score methodology for selection of variables and balancing/matching
- Use multiple imputation for data missingness
- Used firewall/masking of outcome to minimize bias
- Subgroup analysis
- Sensitivity analysis

Key Takeaways

- RCT is the best method to evaluate a treatment effect, if feasible
- When randomization is impractical, infeasible or unethical, external controls/synthetic control arm are recognized as a possible type of control arm
- If an external control arm is used to support regulatory submission, best practices should include:
 - Pre-determined patient selection
 - Pre-specified SAP: careful, detailed and transparent
 - Firewall/blinding to outcome data for biostatisticians while performing matching
 - Adherence to observational study principles to minimize bias and confounding and produce credible/reproducible RWE

Contact Information

Lorraine Fang

- Associate Director,
RWE Statistical Programming
- Email: lorraine.fang@bms.com

Thank you