

# Real World Evidence

## Table of Contents

1

Applying Advanced Data Privacy Methods to Real World Data (RWD)

2

Estimands for RWD/RWE

3

Missing Data Imputation in RWD Exploration of Multiple Techniques on Open-Source Data

4

Quality and Reusability of Real World Data

5

Real World Evidence (RWE) in Japan

6

Real World Data Lineage Implementation Guide

7

Using OMOP and Other Real World Data Standards to Support Regulatory Submissions

# Applying Advanced Data Privacy Methods to Real World Data (RWD)

## Description

This project aims to develop a comprehensive, cross-disciplinary resource to support the application of advanced privacy-preserving techniques to real-world data (RWD). It addresses the growing complexity of integrating diverse data sources and the increasing use of machine learning (ML), artificial intelligence (AI) and large language models (LLMs) within healthcare.

RWD presents significant privacy risks due to its sensitivity, high dimensionality, and potential for re-identification. While methods such as differential privacy, federated learning and synthetic data generation exist, there is limited consolidated guidance on their practical implementation, particularly in AI-driven contexts.

The project will provide actionable guidance on applying these techniques in compliant, scalable and scientifically robust ways. It will leverage open-source datasets (e.g. MIMIC-IV, UK Biobank, ADNI) for validation and demonstration. The resulting resource will support researchers, developers and regulators in advancing privacy-preserving data integration and responsible AI adoption.

Q4 2027

Proposed  
End Date

White paper/  
guideline

Deliverable

The team has submitted three abstracts for the APAC Connect - all accepted - and two abstracts for the US Connect 2027.

Key Achievements  
This Quarter:

Target to have the draft paper stable

Deliverables &  
Targets Planned for  
the Next Quarter:

Abhishek Mishra &  
Elena Valkanova

Lead



Project Status: Green  
Accepting New Members

- The four subteams are working on the methods and models, and they are posting the scripts in GitHub for the project. There is a delay, and there was low attendance at the regular project calls during the summer.

Project Status

# Estimands for RWD/RWE

## Description

This project focuses on establishing best practices for defining estimands in real-world data (RWD) and real-world evidence (RWE), rather than their implementation using data standards. The team brings together expertise from across PHUSE initiatives, including the Implementation of Estimands (ICH E9 (R1)) project, the Estimands Implementation Working Group (EIWG) HTA and the RWE sub-team, and contributors from the RWE Working Group, representing regulators, statisticians, epidemiologists and health economists.

Currently, there is limited guidance on constructing estimands for RWD/RWE and how these differ from or align with those used in randomised clinical trials (RCTs). Existing regulatory guidance, such as the FDA's February 2023 publication on externally controlled trials, provides an initial foundation but lacks detailed practical direction.

This project aims to address these gaps by developing clear, practical guidance and examples, supporting consistent and appropriate use of estimands across RWD/RWE studies and benefiting stakeholders across the pharmaceutical and biotech industries.

**Q2 2027**

**Proposed  
End Date**

**White paper &  
webinar series**

**Deliverable**

**Ksenia Titorenko,  
Matt Baldwin &  
Paramita Chakraborty**

**Leads**

**White paper draft final**

**Key Achievements  
This Quarter:**

**White paper internal  
review and  
incorporation of  
comments  
Last webinar to be  
planned**

**Deliverables &  
Targets Planned for  
the Next Quarter:**



**Project Status: Green**

- **White paper ready for internal review**

**Project Status**

# Missing Data Imputation in RWD Exploration of Multiple Techniques on Open-Source Data

## Description

This project aims to evaluate and compare multiple missing data imputation methods for real-world data (RWD) using a single open-source simulated dataset (Synthea). The goal is to produce a white paper assessing the strengths, limitations and efficiency of different techniques across varying missing data scenarios.

Handling missing data is critical to ensuring RWD is fit to generate reliable real-world evidence (RWE) and meet regulatory expectations, as outlined in ICH E9 and FDA guidance. However, practical guidance on selecting appropriate imputation strategies remains limited.

By systematically comparing models, this project will provide clear, evidence-based recommendations on which approaches are best suited to different types of missingness. It will also identify key challenges and limitations in current methods. The use of open-source tools ensures accessibility and cost-effectiveness, supporting broader industry adoption and improved data quality for regulatory submissions.

**Q1 2027**

**Proposed  
End Date**

**Draft paper in progress;  
there was a delay  
during the summer.  
The team has  
submitted two  
abstracts for the APAC  
Connect 2027 - Tuhin  
will present a poster.**

**Key Achievements  
This Quarter:**

**Completion of literature  
review and adding new  
sections for the  
methods**

**Deliverables &  
Targets Planned for  
the Next Quarter:**

**White paper**

**Deliverable**

**Aren Ghulijanyan,  
Likhita Kolli,  
Marieta Voskanyan &  
Tuhin James Paul**

**Leads**



**Project Status: Amber  
Accepting New Members**

- Please provide an update on your status (e.g. the status of your deliverable - in review, draft in progress - preparing for an upcoming webinar or blog post, or any other relevant activity).

**Project Status**

# Quality and Reusability of Real World Data

## Description

This project focuses on evaluating real-world data (RWD) sources for regulatory use and identifying the documentation needed to support FDA pre-alignment discussions. It will assess available data sources, summarise their advantages and limitations, and outline the requirements for demonstrating data quality, traceability, appropriateness and regulatory readiness.

RWD has long supported research planning, site and patient selection, safety follow-up and market access, but its use in regulatory submissions requires a more rigorous and structured approach. Sponsors must be able to justify their choice of data source, confirm its representativeness and reliability, and provide robust evidence to support study assumptions.

The project will develop practical guidance on selecting suitable data sources, evaluating quality, and preparing key supporting documents for regulatory engagement. Where possible, it will seek alignment with FDA expectations. The outputs will help organisations strengthen study design early, reduce the risk of rejection, and improve readiness for RWD-based regulatory submissions.

**Project  
closed**

**Proposed  
End Date**

- Guideline review blog posts
- Data vendor feasibility blog posts
- White paper

**Deliverable**

**Published final white  
paper**

**Key Achievements  
This Quarter:**

**N/A**

**Deliverables &  
Targets Planned for  
the Next Quarter:**

**Ashwin Rai &  
Berber Snoeijer**

**Leads**



**Project Status: Green**

**Project Status**

# Real World Evidence (RWE) in Japan

## Description

This project focuses on sharing real-world evidence (RWE) use cases relevant to the Japan regulatory and industry context, including applications in accelerating drug development, post-marketing surveillance (PMS) and preventative healthcare strategies.

There is currently a significant gap in awareness, understanding and practical experience with RWE among pharmaceutical professionals in Japan. This initiative aims to address that gap by creating opportunities to share practical, region-specific examples aligned with regulatory requirements and expectations in Japan. By highlighting relevant use cases, the project will support stakeholders in better understanding how RWE can be effectively applied across the product lifecycle. It will also promote knowledge exchange and encourage more consistent and informed use of RWE within Japan. Ultimately, this work will help build capability, improve decision-making and strengthen the integration of RWE into regulatory and industry practices.

Q4 2026

Proposed  
End Date

- Sessions at the Japan SDE
- Deliver a presentation at the Japan Webinar in Oct 2026

Deliverable

Draft detailed plan for the PHUSE SDE in Japan  
Prepare for the Japan Webinar in Oct 2026

Key Achievements  
This Quarter:

Presentation at the PHUSE SDE in Japan  
Japan Webinar in Oct 2026

Deliverables &  
Targets Planned for  
the Next Quarter:

William Kuan,  
Yasunari Sadatsuki  
& Yosuke Nishida

Leads



Project Status: Green

- The project is progressing as planned

Project Status

# Real World Data Lineage Implementation Guide

## Description

This project aims to develop practical guidance for documenting end-to-end data lineage for real-world data (RWD) regulatory submissions. While FDA and EMA guidance emphasise the importance of tracing data from source to submission, there is limited direction on how to document lineage throughout the full data transformation process. Existing initiatives, such as the CDISC RWD Lineage project, focus primarily on capturing source information within Define-XML but do not fully address transformations such as data selection, cleaning, standardisation, attrition, imputation, and AI-driven processing.

The project will create guidance that helps sponsors consistently document, validate and demonstrate RWD lineage from data acquisition through to SDTM and submission. Areas of focus include defining the information required to tell the complete lineage story, ensuring traceability and immutability, addressing AI-related processes, supporting backward data tracking, and aligning with CDISC RWD lineage standards. The guidance will promote more structured, transparent and regulator-friendly RWD submissions.

Q4 2027

Proposed  
End Date

White paper and/or  
Guideline

Deliverable

Project approved

Key Achievements  
This Quarter:

Kick-off meeting and  
first outline of  
deliverable

Deliverables &  
Targets Planned for  
the Next Quarter:

Berber Snoeijer &  
Ingeborg Holt

Leads



Project Status: Green  
Accepting New Members

- Call for volunteers

Project Status

# Using OMOP and Other Real World Data Standards to Support Regulatory Submissions

## Description

This project aims to raise awareness of data standards for real-world data (RWD) to support regulatory use, with a focus on commonly used data sources such as electronic health records, claims data, patient-reported outcomes, and other observational datasets. It will evaluate common data models (CDMs), including OMOP and the OHDSI ecosystem, alongside alternatives such as PCORnet, Sentinel and CDISC standards.

Differences in regulatory guidance and preferred standards across global agencies can lead to inefficiencies, as data may need to be transformed multiple times to meet varying requirements. OMOP, endorsed by the EMA and supported by a mature community, provides a strong reference point for comparison.

The project will deliver a white paper assessing the strengths, limitations and maturity of key data standards, their suitability for different data types and use cases, and the supporting communities and tools. It aims to promote greater alignment, improve data quality and interoperability, and initiate industry dialogue on the use of these standards for regulatory submissions.

Q2 2027

Proposed  
End Date

None - delayed

Final white paper for  
external review by end  
of quarter

White paper

Deliverable

Key Achievements  
This Quarter:

Deliverables &  
Targets Planned for  
the Next Quarter:

Alexa Parliyan,  
Berber Snoeijer &  
Sanket Kalyankar

Leads



Project Status: Amber

- Initial white paper ready for internal review

Project Status