RWD Reliability

James Joseph, EDA CLINICAL
Hands-On Workshop
PharmaSUG 2025
hackmd.io/@explore/SUG25

Real World Data

- electronic health records
- medical claims data
- product and disease registries
- patient-generated data

RWD

Data related to patient health status and/or healthcare delivery, collected from various sources **other than traditional clinical trials**.

> The key distinction is that RWD are not collected in a special research clinic or setting.
> ref

Real World Evidence

Clinical evidence about medical product usage and benefits/risks derived from RWD analysis

---

Trial Designs

Retrospective (comparative)
Prospective (non-randomized, e.g., ECA)
Prospective (randomized, e.g., pragmatic)

Benefits of RWD

Challenges of RWD

- data availability,
- linkages,
- timeliness,
- data accrual,
- quality and integrity,
- study purpose,
- specific data elements,
- generalizability of data,
- assessment of confounding,
- timing of data availability and
- completeness and accuracy of study sample reflecting the target population

---

Data Fitness ^

For all study designs, it is important to ensure the **reliability** and **relevance** of the data used to help support a regulatory decision.

`Reliability` ^

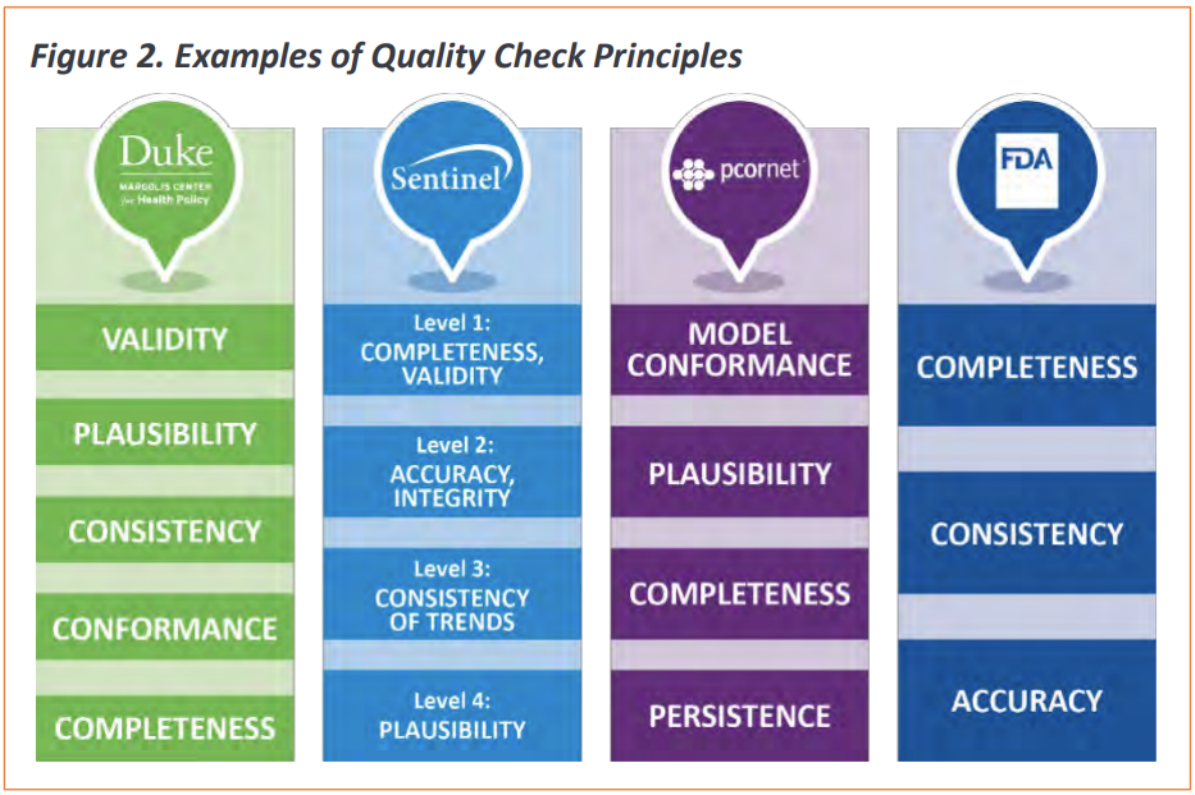
- accuracy,
- completeness, and
- traceability

`Relevance` ^

- **availability** of data for key study variables (exposures, outcomes, covariates) and
- sufficient **numbers of representative patients** for the study.

`Reliability`

"Data Quality”



*consistency –> traceability

Characterize the Data
^

The **format** and **traceability** of EHRs and medical claims data can vary significantly across health care entities (e.g., insurer, practice, provider, data vendor).

`Reliability` ^

- traceability
- completeness, and
- accuracy,

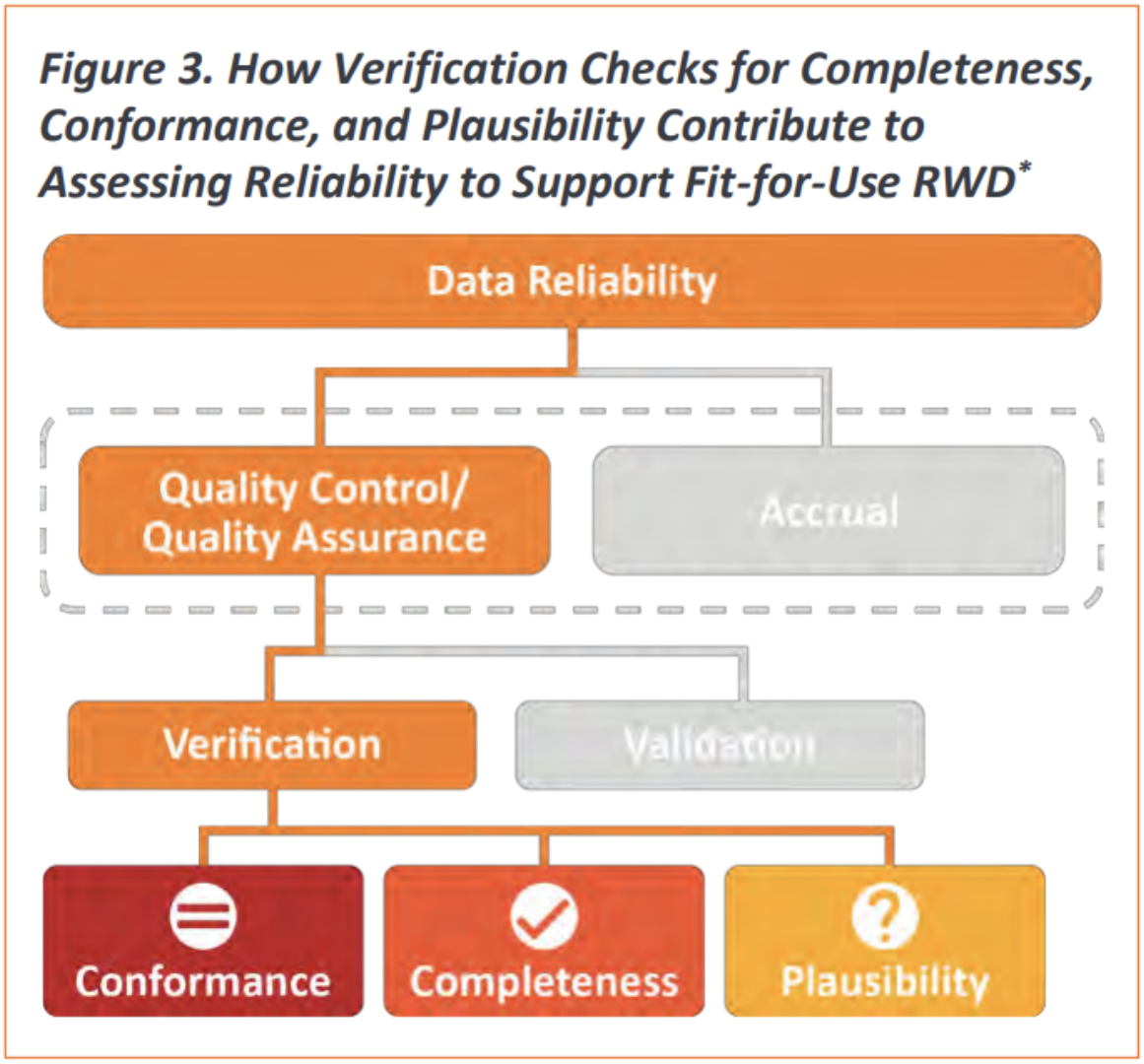
`Reliability: Traceability`

The study protocol and analysis plan should specify the traceability (curation and procedures used throughout the data life cycle) and describe how these procedures could affect the integrity of the data and the overall validity of the study.^

`Reliability: Traceability`

Examine data at each step in the data life cycle...

- characterize data **completeness**, **conformance**, and **plausibility** of data values,
- document QA/QC plan including data transformations
- define procedures to ensure integrity of the data.^



- plausibility –> accuracy

`Reliability: Traceability`

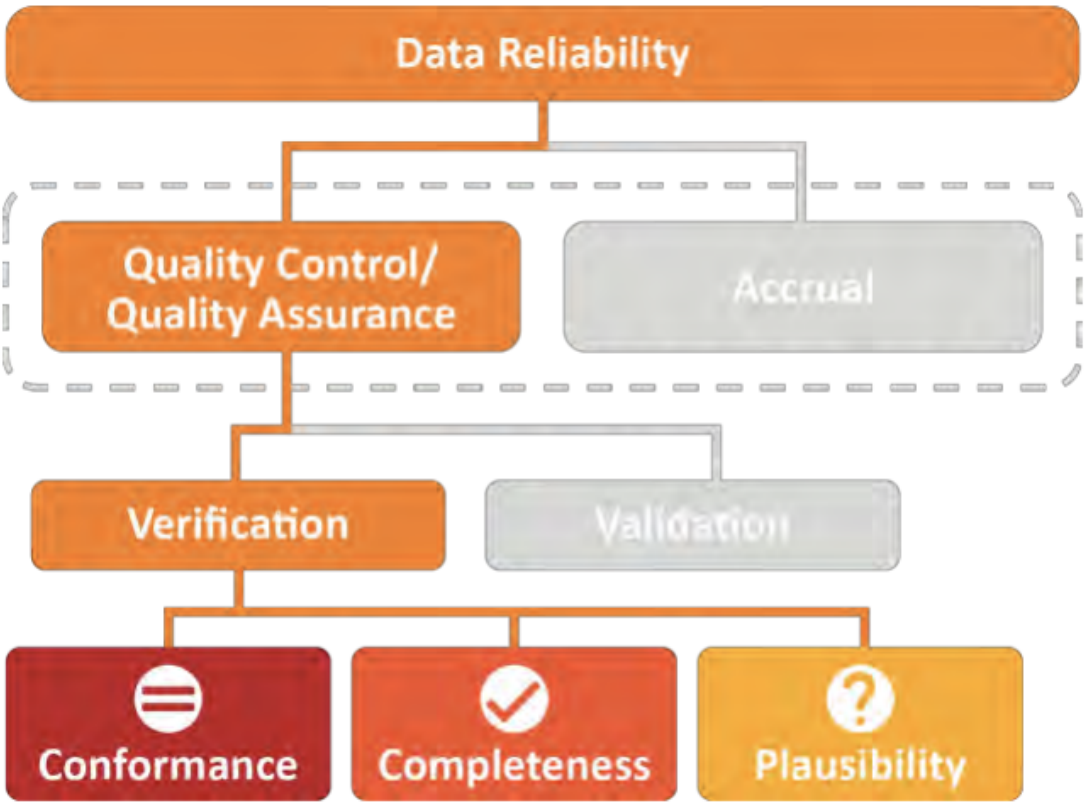
Conformance Checks ^

- [ ] Fields do not conform to data model specifications for data type, length, or name
- [ ] Fields contain values outside of data model specifications
- [ ] Fields have non-permissible missing values

- [ ] More than x% of the time, patient-reported or sensor data are assigned to more than one patient identification number
- [ ] Required fields in a table are not present
- [ ] Required tables in the prespecified data model are not present
- [ ] Expected tables in the prespecified data model are not populated
- [ ] Tables have primary key definition errors
- [ ] Tables contain orphan patient identification numbers
- [ ] Tables contain orphan encounter identification numbers
- [ ] Tables contain orphan provider identification numbers
- [ ] Replication errors between tables
- [ ] More than x% of ICD, IPT, LOINC, RXCUI, NDC codes do not conform to the expected length or content

---

Figure 3. How Verification Checks for Completeness, Conformance, and Plausibility Contribute to Assessing Reliability to Support Fit-for-Use RWD*



`Reliability: Completeness`

- frequencies without reference to data values
- amount of data required determined a priori

`Reliability: Completeness`

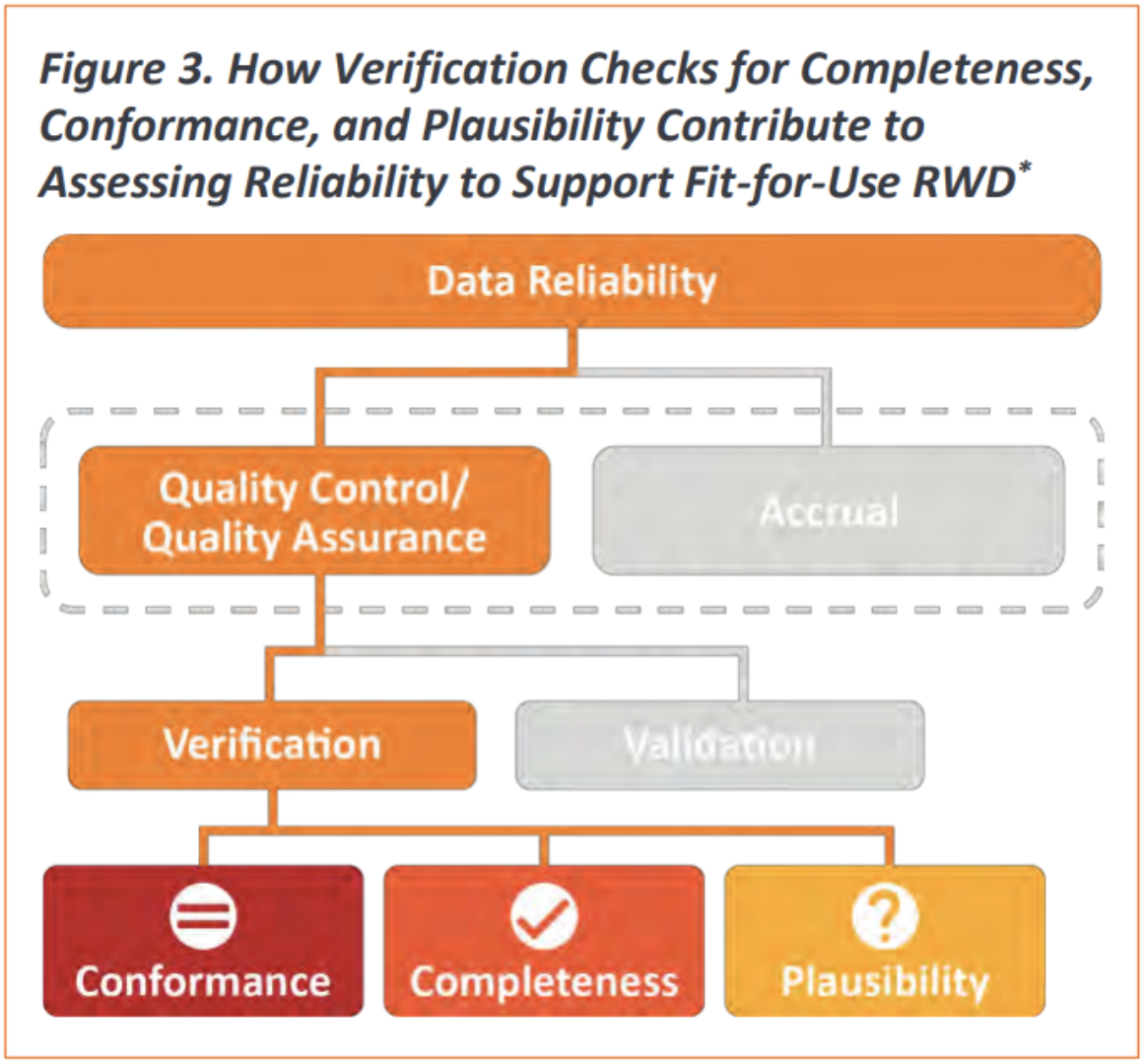
- [ ] Average number of diagnoses records with known diagnosis types per encounter is below the threshold ...
- [ ] Average number of procedure records with known procedure rates per encounter...
- [ ] More than x% of records have missing or unknown values for key variables
- [ ] Less than x% of patients with encounters have diagnosis records
- [ ] Less than x% of patients with encounters have procedures records
- [ ] More than x% of inpatient or emergency department to inpatient encounters with any diagnosis don't have a principal diagnosis
- [ ] Encounters, diagnoses, or procedures in an ambulatory, emergency department, emergency department to inpatient, or inpatient setting are less than x% complete three months prior to the current month
- [ ] Less than x% of prescribing orders are mapped to a RXNORM_CUI which fully specifies the ingredient, strength, and dose form
- [ ] Less than x% of laboratory results are mapped to LAB_LOINC
- [ ] Less than x% of quantitative results for tests mapped to LAB_LOINC fully specify the normal range
- [ ] Data are less than x% complete three months prior to the current month
- [ ] Less than x% of qualitative results for tests mapped to LAB_LOINC fully specify the source of the specimen and standardized unit for quantitative results

---

`Reliability`

~~Missing Data~~

Real world lab test finding:
(1) not ordered by HCP;
(2) ordered but not conducted;
(3) performed, but result was not stored; or
(4) test performed, result stored, but data lost



`Reliability: Accuracy`

Assessment of the believability or truthfulness of data values.

`Reliability: Accuracy`

Evaluate Plausibility by examining

- uniqueness of values
- range and distribution of values
  within a variable
- and two or more variables have an expected context-dependent relationship
- whether time-related and time-varying variables change as expected

---

`Reliability: Accuracy`

- [ ] More than x% of records have future dates
- [ ] More than x% of records are in the highest or lowest y% of biologically plausibility (e.g., age, height, weight)
- [ ] More than x% of patients have illogical date relationships (e.g., death date before birth date, dispense date occurs after death date)
- [ ] The average number of times data was reported per "y" time frame was greater than “x”
- [ ] More than x% of results for selected laboratory tests do not have the appropriate specimen source
- [ ] Median patient-reported or sensor data values for selected variables are statistical or clinical outliers
- [ ] The average number of principal diagnoses per encounter is above the threshold for inpatient and Emergency Department to inpatient
- [ ] The average number of patients increased or decreased by x% over pre-specified sequential time periods
- [ ] The average amount of data per patient increased or decreased by x% over pre-specified sequential time periods
- [ ] More than x% decrease in the number of records for ICD9 or ICD10 diagnosis or procedure codes or CPT/HCPCS procedure codes

---

Side B of Data Fitness

`Relevance`

"... the **availability** of key data elements (exposure, outcomes, covariates) and sufficient numbers of representative patients for the study."

`Relevance`

Challenges of RWD Design

`Relevance`

Over-Fitness

Given **availability** of RWD sources

- inability to randomize study participatnts to treatment arms

---

`Relevance`

When designs are tailored to the specific data available, results from those studies will suffer from being "overfit".

`Relevance`

Overfitness occurs when a study is designed based on patterns that are already apparent in the data (but not necessarily the population).

`Relevance`

Conclusions drawn from study designs that were derived from study data will lack scientific validity because they are not **reproducible or generalizable** to a population that is not represented by the data source(s).

`Relevance`

Avoiding Over-Fitness

(Too Relevant)

Crucial during planning stage:
(1) **predefine** study design and conduct; and
(2) collaborate with the FDA to assess **suitability** of non-interventional study design in the context of the research question ^

`Relevance`

Solution to Over-Fitness

For all key covariates, including confounders and effect modifiers, FDA recommends providing and justifying the validity of operational definitions in the protocol and study report.

`Relevance`

Solution to Over-Fitness

For all key covariates

- quantiative measure
- a priori
- sensitivity/specficity indices

Feasibility for RWD

A Structured Process to Identify Fit-for-Purpose Study Design and Data to Generate Valid and Transparent Real-World Evidence for Regulatory Uses^

`SPIFD2`

TOP

VIEW ONLINE

A Structured Process to Identify Fit-for-Purpose Study Design and Data to Generate Valid and Transparent Real-World Evidence for Regulatory Uses

Gatto, Nicolle M; Vititoe, Sarah E; Rubinstein, Emily; Reynolds, Robert F; Campbell, Ulka B

ISSN: 0009-9236 , 1532-6535; **DOI:** 10.1002/cpt.2883
Clinical pharmacology & therapeutics : CPT. , 2023, Vol.113(6), p.1235-1239

SPIFD2

MINI-REVIEW

A Structured Process to Identify Fit-for-Purpose Study Design and Data to Generate Valid and Transparent Real-World Evidence for Regulatory Uses

Nicolle M. Gatto[1,2,3,*], Sarah E. Vititoe[1], Emily Rubinstein[1], Robert F Reynolds[3,4] and Ulka B. Campbell[1,2]

Generating evidence from real-world data requires fit-for-purpose study design and data. In addition to validity, decision makers require transparency in the reasoning that underlies study design and data source decisions. The 2019 Structured Preapproval and Postapproval Comparative Study Design Framework to Generate Valid and Transparent Real-World Evidence (SPACE) and the 2021 Structured Process to Identify Fit-For-Purpose Data (SPIFD)—intended to be used together—provide a step-by-step guide to identify decision grade, fit-for-purpose study design and data. In this update (referred to as "SPIFD2" to encompass both the design and data aspects) we provide an update to these frameworks that combines the templates into one, more explicitly calls for articulation of the hypothetical target trial and sources of bias that may arise in the real-world emulation, and provides explicit references to the Structured Template and Reporting Tool for Real-World Evidence (STaRT-RWE) tables that we suggest using immediately after invoking the SPIFD2 framework. Following the steps recommended in the SPIFD2 process requires due diligence on the part of the researcher to ensure that every aspect of study design and data selection is rationalized and supported by evidence. The resulting stepwise documentation enables reproducibility and clear communication with decision makers, and it increases the likelihood that the evidence generated is valid, fit-for-purpose, and sufficient to support healthcare and regulatory decisions.

`SPIFD2`

2019 Structured Preapproval and Postapproval Comparative **Study Design Framework** to Generate Valid
and Transparent Real-World Evidence (SPACE)

2021 Structured Process to Identify **Fit-For-Purpose Data**
(SPIFD)

Feasibility for RWD

A Structured Process to Identify Fit-for-Purpose Study Design and Data to Generate Valid and Transparent Real-World Evidence for Regulatory Uses^

`SPIFD2`

Why?

1. Articulate a specific research question
2. Identify fit-for-purpose data
3. Identify study design
4. Support study design and data selection with evidence
5. Ensure documentation enables reproducibility and clear communication with decision makers

---

`SPIFD2`

How?

1. State Research Aim, Question, and Objectives
2. Describe Hypothetical Target Trial (HTT)
3. Describe RWD Study Emulation of HTT
4. Identify Fit-for-Purpose RWD Source
5. Document Final RW Operationalization, Rationale, Validity Concerns, and Approaches to Address These Concerns

---

`SPIFD2`

Scoring RWD Feasibility
by Data Elements

Score	Description
1	Data requirements are not met.
3	Some data requirements are met.
5	All or nearly all data requirements are met.

`SPIFD2`

Scoring RWD Feasibility
by **Budget Constraints**

Score	Description
$	Well Below Budget
$$	Within Budget
$$$	Well Above Budget

`SPIFD2`

Scoring RWD Feasibility
by **Timeline Constraints**

Score	Description
Slow	May cause delays in timeline.
Medium	Expected to meet timeline constraints.
Fast	Expected to be delivered well before timeline constraints (expedited).

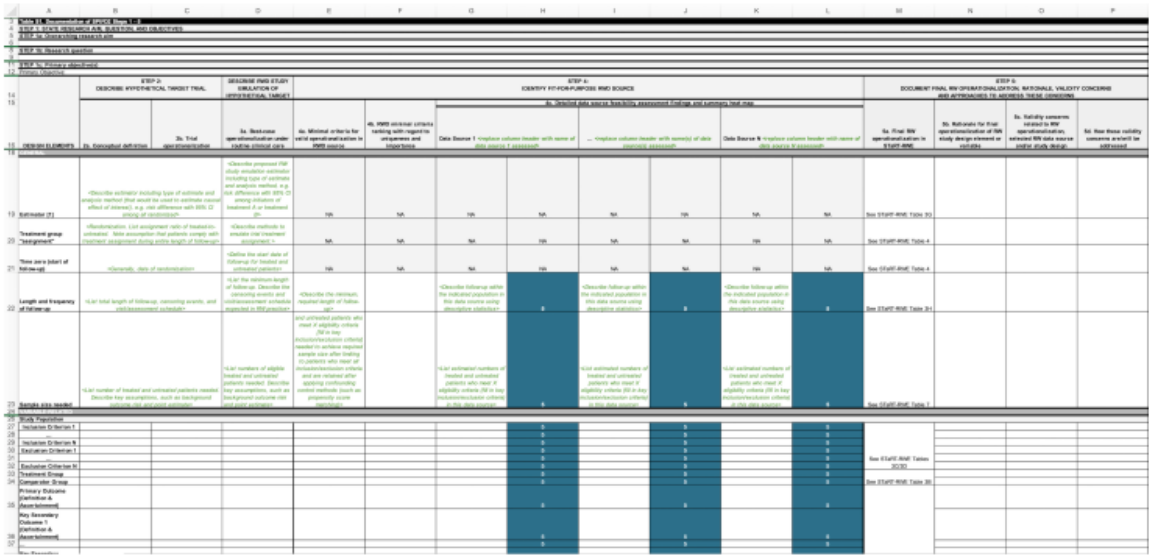


Figure 1 Screenshot of SPIFD2 **Table S1** (partial view) showing merged SPACE+SPIFD steps. SPIFD, Structured Process to Identify Fit-For-Purpose Data; SPIFD2, Structured Process to Identify Fit-For-Purpose Study Design and Data; SPACE, Structured Preapproval and Postapproval Comparative Study Design Framework to Generate Valid and Transparent Real-World Evidence.