

ISCTM | 22nd Annual Scientific meeting

Session 1: Methodological Opportunities and Challenges with Long-term Post-marketing Data

Regulatory Perspective

Wei Hua, MD, PhD, MS, MHS

Director, Division of Epidemiology I

Office of Pharmacovigilance and Epidemiology, Office of Surveillance and Epidemiology

Center for Drug Evaluation and Research, U.S. Food and Drug Administration

February 18, 2026

Disclaimer

- This speech reflects the views of the author and should not be construed to represent FDA's views or policies.

Examples of Source(s) of Safety Concern

- Biological plausibility, theoretical concern
- Animal data
- Clinical trials
- Pharmacovigilance, spontaneous reports, e.g., FDA Adverse Event Reporting System (FAERS)
- Medical literature
- Similar drugs on the market
- Others

Types of Epidemiologic Studies

- Experimental studies (interventional studies), e.g., clinical trials
- Observational studies (non-interventional)
 - Study design
 - Cohort design, case-control design, self-controlled designs
 - Data sources
 - Electronic healthcare data, e.g., electronic health records (EHRs), administrative claims data
 - Lab, radiology, biobank, vital statistics (e.g., birth registries, death registries)
 - Primary data collection, e.g., drug registries, disease registries, surveys and interviews
 - Others, e.g., social media

General Approaches for Postmarket Safety Assessment

- Routine pharmacovigilance*
 - Spontaneous reports, case reports or case series from medical literature, etc.
- Non-interventional (observational) studies
 - Registry-based studies
 - Prospective cohort studies with primary data collection
 - Could be supplemented by EHR or other data sources
 - Healthcare database studies
 - Electronic healthcare data, such as EHR and claims data
 - Could be supplemented by linkage to other data sources and primary data collection

Non-interventional Studies – Key Considerations

- Predefined study question
- Required level of evidence based on regulatory objective of study
- Fit-for-use data
 - Data relevance and reliability
- Appropriate study design and methods
 - Study period, population, inclusion/exclusion criteria, exposure/comparator, outcome, covariates, follow-up, analysis, bias and confounding (impact on study results and how to mitigate)

Study Question

Consider: what is the fundamental concept to measure?

- Does data allow for direct measure of the concept?
 - If not, can the concept be measured indirectly (e.g., surrogate endpoint)?
 - Does the indirect measure of concept change the study question?
-

Regulatory Objective of Study

Determines the required level of certainty

When high level of certainty is typically required:

- To evaluate effectiveness for a new indication of an approved drug.
- To evaluate a comparative safety claim.
- To quantify a known risk with substantial existing information.

When less stringent requirement may be allowed:

- To improve current knowledge on safety when little is known.
-

Using Real-World Data To Generate Evidence – General Principles

- In general, data should be sufficiently accurate and complete to address the regulatory question
- Data alone are not evidence. Relevant and reliable data must be used in conjunction with appropriate study design and analysis
- The study question of interest should be established first, and then the data source and study design most appropriate for addressing the question should be determined
- The study should not be designed to fit a specific data source, because the limitations of a specific data source may restrict the options for study design and limit the inferences that can be drawn
- All essential elements of study design, analysis, conduct, and reporting should be predefined
- For each study element, the protocol and final study report should describe how the element was ascertained from the selected real-world data source, including applicable validation studies

Importance of Understanding Mechanism of Action in Causal Inference Assessment

- Mechanism determines the biological time scale of causation, which drives nearly every major study design decision. Understanding the mechanism helps ensure that study design reflects underlying biology rather than statistical association alone.

For example:

- For outcomes with progressive biological change, consider cumulative exposure/dose in exposure definition; “current use yes/no” only may be inadequate
- Consider induction time, carryover effect, and latency time in follow-up specification (e.g., time 0, lagging strategy, length of follow-up)
- Consider factors that drive confounding in the context of exposure-outcome relationship (e.g., treatment indication, disease severity, comorbidities)

Examples of Current Guidelines That Provide General Principles for Data Source Selection, Study Design, and Reporting

- FDA Guidance on Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making (July 2024)
<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory>
- International Council for Harmonisation (ICH) Harmonized Guideline (M14): General Principles on Planning, Designing, Analyzing, and Reporting of Non-Interventional Studies (Sep 2025) https://database.ich.org/sites/default/files/ICH_M14_Step4_Final_Guideline_2025_0905.pdf
- HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE-ISPOR Task Force (June 2022) <https://pubmed.ncbi.nlm.nih.gov/36215113/>

ISPE: The International Society for Pharmacoepidemiology

ISPOR: The Professional Society for Health Economics and Outcomes Research

FDA Guidance on Real-World Data (July 2024)

Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products Guidance for Industry

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

July 2024
Real-World Data/Real-World Evidence (RWD/RWE)

TABLE OF CONTENTS

I.	INTRODUCTION AND SCOPE.....	1
II.	BACKGROUND	3
III.	GENERAL CONSIDERATIONS	3
IV.	DATA SOURCES	4
A.	Relevance of the Data Source.....	5
B.	Data Capture: General Discussion	6
1.	Enrollment and Comprehensive Capture of Care.....	6
2.	Data Linkage and Synthesis.....	7
3.	Distributed Data Networks	8
4.	Computable Phenotypes.....	10
5.	Unstructured Data	10
C.	Missing Data: General Considerations	11
D.	Validation: General Considerations.....	12
1.	Conceptual and Operational Definitions of Study Variables	12
2.	Selection of Study Variables for Validation.....	12
3.	Validation Approaches.....	13
V.	STUDY DESIGN ELEMENTS.....	14
A.	Definition of Time Periods	15
B.	Selection of Study Population	15
C.	Exposure Ascertainment and Validation.....	16
1.	Definition of Exposure	16
2.	Ascertainment of Exposure: Data Source.....	17
3.	Ascertainment of Exposure: Duration	18
4.	Ascertainment of Exposure: Dose.....	18
5.	Validation of Exposure	19
6.	Dosing in Specific Populations.....	20
7.	Other Considerations.....	20
D.	Outcome Ascertainment and Validation.....	20
1.	Definition of Outcomes of Interest.....	21
2.	Ascertainment of Outcomes	21
3.	Validation of Outcomes.....	22
4.	Mortality as an Outcome.....	26
E.	Covariate Ascertainment and Validation	26
1.	Confounders.....	26
2.	Effect Modifiers.....	27
3.	Validation of Confounders and Effect Modifiers	27
VI.	DATA QUALITY DURING DATA ACCRUAL, CURATION, AND TRANSFORMATION INTO THE FINAL STUDY-SPECIFIC DATASET	28
A.	Characterizing Data.....	30
B.	Documentation of the QA/QC Plan.....	33
C.	Documentation of Data Management Process.....	33
	REFERENCES.....	34



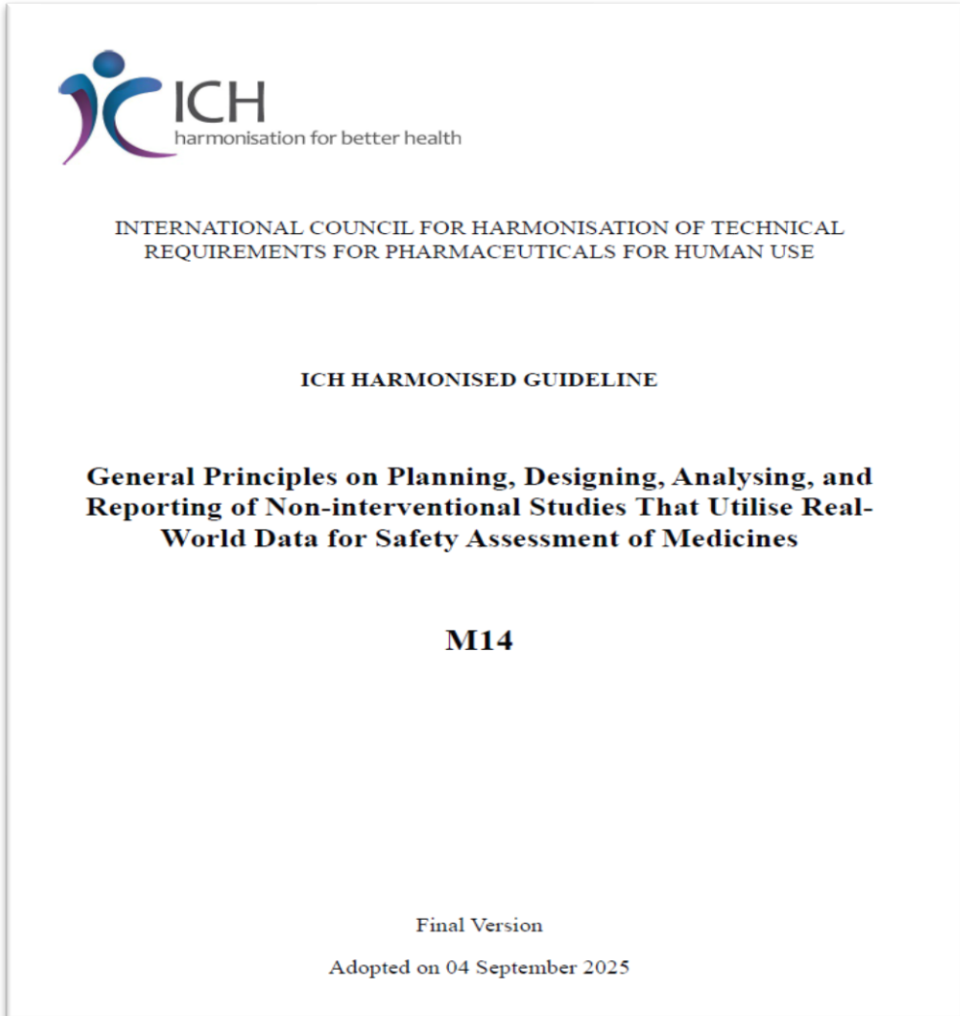
FDA RWD Guidance (July 2024): Key Concepts to Highlight

- Data relevance and reliability
- Study design elements: Definition, ascertainment, and validation
 - Distinguish conceptual and operational definitions

Conceptual	Current medical and scientific thinking regarding the variable of interest; Intended to minimize the effect of variability in practice by different physicians and over time
Operational	Developed based on the conceptual definition to extract the most complete and accurate data from the data source, such as code-based data, lab/imaging, physician's notes, patient survey

- Consider the complexity and impact of misclassification
 - How to determine the extent of validation
- Data quality

ICH M14: General Principles on Planning, Designing, Analyzing, and Reporting of Non-Interventional Studies (Sep 2025)



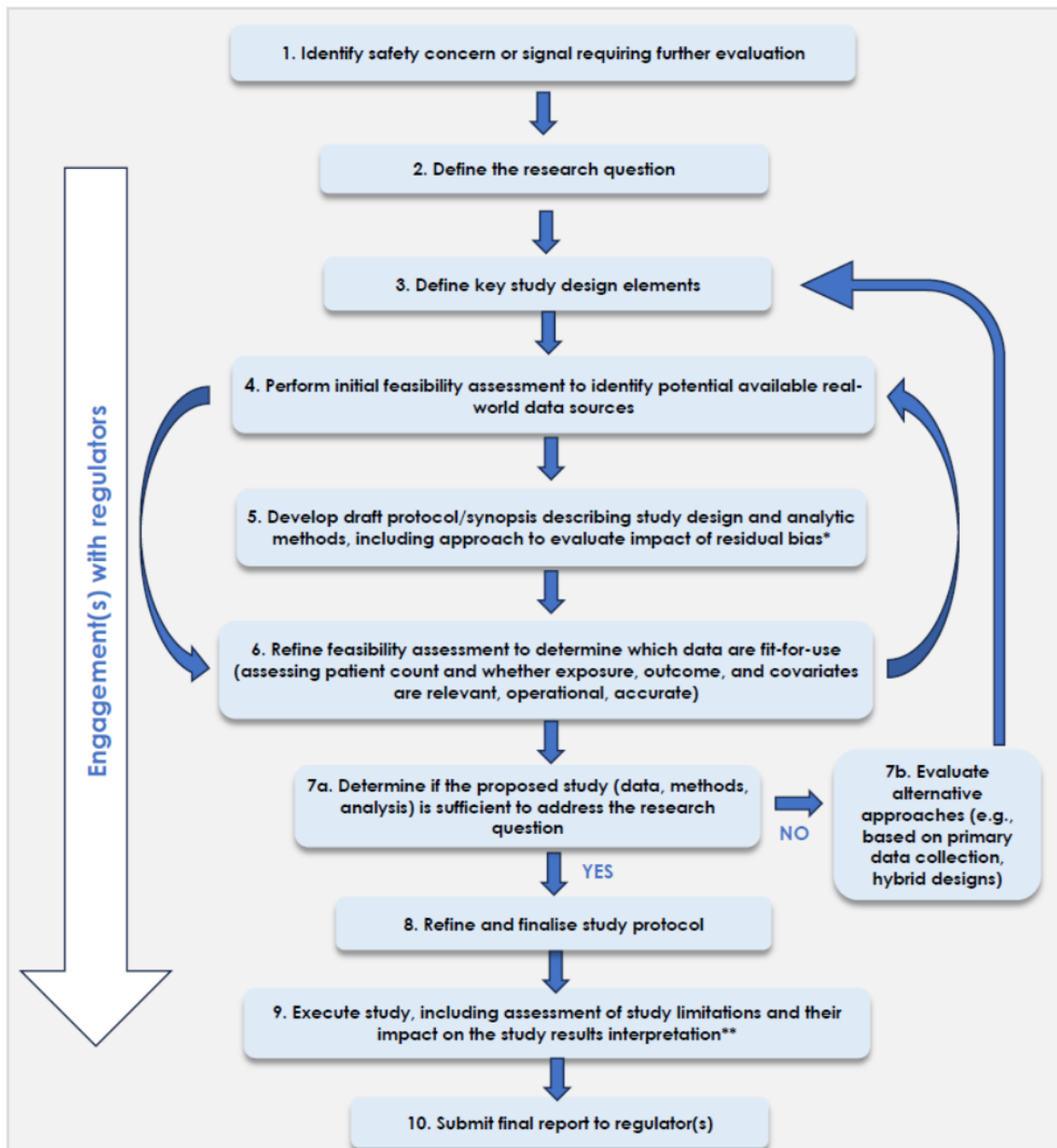
ICH M14 Outlines High-level Best Practice for Non-Interventional Pharmacoepidemiology Studies



Table of Contents

1	INTRODUCTION	1
1.1	Objectives	1
1.2	Background	1
1.3	Scope.....	2
1.4	Studies Conducted for Purposes other than the Safety Assessment of Medicines.....	2
2	GENERAL PRINCIPLES	3
3	CONCEPTUAL FRAMEWORK FOR GENERATING ADEQUATE EVIDENCE USING REAL-WORLD DATA	3
4	INITIAL DESIGN AND FEASIBILITY	6
4.1	Research Question	6
4.2	Feasibility Assessment(s).....	6
5	PROTOCOL DEVELOPMENT	9
5.1	Study Design	9
5.2	Target and Study Population	10
5.3	Data Sources	10
	5.3.1 Characteristics of Data Source Types	11
	5.3.2 Research Networks	14
	5.3.3 Missing Data	16
	5.3.4 Appropriateness of Data Sources in Addressing Safety Questions of Interest	16
5.4	Exposures, Outcomes, Covariates.....	17
	5.4.1 Exposure.....	18
	5.4.2 Outcome	20
	5.4.3 Covariates	21
5.5	Bias and Confounding.....	22
	5.5.1 Selection Bias.....	23
	5.5.2 Information Bias	23
	5.5.3 Time-Related Bias	23
	5.5.4 Confounding.....	24
5.6	Validation of Key Variables	24
6	DATA MANAGEMENT AND CURATION	25
6.1	Data Management Plan	25
6.2	Quality Assurance and Quality Control	25
	6.2.1 Data Quality Management.....	25
	6.2.2 Data Holders.....	26
	6.2.3 Researcher(s)	26
7	ANALYSIS	26
7.1	Statistical Analysis	27
	7.1.1 Analytical approach	27
	7.1.2 Missing Data	28
	7.1.3 Sensitivity Analyses	28
8	REPORTING AND SUBMISSION	28
8.1	Reporting of Adverse Events, Adverse Drug Reactions, and Product Quality Complaints...	28
8.2	Formatting and Content of Study Documents for Submission to Regulators	29
9	DISSEMINATION AND COMMUNICATION OF STUDY MATERIALS AND FINDINGS	29
10	STUDY DOCUMENTATION AND RECORD RETENTION	29
11	CONSIDERATIONS IN SPECIAL POPULATIONS	30
12	GLOSSARY	31
13	LIST OF ABBREVIATIONS	36
14	REFERENCES	38

ICH M14: Conceptual Framework for generating adequate evidence using real-world data



➤ Define research question

➤ Integrated, iterative assessment of three components

- Data fitness-for-purpose
- Appropriateness of study design and methods
- Robust qualitative/quantitative assessment of study limitations and their impact on study validity and interpretation of findings

➤ All three components considered jointly can enable a decision on whether the proposed study, if executed according to the protocol, can generate adequate evidence to address research question

ICH M14: Objectives and Expected Benefits

- Provide recommendations and high-level best practices for the conduct of non-interventional studies, to streamline the development and regulatory assessment of study protocols and reports

Therefore,

- Improve efficiency and transparency in study development, reporting, submission, and review, and resultant regulatory actions
 - Improve the ability of the study protocol and report to be accepted across regulatory authorities
- ❖ The principles presented in this guideline may apply to both safety and effectiveness studies when real-world data are included.

ISPE/ISPOR Joint Task Force: HARPER

HARPER: A standard protocol template with embedded instructions which harmonized across existing guidance and templates

Received: 7 February 2022 | Revised: 17 June 2022 | Accepted: 28 June 2022

DOI: 10.1002/pds.5507

ORIGINAL ARTICLE

WILEY

HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE/ISPOR task force


Shirley V. Wang¹  | Anton Pottegård²  | William Crown³ | Peter Arlett⁴ | Darren M. Ashcroft⁵ | Eric I. Benchimol^{6,7,8}  | Marc L. Berger⁹ | Gracy Crane¹⁰ | Wim Goettsch^{11,12} | Wei Hua¹³ | Shaum Kabadi¹⁴ | David M. Kern¹⁵ | Xavier Kurz⁴  | Sinead Langan¹⁶ | Takahiro Nonaka¹⁷ | Lucinda Orsini¹⁸ | Susana Perez-Gutthann¹⁹  | Simone Pinheiro¹³ | Nicole Pratt²⁰  | Sebastian Schneeweiss¹  | Massoud Toussi²¹ | Rebecca J. Williams²²

TABLE 3 HARPER table of contents

- 1. Title Page
- 2. Abstract
- 3. Amendments and updates
- 4. Timeline
 - Table 1 Milestones and Timeline
- 5. Rationale and background
- 6. Research question and objectives
 - Table 2 Primary and secondary research questions and objective
- 7. Research methods
 - 7.1. Study design
 - 7.2. Study design diagram
 - 7.3. Setting
 - 7.3.1 Context and rationale for definition of time 0 (and other primary time anchors) for entry to the study population
 - Table 3 Operational Definition of Time 0 (index date) and other primary time anchors
 - 7.3.2 Context and rationale for study inclusion criteria:
 - Table 4. Operational Definitions of Inclusion Criteria
 - 7.3.3 Context and rationale for study exclusion criteria
 - Table 5. Operational Definitions of Exclusion Criteria
 - 7.4. Variables
 - 7.4.1 Context and rationale for exposure(s) of interest
 - Table 6. Operational Definitions of Exposure
 - 7.4.2 Context and rationale for outcome(s) of interest
 - Table 7. Operational Definitions of Outcome
 - 7.4.3 Context and rationale for follow up
 - Table 8. Operational Definitions of Follow Up
 - 7.4.4 Context and rationale for covariates (confounding variables and effect modifiers, e.g. risk factors, comorbidities, comedICATIONS)
 - Table 9. Operational Definitions of Covariates

- 7.5. Data analysis
 - 7.5.1 Context and rationale for analysis plan
 - Table 10. Primary, secondary, and subgroup analysis specification
 - Table 11. Sensitivity analyses – rationale, strengths and limitations
- 7.6. Data sources
 - 7.6.1 Context and rationale for data sources
 - Table 12. Metadata about data sources and software
- 7.7. Data management
- 7.8. Quality control
- 7.9. Study size and feasibility
 - Table 13. Power and sample size
- 8. Limitation of the methods
- 9. Protection of human subjects
- 10. Reporting of adverse events
- 11. References
- 12. Appendices

HARPER: Aiming for Transparency and Efficiency

- Help researchers thoroughly consider, then document their decisions and rationale regarding key study parameters that define causal question
- Facilitate decision-making by enabling reviewers to readily assess potential for biases related to the researcher's choices and rationale
- Facilitate reproducibility of results

Takeaway Messages

- Methodological challenges vary by study question of interest
- FDA does not endorse any type of data source or study methodology
- Researchers should identify the data-design combination best suited to address the study question
- To facilitate FDA review, the protocol should provide a detailed description and justification of study design and data source(s)
- Discussion with the relevant FDA review division early in the application process is strongly encouraged (case-by-case)

