23 May 2024
EMA/CHMP/ICH/155061/2024
Committee for Human Medicinal Products

ICH M14 Guideline on general principles on plan, design and analysis of pharmacoepidemiological studies that utilize real-world data for safety assessment of medicines

Step 2b

Transmission to CHMP	11 April 2024
Adoption by CHMP	25 April 2024
Release for public consultation	30 May 2024
Deadline for comments	30 August 2024

Comments should be provided using this template. The completed comments form should be sent to ich@ema.europa.eu

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

INTERNATIONAL COUNCIL FOR HARMONISATION OF TECHNICAL REQUIREMENTS FOR PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED GUIDELINE

General Principles on Plan, Design and Analysis of Pharmacoepidemiological Studies That Utilize Real-World Data for Safety Assessment of Medicines

M14

Draft version
Endorsed on 21 May 2024
Currently under public consultation

At Step 2 of the ICH Process, a consensus draft text or guideline, agreed by the appropriate ICH Expert Working Group, is transmitted by the ICH Assembly to the regulatory authorities of the ICH regions for internal and external consultation, according to national or regional procedures.

M14
Document History

Code	History	Date
M14	Endorsement by the Members of the ICH Assembly under *Step 2* and release for public consultation.	21 May 2024

General Principles on Plan, Design and Analysis of Pharmacoepidemiological Studies That Utilize Real-World Data for Safety Assessment of Medicines

M14

ICH Consensus Guideline

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 3
2 GENERAL PRINCIPLES 3
3 FRAMEWORK FOR GENERATING ADEQUATE EVIDENCE USING REAL-WORLD DATA 4
4 INITIAL DESIGN AND FEASIBILITY 5
4.1 Research Question 5
4.2 Feasibility Assessment(s) 6
5 PROTOCOL DEVELOPMENT 8
5.1 Study Design 9
5.2 Data Sources 10
5.2.1 Appropriateness of Data Sources in Addressing Safety Questions of Interest 11
5.2.2 Characteristics of Major Data Sources 12
5.2.3 Data Standardization 16
5.2.4 Missing Data 17
5.2.5 Data Quality 17
5.2.6 Data Collection and Data Source Sections in the Study Protocol 17
5.3 Target/Study Population 18
5.4 Exposures, Outcomes, Covariates 18
5.4.1 Exposure 20
5.4.2 Outcome 22
5.4.3 Covariates 23
5.5 Bias and Confounding 25
5.5.1 Selection Bias 25
5.5.2 Information Bias 26
5.5.3 Immortal Time Bias 26
5.5.4 Confounding 26
5.6 Validation 27
6 DATA MANAGEMENT 28
6.1 Data Holder 28

6.2 Researchers 29

7 ANALYSIS **29**

7.1 Statistical Analysis 30

7.1.1 Primary Analyses *30*

7.1.2 Missing Data *31*

7.1.3 Sensitivity Analyses *31*

8 REPORTING AND SUBMISSION **32**

8.1 Reporting of Adverse Events, Adverse Drug Reactions, and Product Quality Complaints 32

8.2 Formatting and Content of Study Documents for Submission to Regulatory Authorities 32

9 DISSEMINATION AND COMMUNICATION OF STUDY MATERIALS AND FINDINGS **32**

10 STUDY DOCUMENTATION AND RECORD RETENTION **33**

11 CONSIDERATIONS IN SPECIFIC POPULATIONS **34**

11.1 Pregnancy Studies 34

12 GLOSSARY **35**

13 REGULATORY GUIDELINES REFERENCED **42**

14 NON-REGULATORY GUIDELINES REFERENCED **43**

15 REFERENCES **45**

1 Introduction

1.1 Objectives

The purpose of this document is to recommend international standards for, and promote harmonization of, the general principles on planning, designing, and analyzing observational (non-interventional) pharmacoepidemiological studies that utilize fit-for-purpose data for safety assessment of medicines (drugs, vaccines, and other biological products).

This document outlines recommendations and high-level best practices for the conduct of these studies, to streamline the development and regulatory assessment of study protocols and reports. These recommendations and practices also seek to improve the ability of the study protocol and/or results to be accepted across health authorities and support decision-making in response to study results. The Glossary defines several terms for the purpose of this guideline. Terms that appear in ***bold italic*** type upon first use are defined in the Glossary.

1.2 Background

Pharmacoepidemiological studies have long been a source of data and evidence to support the evaluation of the post-marketing safety of approved ***medicines***.

Signals can arise from a wide variety of data sources. This potentially includes all clinical and scientific information concerning the use of medicines and the outcome of this use, such as product quality, non-clinical and clinical data (including pharmacovigilance and pharmacoepidemiological data). Epidemiological studies are a key component in the detection, characterization and evaluation of safety concerns or signals and may be descriptive or inferential.

Generation of robust evidence to be used for regulatory purposes relies on the reliability and relevance of the data and the application of sound pharmacoepidemiological methods to analyze such data. The use of pharmacoepidemiological studies for regulatory decision-making has increased globally, and multiple guidelines and best practice documents have been developed by health authorities and professional societies. Many countries and regions have published guidelines related to general principles of planning and designing such studies mainly for the purpose of safety assessment of a medicine. In addition, frameworks for study design and conduct are being developed by non-governmental groups, such as The Sentinel Innovation Center with the PRINCIPLED framework and ISPE/ISPOR's HARmonized

Protocol Template to Enhance Reproducibility (HARPER) Initiative, which provide additional detail that is beyond the scope of this guideline [1, 5].

1.3 Scope

While recognizing that there may be slight differences between regions with regard to what constitutes ***real-world data*** (RWD), this guideline includes recommendations for studies utilizing RWD conducted for the purposes of evaluating post-marketing safety of medicinal products. At times, RWD sources alone may be insufficient to answer the research question of interest and researchers will gather additional data for the purposes of the study. For the purpose of this guidance, we refer to data collected for the specific study as primary data collection. Because primary data collection may be relevant to observational studies using RWD, when relevant, this guideline also includes considerations for primary data collection.

It is beyond the scope of this document to provide guidance on whether a clinical trial or a pharmacoepidemiological study is the most appropriate approach, nor is it intended as a comprehensive source of knowledge for pharmacoepidemiological methods. Rather, the intent is to harmonize regulatory guidance documents for the design, planning and execution of pharmacoepidemiological studies, and to facilitate regulatory review. Parties can also consider, as relevant, best practice guidances from other sources to the extent not covered in regulatory guidance (see Non-regulatory Guidelines Referenced).

The following study types are out of scope:

- Pharmacovigilance studies using spontaneous reports obtained from national or global databases (e.g., pharmacovigilance systems at national level);
- Studies involving treatment assignment, including randomized controlled trials, pragmatic trials, single arm clinical trials with treatment assignment defined per protocol, and trials using external comparators; and
- Studies collecting and analyzing ***patient experience data***.

Collecting patient experience data may be a valuable component for post-marketing safety studies to inform on aspects such as notable events, perspectives, needs, and priorities. While a detailed guidance on this is beyond the scope of this guideline, several regulatory guidances have been developed (see Section 13, Regulatory Guidelines Referenced). When studies include patient experience data, the researcher may consult relevant published recommendation for additional information.

Considering the evolving nature of pharmacogenomics, artificial intelligence (AI), and other emerging technologies relevant to the use of RWD, this guideline does not address those topics.

1.4 Studies Conducted for Purposes other than the Safety Assessment of Medicines

The principles presented in this document provide recommendations that may be applicable to post-market pharmacoepidemiological studies conducted for purposes other than evaluation of the safety of medicines, such as utilization and effectiveness studies. The basic principles presented in this guideline may be relevant to these studies when real-world data elements are included.

2 General Principles

The safety profile of a medicine reflects an evolving body of knowledge extending from preclinical investigations through the post-approval lifecycle. Post-approval pharmacoepidemiological safety studies complement other sources of information to provide a better picture of the benefit-risk profile of a medicine as used in clinical practice.

The guideline describes a stepwise process, although the various steps of study design and data source selection are iterative. The process starts with articulating the research question; conducting a systematic process to identify the study population, ***exposure***, outcome, and covariates; identifying minimal data requirements to guide feasibility assessment; assessing the representativeness of the data source to the target population; and considering sources of potential ***bias*** and ***confounding***. After an appropriate data source and/or data collection approach has been identified, the process involves further refining the design, which includes approaches to address study validity. The fit-for-purpose evaluation section of the guideline describes the integration of these activities. Throughout the process, the underlying rationale and justification for exposure, outcome, and confounder definitions, analysis, data management, study implementation, reporting, submission, dissemination of results, and other key decisions should be documented. All operational aspects should be clear and transparent.

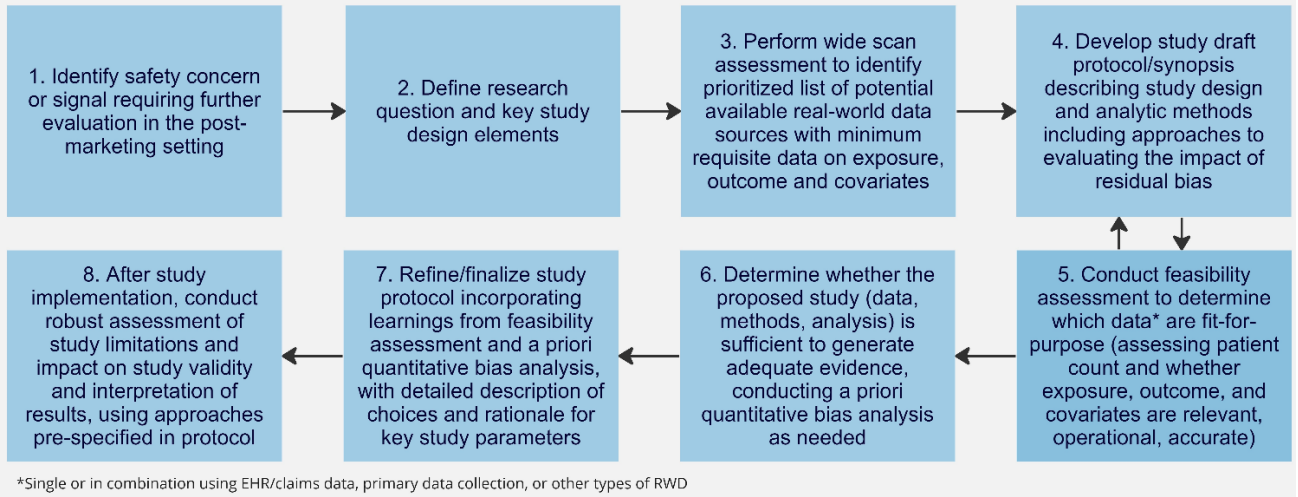
In this guideline we refer to “researcher” as those responsible for designing and executing the study; this may be a regulatory agency, sponsor, contract research organization, academic group, or others. Sponsors of marketing applications and marketing authorization holders are ultimately responsible for all aspects of post-marketing safety studies submitted to regulators.

3 Framework for Generating Adequate Evidence using Real-World Data

The strength of the study generated evidence submitted in support of a regulatory decision depends on the research design and methodology in addition to the relevance and reliability of the underlying data. Within the framework for generating adequate evidence (Figure 1), the research question should be established first, then the study design and data source(s) most appropriate for addressing those questions are determined (2). Researchers should avoid designing a study that conforms to a specific data source, because a specific data source may restrict the options for study design and limit the inferences that can be drawn. In general, to determine if the evidence that will be generated is adequate to answer the research question, the framework should include an integrated assessment of (1) ***data relevance*** and ***data reliability***, (2) appropriateness of the study design and analytic methods, and (3) a qualitative/quantitative robust assessment of study limitations and their impact on the ultimate validity and reliability of the resulting evidence and the interpretation of findings. Integrated assessment of whether the evidence generated through the study is adequate should be considered both during protocol development with a feasibility assessment (e.g., discussion of the impact of theoretical concerns, consideration of possible sources of bias and their potential impact on study validity) and after study implementation with sensitivity analyses pre-specified in the protocol. ***Quantitative bias assessments (analyses)*** may be employed either *a priori* for feasibility assessment, or to facilitate interpretation of study results, or for both purposes. All three components considered simultaneously can enable a decision on whether the study, if executed according to the protocol, can generate adequate evidence to address the specific regulatory question. Studies involving user-generated health data extracted from other sources (e.g., websites, blogs, social media, chat rooms) may not be adequate, but they may provide supportive data to generate hypotheses and contextualize the study results.

Although **Figure 1** depicts a linear process, consideration and evaluation of evidence that is adequate should be iterative [2]. Researchers are encouraged to discuss the attributes of a particular study with the regulatory agency early in the planning process. The ensuing sections of this guideline outline the necessary elements of a study protocol that will allow for a validity assessment.

Figure 1: A framework for generating adequate evidence using fit-for-purpose real-world data to address regulatory questions on the safety of medicines.



4 Initial Design and Feasibility

4.1 Research Question

The research question is a concise statement of the study purpose and the prespecified hypotheses to be tested; the purpose of the study may also be to generate hypotheses for future research. The research question may be formulated by use of the population, intervention (exposure in the case of non-interventional studies), comparator, outcome, and timing (PICOT) template. In the case of non-interventional studies, “intervention” can be considered the same as an exposure. The specific question should be formulated after a review of the literature to identify and understand any knowledge gaps, strengths and weaknesses of prior studies, the expected magnitude of effect, and important confounding factors. When defining the research question, researchers should provide a clear rationale on how it will be addressed by the study objectives. In the protocol, researchers should document and support decisions about the study design and the types of data required/available. Careful formulation of the research question will highlight unknowns that will need to be addressed through information derived from the feasibility assessment and this information may further refine the question and drive protocol development. Researchers may also consider a principled framework for study design and estimation of the risks of a medicine, such as the ***target trial*** approach or the ***estimands*** framework as they initiate work on the research question and conduct initial design and feasibility analyses [3, 5].

4.2 Feasibility Assessment(s)

A feasibility assessment is a systematic process to identify fit-for-purpose data to address a specific research question and to obtain information on the statistical precision of a potential study without evaluating outcomes for treatment arm. When conducting a feasibility assessment, a key goal is to describe and compare the reliability and relevance of the data sources assessed for the research question without evaluating outcomes associated with the medicine under evaluation. Additional detail on potential strengths and limitations of data sources is provided in Section 5, Protocol Development.

Feasibility assessments should be structured in at least two phases:

- An initial scan to determine whether data are available, likely sufficient, and to narrow down data source options; and
- A subsequent, more comprehensive feasibility assessment of the candidate data sources.

After the research question and design elements are established, researchers should specify the minimum criteria required to address the key design elements specific to the research question. This task will require an understanding of RWD source characteristics and the clinical context. Design elements to consider include:

- Data needed to understand and define the study population, exposure, comparison groups, outcomes and covariates;
- Minimum length of follow-up to observe outcome(s);
- The targeted sample size/event rate and expected study precision;
- Geographic region(s) of interest; and
- When feasible, information about the health care system including method of diagnosis, preferred medicines, formulary coverage and prescribing practices.

In the early stages of designing a non-interventional study, expectations regarding access to patient level or analytic data sets should be clarified. Sponsors should obtain any required third-party agreements to access relevant patient-level or analytic data that will be required by the regulatory authority for submission.

Other important elements related to the feasibility assessments can include:

- Whether appropriate codes for a diagnosis are available, especially for rare diseases;
- Whether laboratory confirmation of a diagnosis and/or access to medical records are necessary to validate outcomes or exposures; and

- Whether evidence for the validity of coding algorithms exists.

Depending on the research question, it may be appropriate to specify other important criteria, such as the ability to collect additional information to complement records in the data sources, or link data sources to other types of data (e.g., vital records, cancer registries, vaccine registries). At this point in the initial scan step, it will be possible to identify data sources that are most likely to satisfy the criteria the researcher has specified as important to answer the research question. In some cases, it is possible for the researcher to complete this initial scan step relying on published information, data source descriptions, catalogues of metadata, and occasionally, simple descriptive counts available from the data source.

Once a manageable number of available data sources have been identified as potential candidates for utilization in the study, an in-depth feasibility assessment should be conducted. In some instances, fit-for-purpose data will be identified during the initial feasibility scan, in which case the detailed step will apply to the databases under consideration. In the detailed feasibility step, the researcher can verify that the specific data needed for the key design criteria are available and that there is sufficient evidence of validity and completeness of the minimal design elements in the specific data source.

When selecting a data source, data recency, frequency of data refresh, completeness of follow-up from exposure to outcome should be considered. In addition, the possibility to submit data files generated during conduct of the study to relevant regulatory agencies may need to be determined. Other factors in data source selection may be prior experience with the data, as there may be a trade-off between the time needed to address these factors versus the urgency of obtaining study results.

Analyses that evaluate the potential impact of missingness of data may be conducted to further evaluate the feasibility of conducting a study in a given data source. For example, in a study evaluating the association between hormonal contraceptives and thromboembolism, the impact of missing smoking status information may be evaluated by a review of the literature to determine the association of smoking with exposure and outcome, and then using quantitative bias approaches to evaluate the impact on the study validity for a range of desired effect estimates. A variety of information sources are used to complete this step evaluating the narrowed down list of data sources, and it may often be valuable to request specific information from the ***data holder*** (e.g., number of patients exposed, incidence rate of outcome to conduct sample size calculations, availability of covariates, and other queries of the data to verify the

data source is fit-for-purpose).

After the detailed evaluation is complete, the data sources are compared, and a data source(s) can be selected for the study. Occasionally, at any of the steps, it will be apparent that a specific data source is not suitable to address the research question. In these circumstances, the researcher may conduct a feasibility assessment for primary data collection. This assessment typically includes physician and site queries, including information about the patient population, to determine if a sufficient number of participants can be enrolled and followed for the appropriate timeframe to yield meaningful answers to the research question. Whenever primary data collection studies are proposed, the researcher should consider the time to set up the study which includes time to select sites, undergo ethical approval, enroll, and follow participants, and produce results, and whether this timing is acceptable.

In addition, the specification of an appropriate comparator group (or time period) is a critical part of the study design and an important consideration in the feasibility assessments. The impact that policies for medical or medication coverage could have on the observed level of disease severity of the exposed group and the comparator group must be considered, as should the availability of concurrent comparator data. However, in some situations (such as rare disease population studies) a historical or former ***standard of care*** comparator may be considered. Regulatory guidances provide additional information on the characteristics of an appropriate comparator.

Feasibility assessments are used as context for design decisions in the protocol. In discussion with, and where required by a regulatory authority, submission of the feasibility assessment report can either be a standalone document, or an annex to the protocol. This report should describe the data sources evaluated when designing the study, including results from feasibility evaluations or exploratory analyses of those data sources. Sponsors should provide a justification for selecting or excluding relevant data sources from the study.

The final approach should comply with applicable regulatory requirements. Detailed frameworks, templates, and checklists for conducting feasibility assessments are available in scientific publications.

5 Protocol Development

The design and conduct of pharmacoepidemiological safety studies typically require the participation of subject matter experts. An experienced, multidisciplinary study team with the

appropriate expertise is crucial to the successful execution of a safety study and the protocol should include a description of the expertise and credentials of the study team. These personnel provide essential input in a number of areas, including:

- Development of exposure, outcome and covariate definitions, with appropriate medical expertise to understand disease manifestation, causal pathways, and current clinical practices;
- The unique features of existing electronic healthcare data sources based on their intended purpose and methods for collecting data;
- Disease area billing and coding practices;
- Specific characteristics around primary data collection; and
- Addressing potential data privacy and security concerns raised when accessing health care data.

Completeness of data capture, bias in the assessment of exposure, outcome and covariates, variability among data sources, the impact of changes over time in the data, governance and conditions of access to data, and the healthcare system of the country or region covered by the database are important elements that can affect the choice of the data source(s) for the study and need to be addressed in the study protocol.

5.1 Study Design

Pharmacoepidemiological safety studies usually aim to estimate the incidence of an adverse outcome in a population of interest and to evaluate its association with exposure to a medicine.

Several study designs are commonly used in observational pharmacoepidemiological safety studies, including cohort, case-control, and self-controlled studies. The selection of the most appropriate study design depends on multiple factors including the research question of interest and what is known about the postulated relationship between exposure to a medicine and the specific safety outcomes of interest (e.g., acute vs. latent outcome, biologic ***plausibility***).

Identifying the appropriate comparator population (designed to represent the counterfactual experience) is a critical element of study design. Examples of comparators may include users of other medicines, non-users, historical controls, or the patient themselves in self-controlled designs. Considerations for comparator selection may include the specific indication within a disease, contraindications, disease severity or comorbidity, and the treatment sequence. It is important to maximize and evaluate the comparability of the exposed and comparator

populations to reduce issues related to confounding by indication.

Researchers should discuss their rationale for selecting a particular study design in the study protocol and final report. Researchers should also consider developing graphical representations (such as a study design diagram) to clarify the analysis plan and time components such as inclusion period, lookback period, follow-up period, overall study period. Visualization of design details helps to clarify and communicate the study design to a broad audience of decision makers [3]. The proposed study design should be discussed with health authorities early in the process to ensure that the proposed study may provide adequate evidence for regulatory decision-making.

After initial feasibility analyses, all essential elements of study design, analysis, conduct, and reporting should be prespecified. For each study element, the protocol and final study report should describe how that element was ascertained from the selected data source in studies utilizing secondary data, including applicable validation studies.

5.2 Data Sources

Before using any data source in support of regulatory decision-making, sponsors should consider whether the data are fit-for-purpose by assessing the data's relevance and reliability. For the purposes of this guidance, the term relevance includes the availability of key data elements (patient characteristics, exposures, outcomes) and a sufficient number of representative patients for the study (target population), and the term reliability includes ***data accuracy***, ***completeness***, ***provenance***, and ***traceability***. The protocol should provide discussion and documentation of these key data characteristics.

Several data source characteristics need to be considered in pharmacoepidemiological studies, as they may affect the study design and the interpretation of the results. These include differences in coding systems used across databases, standardization of data elements, and settings of care captured (e.g., primary, hospital, specialty, rehabilitation). Patients, providers, or healthcare systems may have different motivations (monetary, social, or otherwise) for data collection or participation, and billing practices for reimbursement, which may impact the characteristics of the underlying data and further inform study design and interpretation.

In recent years, federated networks of RWD sources have been developed in various regions. When utilizing multiple data sources, either as a network or through data linkage, researchers should consider the steps taken to harmonize data across institutions or data sources (see ***Federated Data Networks***). Some of these networks have been specifically designed to support

scientific evaluations and regulatory decision-making, allowing a growing number of studies to include data from these federated networks, often from different countries. It is essential to understand the strengths and limitations of the chosen data source(s).

5.2.1 Appropriateness of Data Sources in Addressing Safety Questions of Interest

Researchers should demonstrate an understanding of the data source(s) and its appropriateness to address specific research questions. This understanding of the chosen data source(s) including the relevance and reliability of the data to address the specific research question, in conjunction with an appropriate study design and analysis, is key to providing accurate evidence. During development of the protocol, as informed by the feasibility assessment(s), researchers should describe the following key aspects of the proposed data source(s) to support the demonstration of their relevance, the selection rationale and how they might affect the generalizability of the study results to the targeted patient population:

- How well the selected data source captures study elements (e.g., whether a variable is captured, and if so, the degree of completeness);
- The capability to validate the outcome and other key study elements (e.g., exposures, key covariates, inclusion/exclusion criteria);
- The historical experience with use of the selected data source for research purposes, including references for publications citing relevant previous use for pharmacoepidemiology studies which may demonstrate fit-for-use characteristics or other elements to support use of the data source for the proposed study;
- Time to data availability, frequency of data refresh;
- The relevant healthcare system factors, such as medication tiering (e.g., first-line, second-line), formulary decisions, and patient coverage, can influence the degree to which patients on a given therapy in one health care system might differ in disease severity, or other characteristics, from patients on the same therapy in another healthcare system;
- The key patient characteristics which might act as potential confounders, including age, socio-economic status, health conditions, risk factors for the outcome, health system (e.g., private or public/governmental healthcare); and
- Potential limitations of the data source for addressing the research question.

5.2.2 Characteristics of Major Data Sources

Regardless of the data source(s) used, information on the context of the evidence generation should be obtained (e.g., geographic location, setting in which the data were generated, period during which the data were collected, and demographic information such as the age and sex distribution of populations included in the data source). Examples include data derived from EHRs, ***administrative healthcare claims data (claims data)***, data from patient registries, patient-generated data, and data gathered from other sources that can inform on health status, such as interviews, mail surveys, computerized or mobile-application questionnaires, measurements through digital health technologies (see Digital Health Technologies). Although there are regional differences, such as medical practice, below are general considerations for common data types.

Electronic Health Record (EHR) Data

Electronic Health Record (EHR) data are captured by healthcare institutions, and these data reflect episodic care as captured within that specific institution and may not reflect the patient's complete medical history, because they may miss data from other settings of care. Given that components and formats of data may differ among medical institutions, standardization of data formats is often a major issue in a study when integrating data from multiple institutions.

Key clinical information are often unstructured data within EHRs, either as free text fields (such as healthcare practitioner notes) or as other non-standardized information in computer documents (such as PDF-based radiology reports). Free text may be used to further characterize exposure and outcome (e.g., review of patient profiles) in EHR-based data sources. To enhance the efficiency of data abstraction, a range of approaches, including both existing and emerging technologies (e.g., natural language processing, computer vision for images or laboratory results evaluation) are increasingly being used to convert unstructured data into a computable, structured data format.

When making secondary use of EHR data from multiple medical institutions, any differences in components and format of these data, including codes used (such as disease names, drug names, and laboratory test items) should be harmonised and the approach documented in the protocol. EHR data typically capture the medical encounter with the health care provider but may not reflect the actual delivery of healthcare (e.g., medicines that are ordered but not dispensed or administered) and may require additional linkage (e.g., to pharmacy records). In addition, obtaining comprehensive history of medicine use, or medical care data on patients

with certain types of privacy concerns (e.g., sexually transmitted infection, substance use disorders, mental health conditions) can be challenging. Nevertheless, failure to capture these data can result in inaccurate or incomplete data.

Claims Data

Healthcare claims databases are often large and capture healthcare services for all individuals covered by a health insurance program(s). Typically, once claims for all healthcare provided to individuals within a health insurance program are fully adjudicated (i.e., final payment decisions made by insurance companies or claims processors), they are aggregated into a database that reflects a more complete view of services. Some databases will contain a mix of open (in-process) and closed (paid/denied) claims and the researcher should understand the dynamic nature of the data in these cases. Without linkage to other data sources, it is often not possible to obtain information about healthcare visits, results from laboratory testing, outcomes of offspring in pregnancy studies, many vaccinations, injuries from accidents, and other care not covered by health insurance. These issues may be due to numerous factors, including health insurance coverage policies and seeking medical care outside of the insurance system (e.g., self-paid/self-care treatments, procedures insured by worker's accident insurance, and motor vehicle liability insurance).

Registries

A ***registry*** is an organized system that collects prespecified uniform data from a population defined by a specific disease, condition, or exposure. Registries may be further described as "Patient Registries" or "Product Registries" to indicate defining characteristics for registry entry. The former highlights a focus on collecting data from patients with a certain disease, specific populations, such as pregnant or lactating people, or individuals with a specific condition, such as a birth defect or a molecular or genomic feature. The latter is a system by which sponsors collect data on patients exposed to a specific health care product or class of products.

An already established registry may be used to collect data for reasons other than originally intended. If a study makes secondary use of registry data, the same considerations and fit-for-purpose assessment relevant to secondary sources such as EHR and claims data should be applied to evaluate the suitability of the registry to answer the research question, e.g., taking into account the registry population, data elements collected, including longitudinally, frequency of data assessments, and calendar time, level of data quality, and governance

(including aspects on data sharing and data access). Additional considerations may include the type of registry and the impact of methods involved in patient selection on the representativeness of the population relative to the target population (such as geographic factors, total number of patients in the registry, number eligible, number of new patients entering the registry per year and number lost per year with reasons for exit). If data necessary to answer the research question(s) are not routinely collected within established registries, linkage to external data sources or supplemental data collection through other means should be explored. In some cases, de novo registries or other study designs may be required (e.g., need for an adequate comparator population in a single-arm product registry, key measures of exposure or covariates such as duration, dose and route of therapy administration, or intractable channeling bias requiring randomization).

Data Collected by Digital Health Technologies (DHT)

Digital health technologies (DHTs) are systems that use computing platforms, connectivity, software, and/or sensors for health care and related uses. These technologies span a wide range of uses, from applications in general wellness to applications as a medical device. They include technologies intended for use as a medical product, in a medical product, or as an adjunct to other medical products (devices, drugs, and biologics). They may also be used to develop or study medical products. Technological advances have increased the range of data sources that can be used to complement traditional ones and may provide insights into or relevant to safety (and effectiveness) of health interventions. These technologies should be subject to the same fit-for-purpose assessments as other data sources. There may be a need to specify DHTs (e.g., version, software, hardware, manufacturer), or to harmonize data across different types of devices. Depending on the data source maturity, greater validation work may be needed.

Federated Data Networks

Federated Data Networks (FDNs) enable distributed analyses combining data or results across multiple databases. When choosing to use a FDN for a study, there are specific issues unique to these systems that should be considered, such as the FDN's transformation of data into common data models (CDMs), and the differences between systems from which the data arise. Governance of federated networks (centralized or decentralized) also needs to be taken into account, as it has an impact on the operational aspects of a study.

Under the FDN framework, different approaches can be applied for combining data or results from multiple databases. A common characteristic of all approaches is the fact that data partners

maintain physical and operational control over electronic data in their existing environment and therefore the data extraction is always done locally. Differences, however, exist in the following areas: use of a common protocol; use of a CDM; and where and how the data analysis is conducted.

The choice of data captured in a CDM is optimized for the types of data measures and detail needed for the intended use. Therefore, data in CDM-driven networks rarely contain all of the source information present at the individual databases, and the data elements chosen for a given CDM network may not be sufficient for all research purposes or questions.

FDNs can provide unique advantages that can assist with addressing drug safety questions, such as:

- Decreasing the time to conduct a study, either through pre-developed analyses, or by increasing the size of study populations as this shortens the time needed to obtain the desired sample size. Large sample sizes may facilitate research on rare events, rare diseases, and less common drug exposures;
- Multi-database studies may provide additional knowledge on whether a drug safety issue exists in different populations or across countries and thereby may reveal causes of differential drug effects, inform the generalizability of results, the consistency of information and the impact of biases on estimates;
- Heterogeneity of treatment options and utilization patterns across institutions, communities or countries may allow for a more complete understanding of the effect of individual medicines; and
- Involvement of experts from various countries addressing terminologies, coding in databases and research practices provides opportunities to increase consistency of results of pharmacoepidemiological studies.

Data Linkage

Data linkage can be used to increase the breadth and depth of data on individual patients over time and may be utilized to allow access to other data sources to support validation efforts. Linkage of data sources such as cancer or mortality registries linked to claims or EHR may result in a higher quality study by including data not in the original data source. It is important to have a comprehensive understanding of the data and to assess the accuracy and completeness of the linkage and the resulting linked data. In some circumstances, chart review or text entries in electronic format linked to coded entries can be useful for exposure, outcome, and covariate

identification.

Conceptually, a data linkage may be undertaken within a database (e.g., mother–infant linkage) or across databases (e.g., vital records, biobank). If the study involves a data linkage, the protocol should describe each data source, the information that will be obtained, linkage methods, and the accuracy and completeness of data linkages over time. If the study involves generating additional data (e.g., interviews, surveys, computerized or mobile-application questionnaires, measurements through digital health technologies), the protocol should describe the methods of data collection and linkage, explanations of the data elements used for linkage, and what will be done if imperfect linkage exists, or contradictory data are found across linked data sources.

5.2.3 Data Standardization

Data standardization is relevant to multi-database studies, including federated data networks. There are several challenges to consider in standardizing study data derived from RWD sources. These challenges to standardization include but are not limited to:

- The type of information the sources contain (e.g., diagnoses, procedures, medications);
- The variety of RWD sources and the level of consistency in formats and coding languages, including differences in source data captured regionally and globally using different standards and terminologies;
- Differences in healthcare systems, such as business processes and local healthcare practice patterns, database structure, vocabularies, coding systems, and de-identification methodologies used to protect patient data when shared.

Coding systems for diagnoses, medicines, and laboratory data, among others, are updated regularly. Therefore, plans for mapping coding systems as they evolve/change should be addressed at the protocol stage. Moreover, care should be given when re-using a code list from another study, as code lists reflect the individual study objectives, methods, and the time in which they were created.

A free-text/unstructured component exists in some databases, and can be used to define inclusion criteria, exclusion criteria, exposures, outcomes, follow-up, and covariates. Each free-text component may be transferred into a structured table which prompts users to specify what is measured, the timing of measurement, the care setting, type of codes that are used to define the measure as well as the sources for any algorithms used to derive study measures,

e.g., defining exposures, outcomes, or covariates. The process for creating a structured variable from unstructured data should be provided in the study documentation.

5.2.4 Missing Data

Missing data are data value(s) that are not captured in the data source of interest. There are two scenarios where data can be missing. The first scenario is the data are intended to be collected but were not collected. The second scenario is the data are not intended to be collected in the data source and therefore not available. A record in EHR systems or administrative claims databases is generated only if there is an interaction of the patient with the health care system. Lack of information such as a laboratory result or prescription, could be caused by different circumstances, such as (1) it might not have been ordered by the health care provider; (2) it may have been ordered but not conducted; (3) or it may have been conducted, but the result (test, dispensing) is not recorded; or (4) there is evidence of the healthcare interaction and the result was stored in the data source, but data were not in an accessible format, or lost in the transformation and curation process when the final study-specific dataset was generated. Approaches to handle missing data are described in further detail in Section 7, Analysis.

5.2.5 Data Quality

Fundamental determinants of data quality at each step in the evidence generation process, such as governance and documentation need to be addressed before finalizing the protocol. Depending on the data source, pharmacoepidemiologic data may lack strict ***quality control (QC)*** over the process of recording, collection, and storage. This can lead to incomplete data, missing key variables, or inaccurate records. The presence of such quality defects will affect subsequent ***data curation***, applicability, and traceability of data.

Compared to Good Clinical Practice (GCP), procedures for pharmacoepidemiologic data quality control and ***quality assurance (QA)*** follow guidances specific to pharmacoepidemiologic data, and detailed quality standards to be fulfilled should be in accordance with local or regional regulatory requirements (see Section 6, Data Management for more details on QA and QC).

5.2.6 Data Collection and Data Source Sections in the Study Protocol

The protocol should describe the data source(s) used and how it/they are fit-for-use to address the research question of interest. In addition, the protocol should state any coding systems used for classification of the exposure and outcomes (e.g., anatomical therapeutic chemical (ATC),

International Classification of Disease (ICD), and any methods used for data linkage). Data collection methods and procedures should be described.

For studies that use data from multiple data sources or study sites (e.g., federated data, meta-analysis, or data pooling), researchers should describe the rationale and procedures for how data from different sources can be obtained and integrated with acceptable quality, given the potential for heterogeneity in population characteristics, clinical practices, and coding across data sources.

For studies with primary data collection, the identification, processing and reporting of adverse events occurring in the course of treatments should be described in the protocol, in accordance with relevant jurisdictional laws and regulations (see Section 8, Reporting and Submission).

5.3 Target/Study Population

The target population is the population about which one wants to make an inference (e.g., children aged 12-16 with attention deficit hyperactivity disorder). The study population is intended to be representative of the target population from which data will be evaluated to answer the research question. The study population is defined via inclusion and exclusion criteria (e.g., demographic factors, medical conditions, disease status, severity, biomarkers) and identified based on the following elements, among others:

- Time points, such as index dates for inclusion in the study, defined lookback period (e.g., to identify new users);
- Key variables (see Feasibility Assessment(s)]) used to select the study population and how they should be validated (see Bias and Confounding); and
- The completeness and accuracy of the data elements to fulfil the inclusion and exclusion criteria (see **Error! Reference source not found.**).

5.4 Exposures, Outcomes, Covariates

If the initial feasibility assessment has indicated that the exposures, outcomes, and covariates of interest are likely to be adequately captured in the selected data sources, then defining and operationalizing these elements should proceed. This process generally starts with the creation of a ***conceptual definition*** which is initiated at the time of initial database selection. The conceptual definition should reflect current medical and scientific thinking regarding the variable of interest, such as: (1) clinical criteria to define a condition for population selection or as an outcome of interest or a covariate; or (2) measurement of drug intake to define an

exposure of interest. The conceptual definition should include a detailed description of the data elements that would characterize the exposure, outcome, or covariates.

Utilizing the key data elements identified during the feasibility phase, this conceptual definition is then developed into the ***operational definition***. An operational definition should be developed based on the conceptual definition to extract the most complete and accurate data from the data source. In many studies using EHRs or claims data, the operational definition will be a code-based electronic algorithm using structured data elements. In other studies, the operational definition may be derived from extracting relevant information from unstructured data or constructing an algorithm that combines structured and unstructured data elements. Operational definitions can also specify additional data collection, such as a patient survey, when appropriate. Researchers should consider the following areas when developing exposure, outcome, and covariate definition(s):

- Whether it is possible to translate a conceptual definition of the exposures, outcomes, and covariates into one that can be operationalized in selected data source(s);
- Whether the operational definition adequately captures all elements of the conceptual definition; and
- Whether the operational definitions and the performance characteristics are adequate in the chosen data source(s) based on the research question (see Validation).

The conceptual definition is often referred to as the ***phenotype.*** The protocol should include a detailed description of the operational definition, sometimes referred to as the computable phenotype (including the coding system and rationale) and the associated limitations (e.g., measurement bias, proxies), the potential impact of misclassification, and how these limitations can be mitigated through the study design and analysis. For unstructured data, a detailed description, rationale for use, search criteria to identify outcomes/exposures/covariates, and the list of codes or concepts should be provided. The operational definitions should be documented in the protocol and/or the statistical analysis plan.

Operational definitions developed for one data source or study population might perform differently in other sources or populations in terms of sensitivity and specificity due to database-specific characteristics as well as variations in the disease epidemiology across populations and databases. If utilizing or adapting a definition used or validated in other studies or databases, applicability must be justified.

When identifying exposures and outcomes in a database for a specific study, data related to

these types of information are usually coded. When selecting a data source, appropriateness of the coding system for defining the exposures and outcomes should be confirmed.

The following elements require consideration during protocol development and are described in more detail below:

- Data source/type and structure;
- Development of exposure, outcome and covariate definitions and the method used to identify them;
- Development and performance of the operational definitions, including time points, data types (structured, unstructured), variable types (categorical, continuous), transformation of variable types, code types, mapping of dictionary codes (e.g., ICD-10 to MedDRA) when applicable, and the mechanism of evaluation (selection of gold standard) and performance measures;
- Mapping of the available data elements against those required for the research question;
- Documentation of variable validity and appropriateness of applying previously used algorithms to the database/population of interest; and
- Potential impact of misclassification on study validity and interpretation.

5.4.1 Exposure

Conceptual Definition

An exposure is the medicine of interest (and dosing or regimen) being evaluated in the proposed study. The product of interest is referred to as the treatment and may be compared to no treatment, standard of care, another treatment, or a combination of the above.

Exposure definitions can have differing levels of granularity, such as ever exposed vs. never exposed, duration of exposure, user type (e.g., incident vs. prevalent), exposure windows (e.g., current vs. past exposure), also referred to as risk periods or risk windows, multiple exposure (e.g., use of more than one medicine or concomitant vaccinations), treatment switching, sequencing (e.g., first line or second line) or dosage (e.g., current dosage, cumulative dosage over time). Consideration should be given to both the requirements of the study design and the availability of data. The exposure definition should include information about the medicine dose, brand, formulation, strength, route, timing, frequency, and duration (as applicable). It may also be necessary to describe the manufacturer as part of the product identification (e.g., for a medicine with the same active substance name made by different manufacturers). This

may require an understanding of the pharmacological or biological properties of the medicine, or members of the product class.

Operationalizing Exposure

By Medicine Type, Route, and Setting

When translating the conceptual definition to the operational definition, there are uncertainties that should be considered, and justified in a discussion of strengths and limitations in the protocol. For example:

- Medicines that are prescribed are not necessarily dispensed;
- Medicines that are dispensed are not necessarily used or administered;
- Patient compliance and ability to provide an accurate account of intake;
- Exposures that are not captured in the data source such as samples, low-cost medicines, non-prescription medicines, and immunizations offered in the workplace; and
- Coding systems used to identify exposures (e.g., NDC, RxNorm, HCPCS).

The setting of administration should be considered carefully. Infusions may be administered in private clinics or on an outpatient basis (e.g., home care) in addition to in-hospital, so setting and treatment patterns should be considered in terms of potential requirements for data linkages.

For vaccines, it is essential to have granular information on brand, dose schedule, coadministration with other vaccines, and sometimes batch number or administration route and site. These data may require linkage to vaccination registries.

By Medicine Dose, Timing, and Duration of Exposure

The exposure (i.e., dose, dosage regimen) to the medicine of interest should be well-defined and measured. Consideration should be given to the timing of exposure for medicines (e.g., the relevant exposure window, relative to the onset of the outcome), and this may be especially difficult when “as needed” or non-prescription medicines are an exposure of interest. When defining the exposure period, it is necessary to decide whether the start date of exposure (the index date) is the date of prescription, the date of dispensing, or the date of administration. Because patients may not refill their prescriptions exactly on time or, alternatively, may refill their prescriptions early, gaps or stockpiling in therapy may exist and may be reflected in the data. Allowable gaps between dispensing to construct exposure episodes and the gaps between exposure episodes should be considered when determining whether an exposure period is

continuous. Conditions for the completion of an exposure period should also be considered and explicitly defined (e.g., no record of a new prescription in preceding six months), noting limitations such as the potential of a drug being prescribed to a patient in another setting that may not be captured in the dataset used for the study.

5.4.2 Outcome

Conceptual Definition

Defining an outcome of interest should be based on the clinical, biological, psychological, and functional concepts of the condition, as appropriate. This conceptual definition should reflect the medical and scientific understanding of the condition. Considerations for how outcomes should be identified will include whether cases can be identified as true incident (vs. prevalent), the latency, and whether the outcome presents with exacerbations or as recurrent episodes. The definition should include a detailed description of the data elements that would confirm the outcome (e.g., signs, symptoms, laboratory and radiology results).

Clinical outcome definitions should contain diagnostic criteria, measuring methods and their quality control (if any), measurement tools (e.g., the use of questionnaire scales), calculation methods, measurement time points, variable types, transformation of variable types (e.g., from quantitative to qualitative variable), and mechanism of endpoint event evaluation (e.g., the operation mechanism of the endpoint event committee).

Operationalizing Outcome

An operational definition is one that can be implemented independently using the data available in the proposed study with acceptable performance to meet the goals of the study. The conceptual definition is operationalized using diagnosis and procedure codes (e.g., ICD-9-CM, ICD-10, Read, MedDRA), laboratory tests (e.g., Logical Observation Identifiers Names and Codes [LOINC]) and values, unstructured data (e.g., physician's encounter notes, radiology, or pathology reports), or measurement tools such as questionnaire scales. Consideration for changes in coding or the underlying EHR systems over time is essential. If unstructured data are used, detailed description and rationale for the methods and tools utilized and validation of those methods should be provided.

Single appearances of a diagnosis code may indicate a rule out diagnosis or lack adequate specificity. Instead, consider whether a valid definition of outcome can be achieved by combining medicines, laboratory data, and medical procedures used for diagnosis or treatment,

rather than operationalizing the outcome only based on the disease or diagnosis (e.g., a thromboembolism diagnosis code plus treatment with anticoagulant, anaphylaxis code plus use of epinephrine). In some studies, in which the outcome is complex to define, information on the specialty of the physician making the diagnosis might help provide additional reassurance regarding the quality of the information used to determine the outcome. Mortality as an outcome may not be included in electronic health care data unless the patient was under medical care at time of death. Linkage to external vital statistics resources may be necessary.

When considering use of previously developed operational definitions, researchers should consider secular trends in disease, diagnosis, or changes in coding practices that may necessitate assessment using more recent data. Published ***case definitions*** of outcomes may be used but are not necessarily compatible with the information available in a given RWD data set. When proposing to use an operational definition that has been assessed in a prior study, ideally select those assessed in the same data source and in a similar study population. In addition, the quality of prior studies to establish sensitivity, specificity, and predictive values should always be evaluated. Applicability of a definition used in a prior study or validated in another data source will depend on an assessment of its external validity with a justification provided in the protocol.

When conducting a study using data from multiple data sources (databases, institutions, sites), define the outcome considering the data differences between sources, such as diagnosis coding, laboratory reference ranges and medication records. A complete understanding of the timing and relationship between these elements is essential. For example, there are situations where the start date of treatment on the claims data and the date of diagnosis on the EHR data may not match for the same patient.

When outcomes to measure patient experience are included (e.g., quality of life, subjective severity of disease), the protocol should specify how the outcome measure is defined, constructed, and validated, and the procedures for data collection. The reason for the data collection and the nature of the healthcare system that generated the data should also be described as they can impact the quality of the available information and the presence of potential biases.

5.4.3 Covariates

Covariates are variables that are neither an exposure nor an outcome of interest, but instead are measured because they either characterize a population or are a potential confounder or effect

modifier to account for in study design or analysis. As with exposure and outcome, the definition moves from a conceptual definition to an operational definition based upon clinical, biological, psychological, and functional concepts, as appropriate. The definition should include a detailed description of the data elements used to construct the covariate.

The potential for confounding and effect measure modification should be considered and planned for during protocol development. For example, the potential for ***effect modification*** by demographic variables (e.g., age, sex, race, ethnicity), other exposures (e.g., biologically active herbals) or pertinent comorbidities should be documented in the study, and relevant effect modifiers should be available in the chosen data source.

- **Confounding:** Confounding occurs when the estimate of measure of association is distorted by the presence of another factor. For a variable to be a confounder, it must be associated with both the exposure and the outcome, without being in the causal pathway.
- **Effect Modification:** Effect modification occurs when the effect of a single exposure on an outcome depends on the values of another variable, i.e., the effect modifier, which does not necessarily need to be involved in the causal pathway.

Conceptual Definition

Definitions of covariates needed in a study should be identified and a determination made on whether it can be directly operationalized in a given data source. When the covariate is not available in the chosen data source, researchers may consider whether proxies for the covariate are appropriate.

Operationalizing Covariates

Moving from a conceptual to operational definition proceeds as with exposure and outcome. Covariates may be used to characterize cohorts, to develop propensity scores, to stratify or match patients, evaluate effect modification and adjust for confounding. Covariates are typically identified and assessed during the period before the start of the exposure of interest (baseline). Assessment of baseline covariates can be performed using different periods of time. The length of this lookback period is selected by considering factors such as changes in coding or medical practice, expected frequency of medical encounters, relevance to the research question, and the impact on study power. Covariates and may also be assessed during the observation period, either as static or time varying variables. Reliable assessment of covariates is therefore essential for the validity of results, including the timing of assessments for each of

the covariates. A given database may or may not be suitable for studying a research question depending on the availability of information on these covariates. Researchers should provide the developed operational definition, including codes and settings of care, for all covariates in the protocol.

5.5 Bias and Confounding

To obtain a valid and precise estimate of the effect of exposure on the outcome of interest, studies must address two sources of error. Unlike random error, systematic error (bias, confounding) cannot be addressed by increasing sample size. Rather, it is typically addressed in the design, conduct, and analysis stages. From the epidemiological standpoint, it is useful to differentiate the concepts of bias (e.g., selection bias, information bias, resulting from design or measurement errors) and confounding because they arise from distinct mechanisms and may be addressed by distinct methods and approaches in study design and analysis. The design and analysis stages should include evaluation of any potential biases such as information bias and selection bias which can be due to the inclusion/exclusion criteria or loss to follow-up, as well as evaluation of any confounding that may arise, especially if some data elements cannot be collected or measured. Therefore, the handling of missing data should also be prespecified in the Data Management section (see Section 6, Data Management) or Analysis section (see Section 7.1, Statistical Analysis) of the protocol.

The proposed data source should be evaluated to determine whether it is adequate to capture information on important factors so that bias and confounding may be adequately controlled. Linkage with other data sources or additional data collection to expand the capture of important variables that are unmeasured or imperfectly measured in the original data source should be considered. Sources of bias and confounding should be considered, and decisions to address should be justified during the design stage with a plan to evaluate the influence of bias and confounding; these should be included in the protocol, analysis plan or final report.

5.5.1 Selection Bias

There are different types of selection bias such as referral bias, self-selection bias, prevalent user bias, and loss to follow-up (time-related bias). Different forms of selection bias may be addressed in either the design (preferred) or analysis stages.

A common type of selection bias is prevalent user bias, which can arise when prevalent users of a medicine are included in a pharmacoepidemiologic study, i.e. patients already taking a

therapy for some time before study follow-up began. Prevalent users are 'survivors' of the early period of pharmacotherapy that is not captured in the study. This can introduce selection bias if the risk varies with time. For example, subjects who initiated a new medicine, experienced a safety event, and then discontinued the medicine may not be included in the study, thereby leading to a potential underestimation of the risk among the treated.

5.5.2 Information Bias

Information bias arises when misclassification of binary or categorical variables or mismeasurement of continuous variables exists. Whereas internal validity should always be optimized, and misclassification of key variables should be minimized to accurately estimate the effect of exposure on the outcome, some degree of misclassification may be acceptable in some studies depending on the study question and regulatory purpose and should be determined on a case-by-case basis. Overall, the extent of variable validation (see Validation) should be determined by the necessary level of certainty and the implication of potential misclassification on study inference. As discussed in Section 3, a plan to use ***quantitative bias analysis*** may be useful when evaluating the direction and magnitude of biases to inform strategies for bias mitigation, and how the study biases may influence the interpretation of the study (see Analysis).

5.5.3 Immortal Time Bias

Immortal time bias refers to a period of cohort follow-up time during which an outcome of interest cannot occur. Selection of an appropriate index date is essential to avoid the risk of immortal time bias and other time-related biases. This risk may be mitigated by design frameworks (see Research Question), as this approach aligns assessment of eligibility and baseline information with start of follow-up [4].

5.5.4 Confounding

Researchers are typically unable to capture all potential confounders that are relevant to a research question, introducing the potential for unmeasured or residual confounding. In pharmacoepidemiology, commonly considered confounding factors include demographics, indication for treatment, disease severity, previous medication use, and comedications, comorbidities, prognostic characteristics, frailty, and others, depending on the study question. A number of approaches are available to address or evaluate unmeasured confounding, including high dimensional propensity scores, negative controls, and linkages to external data

sources such as surveys that include data on confounders unmeasured in the study database. The presence and impact of potential confounding factors should be considered in the study design phase. Directed acyclic graphs can be used to understand the relations between the variables and identify potential confounding and intermediate effects in a longitudinal study, and the impacts of these assessed using quantitative bias analysis, as discussed in the Analysis section (see Analysis).

5.6 Validation

Validity is the extent to which a concept (variable) is accurately measured in a study by the operational definition. Validation of exposure, outcome, and key covariates is important for internal validity of pharmacoepidemiological studies. There are various approaches to validation, which may be data-source specific. These may include complete verification, partial verification, clinical expert review, review of patient claims, or profile history. Validation efforts should be commensurate with the level of evidence required, such as validating the outcome variable for all potential cases or non-cases or verifying the performance of an operational definition to identify cases and non-cases. For databases routinely used in research, documented validation of key variables may have been done previously. Any extrapolation of a previous validation study should however consider the effect of any differences in prevalence and inclusion and exclusion criteria, the distribution and analysis of risk factors as well as subsequent changes to health care, procedures, and coding. Sponsors should have early interactions with regulatory authorities to discuss and agree upon a proposed validation approach, such as partial vs. full, or adoption of definitions validated previously. A justification for validation approach should include the data source, population, time frame, performance, reference standard, and a discussion of the applicability of the proposed operational definitions considering the level of evidence required.

Validation studies should be conducted under a separate protocol. When validating an operational definition, prespecify the metrics to be reported (e.g., sensitivity, specificity, positive predictive value, negative predictive value, kappa statistic), and describe how they will be measured. The trade-off between false-positive and false-negative cases should be considered when selecting an operational definition and identifying the proper validation approach to support internal validity. For instance, when cases are rare, one may need to select a highly sensitive operational definition and then validate all potential cases.

If several operational definitions are under consideration, the performance of each should be

evaluated and quantified using bias analysis in the design stage. This is distinct from common sensitivity analyses conducted in the analysis stage.

6 Data Management

Management of data quality for a pharmacoepidemiological study depends on various factors, including the source of the data and the planned use of the study results. A data management and/or data curation plan should be developed prior to study initiation. The quality assurance and quality control (QA/QC) plan should be developed before an analysis is undertaken and the various factors (e.g., data management by the data holder, quality defects of the data, inadequate data processing and analysis, or inadequate training) influencing quality should be identified and addressed to preserve the integrity of the study. Detailed quality standards to be fulfilled should be in accordance with local or regional regulatory requirements.

To facilitate regulatory review, where submission of datasets is a regulatory requirement, a description of the context, content, structure of files, and steps used to create the files should be included. Datasets must be retained in accordance with relevant regulatory requirements in the region(s) to which they will be submitted. Any migration of data and documents to new media or a new format should be verified to ensure long-term readability and to maintain integrity.

Data Management Plan

Data quality assurance processes, policies, and procedures should account for potential risks to data quality, including errors in interpretation or coding; errors in data entry, transfer, or transformation accuracy; errors of intent; inadequate training; ***data completeness***; and ***data consistency***.

A description of data storage, management, and statistical software should be included in the protocol. All procedures used to obtain, verify, and promote the integrity of the analytic dataset should be recorded in sufficient detail so that they can be replicated. Data security should always be maintained by limiting access to authorized individuals.

Quality Assurance and Quality Control

6.1 Data Holder

QA and QC procedures used by the ***data holders*** may include ensuring reliability of data collection and management; the frequency and type of any data error corrections or changes in

data adjudication policies implemented by the data holders during the relevant period of data collection; peer-reviewed publications examining data quality and/or validity, updates and changes in coding practices (e.g., International Classification of Diseases, ICD codes) across the study period that are relevant to the outcomes of interest; changes in key data elements during the study time frame and their potential effect on the study; the extent of missing data over time (i.e., the percentage of data not available for a particular variable of interest), and procedures (e.g., exclusion, imputation) employed to handle these issues.

6.2 Researchers

While the data holder maintains control of the data and is responsible for the underlying data quality, the researchers are responsible for aligning QC and QA procedures with data holders to ensure transparency, understanding of data strengths/limitations, and meeting the standards of quality criteria required by the regulatory authorities. Further, the researcher is responsible for the management and quality assurance of all data cleaning, processing, and analytic datasets. To balance the need for sufficient quality assurance with reasonable resource expenditure for a particular purpose, a risk-based approach to quality assurance is recommended. Issues that are essential to determining the reliability and relevance of the data should be addressed in the protocol, and include QA/QC procedures for data accrual, curation, and transformation into the final study-specific dataset.

The researcher is responsible for implementing and maintaining QA/QC systems with written procedures. This is to ensure that studies are conducted, and results are generated, documented, and reported in compliance with the protocol, regional laws, ethical considerations, and the applicable regulatory requirement(s). Documentation of these processes may include, but are not limited to electronic documentation (i.e., metadata-driven audit trails, quality control procedures) of data additions, deletions, or alterations from the data source to the final study analytic dataset(s). Researchers should also document changes to data and the potential impacts of these changes for conducting this specific study. Methods for quality assurance and quality control of analytic programming should be described in the protocol, such as the process to inspect code and/or replicate code, or whether an analytical code that was previously QA/QC'ed is being used.

7 Analysis

The analytic strategy includes descriptive and inferential analyses to address the study

objectives, while accounting for potential sources of bias and confounding. In addition, the strategy should also include an empirical evaluation of unmeasured or mismeasured confounding and other sources of bias. The statistical analysis should be prespecified, reflect the information gained from the feasibility assessments(s), and be developed to meet the study objectives. An overview of the ***statistical analysis plan (SAP)*** should be provided in the protocol. The complete SAP should be provided as a standalone document, or as a detailed section of the protocol. It is recommended that the approach chosen should be discussed with health authorities, keeping in mind that the protocol and SAP are highly interdependent. The SAP should provide sufficient detail to allow replication of the study to help ensure confidence in the results.

In some studies, data driven analyses may be performed and it is important to distinguish between those that are pre-specified and those that are *post-hoc*. Pre-specified analyses such as those used for covariate selection should be documented in the protocol and analysis plan and deviations from the plan documented in the final report. Post-hoc analyses are often conducted in response to observations in the data to help in the interpretation of results and should be described and justified in the final report.

Researchers should consider developing a timeline of the analyses that will be performed during the conduct of the study (e.g., accrual, descriptive analyses, inferential analyses, sensitivity analyses, and quantitative bias analysis).

Proactive planning is required when conducting a multi database study or when using an FDN, as the analytic strategy is impacted by the types of databases or the FDN under consideration. Specific issues may need to be considered in the analysis plan, such as the independent related analyses performed in each data source or FDN which may require meta-analytic techniques.

7.1 Statistical Analysis

7.1.1 Primary Analyses

The analysis should be directed towards the unbiased estimation of the epidemiological parameters of interest (e.g., risk or rate differences and risk or rate ratios). The analysis section is where a description and justification for the chosen approaches for the statistical analyses should be described, including the assumptions and conditions.

The following aspects and elements may be considered for inclusion: descriptive analyses, subgroup analyses, methods of estimation and associated assumptions needed for analysis,

estimate of the anticipated study size/power/statistical precision, plans to control for confounding and bias (e.g., misclassification, selection bias, information bias, time-related bias, and impact on validity of results), assessment of population comparability, sensitivity analyses, type I error control (e.g., for sequential analysis), assessment of representativeness and plans for handling missing data.

If the analysis proposes to use machine learning or other derivation methods, specify the assumptions and parameters of the computer algorithms used, the data source from which the information was used to build the algorithm, whether the algorithm was supervised (i.e., using input and review by experts) or unsupervised, and the metrics associated with validation of the methods.

7.1.2 Missing Data

Researchers should develop the protocol and the statistical analysis plan with an understanding of reasons for the presence and absence of information in the underlying data; consider data linkage and or imputation to address missing data, and address the implications of the extent of missing data on study findings (see Bias and Confounding). Descriptive analyses should be included to characterize missing data. Assumptions regarding the missing data (e.g., missing at random, missing not at random) underlying the statistical analysis for study outcomes and important covariates should be supported and the implications of missing data considered. The analysis should address missing data in line with the methods described in the statistical analysis plan. The extent and implications of missing data on study findings should be described.

7.1.3 Sensitivity Analyses

When planning for sensitivity analyses, a rationale for each analysis should be provided with the strengths and limitations of each analysis. Sensitivity analyses should be conducted to examine the effect of varying potentially critical assumptions of the analysis, such as those relating to design, estimands, exposure definition, outcome definition, missing data, and limitations of the data source(s) selected. The analyses can facilitate better interpretation of study results in light of the extent of uncertainty noted.

Quantitative bias analysis evaluates the impact of potential bias on the measure of association. The protocol should pre-specify the indices (e.g., sensitivity, specificity, positive [PPV] and negative [NPV] predictive values) that will be used for quantifying bias and describe how the

selected indices will be measured when validating variables of interest. The precision of the bias-adjusted effect estimates should be quantified using confidence intervals. These analyses may facilitate interpretation of study results.

8 Reporting and Submission

8.1 Reporting of Adverse Events, Adverse Drug Reactions, and Product Quality Complaints

Adverse events, adverse drug reactions, and product quality complaints identified during the conduct of a study may require reporting to the relevant regulatory authority. Reporting requirements may vary by party (e.g., marketing authorization holder (MAH), other sponsor or applicant, investigator, or independent research group) and by region, due to differences in regulatory reporting requirements. The ICH E2D guideline on Post Approval Safety Data Management provides guidance for MAHs on reporting of individual case safety reports for adverse events and adverse drug reactions. For other reporting requirements (and for parties outside the scope of ICH E2D), refer to applicable laws and regulations and, as appropriate, pharmacovigilance guidelines.

8.2 Formatting and Content of Study Documents for Submission to Regulatory Authorities

Sponsors should discuss with regulators the required study documents and timetables for submission. These documents may vary by regulator, can include the feasibility assessments, protocol, analysis plan, and interim and final reports. In the absence of specific formatting and content required by regulators, sponsors may utilize frameworks developed by the scientific community as a guide for document development, such as ISPE/ISPOR's HARmonized Protocol Template to Enhance Reproducibility (HARPER) [1, 5].

9 Dissemination and Communication of Study Materials and Findings

For transparency, to support scientific exchange, and to allow the conduct of reproducible research, even where not mandated by regulatory requirements, it is encouraged that researchers make study materials publicly available. It is encouraged that the protocol and statistical analysis plan be made publicly available in appropriate registers before study initiation, and study reports upon completion. Further vehicles for dissemination and communication of study results may include non-regulatory submission in scientific fora,

scientific publications, and patient or practitioner-focused communications.

Several guidelines exist that provide recommendations for reporting studies in medical literature. These include The Reporting of studies Conducted using Observational Routinely collected health Data (RECORD) Statement, RECORD-PE, and "Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals," established by the International Committee of Medical Journal Editors (ICMJE). In addition, when publishing the contents of the study, the contents of the report should be summarized so that the publication is consistent with the report. To avoid publication bias, it is recommended that the results be published even if negative or inconclusive study results are obtained with respect to study objectives or hypotheses.

Results of the research should be communicated to the study participants (for example, when primary data collection is used), the public, and patients, so that they may be aware of and understand the study results and their implications. Communications should include a factual summary of the overall study results in an objective, balanced and nonpromotional manner, including relevant safety information and any limitations of the study.

10 Study Documentation and Record Retention

Key documents and records related to the planning, conduct and results of a study should be kept in compliance with applicable standards and jurisdictional requirements. Key principles for studies utilizing RWD in post-marketing safety studies are similar to those for GCP (especially for primary data) and Good Pharmacoepidemiological Practice (especially for ***secondary use of data***).

- All study information, documents and records, should be recorded, handled, stored and archived in a way that allows its accurate reporting, interpretation, verification, and that ensures confidentiality and patient privacy in compliance with applicable privacy laws;
- Systems are in place to ensure completeness of the study documentation, to prevent accidental or premature loss, prevent unauthorized access, alteration, destruction, disclosure or dissemination; and ensure that an audit trail is maintained;
- Needed systems are in place with procedures that assure the quality of every aspect of the documentation of study development, conduct, and reporting;
- Study information should be readily available and directly accessible upon request by regulatory authorities (e.g., internal or regulatory inspection ready) with risk-based

quality checks or review processes to ensure that the primary record system is being maintained up-to-date and that all key documents are appropriately filed; and

- All information retained at least for the duration of time required by applicable regulatory requirements.

11 Considerations in Specific Populations

Specific (special) populations are often not enrolled in pre-approval clinical studies and include pregnant and lactating people, infants, children, adolescents/young adults, older adults, immunocompromised patients, and people with disabilities and/or rare disorders. Therefore, post-market pharmacoepidemiological studies may provide valuable information supporting the benefit/risk assessment of medicines in these populations. Studies in these populations may require unique considerations when defining the study population, in addition to other considerations applying to any studies (such as definition of exposure, confounders and outcomes). Examples of challenges include low sample sizes for rare diseases; multiple comorbidities and polypharmacy for older adults; and difficulty in identifying cases or disease characteristics (e.g., duration and severity) in immunocompromised patients.

11.1 Pregnancy Studies

Specific challenges of secondary use of data in pregnancy studies include identification of pregnancies, complexity of outcomes, and need for validated algorithms to identify gestational age or date of conception, and maternal and infant outcomes. These challenges may necessitate linkage within the data source (e.g., mother-child link) or complementary data sources such as birth registries. Pregnancy registries can provide more granular clinical information on timing of exposure, gestational age, outcomes, and covariates; however, there are challenges with such registries, including difficulty with enrolment and retention of participants and selection bias.

The dichotomous approach of ever- vs. never exposed does not reflect exposure patterns in pregnancy and approaches to address varying risks by trimester should be considered. Attention should be given to definition of risk windows, measurement of both conception and pregnancy start dates, and patterns of medicine use (e.g., start and end dates, dose, frequency, duration). A valid estimate of gestational age, from which a conception date may be estimated, is critical for determining the timing of exposure and may require availability of linked data such as ultrasound or laboratory data. Accurate information about gestational timing of exposure(s) can help identify critical exposure periods and inform biological plausibility of specific effects.

Exposure information in the time period just before pregnancy is often also important, especially for products with a long half-life.

Outcomes include outcomes during pregnancy that affect maternal health (such as preeclampsia or gestational diabetes), spontaneous abortions, birth/neonatal outcomes, and child developmental outcomes which may extend for several years after birth. The protocol should state *a priori* criteria for defining the outcomes of interest, including their severity (e.g., *major* birth defect), and take into account that many adverse pregnancy outcomes have substantial variation over the course of pregnancy. There are unique challenges in outcome identification for pregnancy studies and use of standard classification systems should be considered. Preterm birth and "small for gestational age" are reliably available in registries, but in administrative data may be identified through diagnostic codes or calculated using gestational age and birth weight data. Depending on the data collection methods, birth defect surveillance registries are useful as they have already been adjudicated for live births, stillbirths/fetal deaths, and elective terminations. Major congenital malformations may be recorded in the mother's record, the infant's record, or both.

Bias and confounding in pregnancy studies include, but are not limited to, family history and confounding by indication. The analysis plan should take into account time-varying covariates in relation to the timepoint in the pregnancy.

12 Glossary

This Glossary relies on definitions sourced from ICH Guidelines, supplemented by regulatory documents, and then, relevant non-regulatory best practice documents from other sources such as professional society best practice documents and the literature.

Administrative Claims Data:
Data that arise from a person's use of the healthcare system and reimbursement of healthcare providers for that care.

(FDA, United States. Guidance Pharmacoepidemiologic safety studies using electronic data)

Bias:
A systematic deviation in results from the truth.

(Proposed by CIOMS Working Group X. Bias (CIOMS X: Meta-analysis 2016 | Japanese)

Case Definition:
The clinical, biological, psychological, and functional concepts of the condition, that reflect the medical and scientific understanding of the condition.

(FDA, United States. Guidance Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision Making for Drug and Biological Products)

Common Data Model

A mechanism by which raw data are standardised to a common structure, format and terminology independently from any particular study in order to allow a combined analysis across several databases/datasets. Standardisation of structure and content allows the use of standardised applications, tools and methods across the data to answer a wide range of questions

(A Common Data Model for Europe? – Why? Which? How? – workshop report EMA/614680/2018)

Conceptual Definition:
Explains a study construct (e.g., exposure, outcomes, covariates) or feature in general or qualitative terms.

(FDA, United States. Guidance Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products)

Confounding:
Confounding results from the presence of an additional factor, known as a confounder or confounding factor, that is associated with both the exposure and the outcome, and is not in the causal pathway between exposure and the outcome. Confounding distorts the observed effect estimate for the outcome and the exposure under study.

(The European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology)

Data Accuracy:
The degree of closeness of the measured value to the nominal or known true value under prescribed conditions (or as measured by a particular method).

(M10 EWG Bioanalytical Method Validation and Study Sample Analysis -- Step 4 (final); 24 May 2022)

Data Completeness:
The “presence of the necessary data” (National Institutes of Health 1263 Collaboratory 2014).

(FDA, United States. Guidance Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products)

Data Consistency:
Relevant uniformity in data across clinical sites, facilities, departments, units within a facility, providers, or other assessors.

(FDA, United States. Guidance Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products)

Data Curation:
The curation of the source data for the purpose of statistical analysis of specific clinical research questions. Data curation includes, but is not limited to, the following aspects: data extraction (including multiple data sources), data security processing (de-identification or anonymization, and protection from data corruption, leaking, theft, tampering, or unauthorized access), data

cleaning (edit check and outliers processing, data completeness processing), data conversion (common data models, normalization, natural language processing, medical coding, derived variable calculation), data quality control, data transmission and storage.

(NMPA, China. Guideline on Using Real-World Data to Generate Real-World Evidence (Trial Version) English Translation)

Data Holder:
A legal person, including public sector bodies and international organizations, or a natural person who is not a data subject with respect to the specific data in question, which, in accordance with applicable law, has the right to grant access to or to share certain personal data or non-personal data

(Article 2(8) of the REGULATION (EU) 2022/868 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act))

Data Provenance:
An audit trail that "accounts for the origin of a piece of data (in a database, document or repository) together with an explanation of how and why it got to the present place."

(FDA, United States. Guidance Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products)

Data Relevance:
Data relevance includes the availability of key data elements (exposure, outcomes, covariates) and sufficient numbers of representative patients for the study.

(FDA, United States. Guidance Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision Making for Drug and Biological Products)

Data Reliability:
Data reliability includes data accuracy, completeness, provenance, and traceability.

(FDA, United States. Guidance Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision Making for Drug and Biological Products)

Data Traceability:
Permits an understanding of the relationships between the analysis results (tables, listings, and figures in the study report), analysis datasets, tabulation datasets, and source data.

(FDA, United States. technical specifications document Study Data Technical Conformance Guide (October 2019))

Digital Health Technology (DHT):
A system that uses computing platforms, connectivity, software, and/or sensors for health care and related uses. These technologies span a wide range of uses, from applications in general wellness to applications as a medical device. They include technologies intended for use as a medical product, in a medical product, or as an adjunct to other medical products (devices, drugs, and biologics). They may also be used to develop or study medical products.

(FDA, United States. Digital Health Technologies for Remote Data Acquisition in Clinical

Investigations Guidance for Industry, Investigators, and Other Stakeholders)

Effect Modification:
Effect modification occurs when the effect of a single exposure on an outcome depends on the values of another variable, i.e., the effect modifier, which does not necessarily need to be involved in the causal pathway. Interaction occurs when there is interest in the causal effect of two exposures on an outcome and how the effect of either exposure depends upon the value of the other exposure.

(ENCePP)

Electronic Health Record Data:
A collection of an individual patient records contained within an EHR system. A typical individual EHR may include a patient's medical history, diagnoses, treatment plans, immunization dates, allergies, radiology images, pharmacy records, and laboratory and test results.

(FDA, United States. Data Standards for Drug and Biological Product Submissions Containing Real-World Data)

Estimand:
A precise description of the treatment effect reflecting the clinical question posed by the trial objective. It summarizes, at a population level, what the outcomes would be in the same patients under different treatment conditions being compared.

(ICH E9-R1 - Addendum: Statistical Principles for Clinical Trials, Glossary).

Exposure:
An exposure is the medicinal product or regimen of interest being evaluated in the proposed study (ICH M14 Expert Working Group).

Federated Data Network:
A series of decentralized, interconnected nodes, which allows data to be queried or otherwise analyzed by other nodes in the network without the data leaving the node it is located at. Examples of FDNs include DARWIN EU, Sentinel, CNODES, OHDSI, and MID-NET.

(Hallock H, Marshall SE, 't Hoen PAC, Nygård JF, Hoorne B, Fox C, Alagaratnam S. Federated Networks for Distributed Analysis of Health Data. Front Public Health. 2021;9:712569.)

Medical Claims Data:
A compilation of information on medical claims submitted to insurance companies for reimbursement of medical expenses for treatments and other interventions. Medical claims data use standardized medical codes, such as the World Health Organization's International Classification of Diseases Coding (ICD-CM) diagnosis codes, to identify diagnoses and treatments.

(FDA, United States. Data Standards for Drug and Biological Product Submissions Containing Real-World Data)

Medicine:
Any substance or combination of substances intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease.

(Section 201(g) of the Federal Food Drug and Cosmetic Act (FD&C Act).)

Operational Definition:
The data-specific operation or procedure a researcher followed to measure constructs in a particular study.

(FDA, United States. Guidance Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision Making for Drug and Biological Products (Draft))

Patient Experience Data
Data that are collected by any persons and are intended to provide information about patients' experiences with a disease or condition. Patient experience data can be interpreted as information that captures patients' experiences, perspectives, needs, and priorities related to (but not limited to): 1) the symptoms of their condition and its natural history; 2) the impact of the conditions on their functioning and quality of life; 3) their experience with treatments; 4) input on which outcomes are important to them; 5) patient preferences for outcomes and treatments; and 6) the relative importance of any issue as defined by patients.
(Title III, section 3001 of the 21st Century Cures Act, as amended by section 605 of the FDA Reauthorization Act of 2017 [FDARA])

Phenotype / Phenotype Algorithm:
Observable and measurable information that is relevant to health or healthcare such as a disease (e.g., type 2 diabetes), a blood pressure measurement, a blood sugar value or an antibiotic prescription. It can be used to define any patient characteristics, from exposure to outcome. The translation of the case definition into an executable algorithm that involves querying clinical data elements from the EHRs is the Phenotyping algorithm. These algorithms identify and extract data from health records using clinical codes (for example ICD-10 or SNOMED). They can also be referred to as "electronic phenotype" or "computable phenotype".

(www.ohdsi.github.io (The Book of OHDSI))

Plausibility:
The believability or truthfulness of data values.

(FDA, United States. Guidance Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision Making for Drug and Biological Products citing Kahn et al. 2016).

Primary Data Collection:
Data collected specifically for the present study.

(Adapted from ICH E8)

Quality Assurance (QA):
All those planned and systematic actions that are established to ensure that the *study* is performed and the data are generated, documented (recorded), and reported to an appropriate quality standard and applicable regulatory requirements.

(Adapted from E6(R2) Good Clinical Practice (GCP) - *Step 4* (final); 9 November 2016 – Glossary)

Quality Control (QC):
The operational techniques and activities undertaken within the quality assurance system to verify that the requirements for quality of the study-related activities have been fulfilled.

(Adapted from E6(R2) Good Clinical Practice (GCP) -- *Step 4* (final); 9 November 2016 – Glossary)

Quantitative Bias Analysis:
Quantitative bias analysis is an overarching term applied to methods that estimate quantitatively the direction, magnitude, and uncertainty associated with systematic errors that influence measures of associations.

(Lash TL, Fox MP, Cooney D, Lu Y, Forshee RA. Quantitative Bias Analysis in Regulatory Settings. Am J Public Health. 2016;106(7):1227-30.)

Real-World Data (RWD):
Data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources.

Examples of RWD include data derived from electronic health records (EHRs); medical claims and billing data; data from product and disease registries; patient-generated data, including from mobile devices and wearables; and data gathered from other sources that can inform on health status (e.g., genetic and other biomolecular phenotyping data collected in specific health systems).

(Adapted from FDA, United States. Guidance Real-World Data: Assessing Registries To Support Regulatory Decision-Making for Drug and Biological Products DECEMBER 2023 and

FDA, United States. Draft Guidance Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision Making for Drug and Biological Products)

Real-World Evidence
The clinical evidence about the usage and potential benefits or risks of a medicinal product derived from analysis of RWD.

(FDA, United States. Guidance Real-World Data: Assessing Registries To Support Regulatory Decision-Making for Drug and Biological Products DECEMBER 2023)

Registry:
A registry is an organized system that collects prespecified uniform data from a population defined by a specific disease, condition, or exposure.

(Adapted from: FDA Real-World Data: Assessing Registries To Support Regulatory Decision-Making for Drug and Biological Products DECEMBER 2023 and

EMA Guideline on registry-based studies 24 September 2020)

Secondary Use of Data:
Use of existing data for a different purpose than the one for which they were originally collected.

(EMA Guideline on registry-based studies)

Standard of Care:

Treatment that is accepted by medical experts as a proper treatment for a certain type of disease or condition and that is widely used by healthcare professionals. Also called best practice, standard medical care, or standard therapy.

(National Cancer Institute Dictionary)

Statistical Analysis Plan:

A statistical analysis plan is a document that contains a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and includes detailed procedures for executing the statistical analysis of the primary and secondary variables and other data.

(E9 Statistical Principles for Clinical Trials -- Step 4 (final); 5 February 1998 – Glossary)

Target Trial:
A hypothetical randomized trial that would answer the question of interest if it were feasible.

(Adapted from: National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Care Services; Committee on Developing a Protocol to Evaluate the Concomitant Prescribing of Opioids and Benzodiazepine Medications and Veteran Deaths and Suicides. An Approach to Evaluate the Effects of Concomitant Prescribing of Opioids and Benzodiazepines on Veteran Deaths and Suicides. Washington (DC): National Academies Press (US); 2019 Sep 24. 2, Specifying the Target Trial.)

1071

13 Regulatory Guidelines Referenced

European Medicines Agency. Clinical Trials Regulation (EU) NO 536/2014; 2014 Apr.

European Medicines Agency. Guideline on Good Pharmacovigilance Practices (GVP) – Module III. Revision 1; 2014 Sep.

European Medicines Agency. Guideline on Good Pharmacovigilance Practices (GVP) – Module IX. Revision 1; 2017 Oct.

European Medicines Agency. Guideline on Registry Based Studies. Amsterdam (NL); 2021 Oct.

European Medicines Agency. Guideline on the Content, Management and Archiving of the Clinical Trial Master File (Paper and/or Electronic). London (GB); 2018 Dec.

European Medicines Agency. ICH E2D Post-Approval Safety Data Management. London (GB); 2004 May.

European Medicines Agency. ICH E8 (R1) General Considerations for Clinical Studies. Revision 1. Amsterdam (NL); 2021 Oct.

European Medicines Agency. ICH E9 (R1) Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the guideline on statistical principles for clinical trials. Amsterdam (NL); 2020 Feb.

European Medicines Agency. ICH Guideline for Good Clinical Practice E6(R2). London (GB); 2016 Dec.

European Medicines Agency. ICH Topic E9 Statistical Principles for Clinical Trials. London (GB); 1998 Sep.

European Medicines Agency. Patient experience data in EU medicines development and regulatory decision-making. Outcome of the workshop on 21st September 2022. 17 October 2022, EMA/786952/2022.

European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP). ENCePP Checklist for Study Protocols. Revision 4. EMA; 2018 Oct.

European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on Methodological Standards in Pharmacoepidemiology. Revision 11. Amsterdam (NL): EMA; 2023 July.

FDA, United States. Draft Guidance for Industry: Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products; 2021 Sep.

FDA, United States. United States. Final Guidance: Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data; 2013 May.

FDA, United States. Framework for FDA's Real-World Evidence Program. Silver Spring (US); 2018 Dec.

FDA, United States. Guidance for Industry: Considerations for the Use of Real-World Data and Real World Evidence To Support Regulatory Decision-Making for Drug and Biological Products; 2023 Aug.

FDA, United States. United States. United States Guidance for Industry: Data Standards for Drug and Biological Product Submissions Containing Real-World Data; 2023 Dec.

FDA, United States. Guidance for Industry: Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products; 2023 Dec.

FDA, United States. Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making; 2023 July.

Health Canada, Canada. Elements of Real-World Data/Evidence Quality Throughout the Prescription Drug Product Life Cycle; 2019 Mar.

ICH. ICH Reflection Paper Proposed ICH Guideline Work to Advance Patient Focused Drug Development; 2021 Jun.

NMPA, China. Guidance for Real-World Data Used to Generate Real-World Evidence; 2021 Apr.

NMPA, China. Guidance on the Use of Real-World Evidence to Support Drug Development and Regulatory Decisions; 2020 Jan.

NMPA, China. Guidance on Using of Real-World Study to Support Pediatric Drug Development and Regulatory Evaluation; 2020 Aug.

MHLW/PMDA, Japan. Basic principles on the utilization of health information databases for Post-Marketing Surveillance of Medical Products; 2017 Jun.

MHLW/PMDA, Japan. Guidelines for the Conduct of Pharmacoepidemiological Studies in Drug Safety Assessment with Medical Information Databases; 2014 Mar.

World Health Organization. A Practical Handbook on the Pharmacovigilance of Medicines Used in the Treatment of Tuberculosis: Enhancing the Safety of the TB patient. Geneva (CH): WHO Press; 2012.

www.fda.gov (Patient-Focused Drug Development Glossary 2018.)

www.fda.gov FDA Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making 2024.

www.imi-prefer.eu (How to do a patient preference study.)

www.reganudall.org (Real World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision Making for Drug and Biological Products Guidance for Industry 2021 Nov.)

14 Non-regulatory Guidelines Referenced

International Society for Pharmacoepidemiology. Guidelines for Good

Pharmacoepidemiology Practices (GPP). Washington, DC (US): ISPE; 2015 Jun.

The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement.

The REporting of studies Conducted using Observational Routinely-collected health Data Statement for Pharmacoepidemiology (RECORD-PE).

"Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals," established by the International Committee of Medical Journal Editors (ICMJE).

15 References

1. Wang SV, Pottegard A, Crown W, Arlett P, Ashcroft DM, Benchimol EI, et al. 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. Pharmacoepidemiol Drug Saf. 2023;32(1):44-55.
2. Gatto NM, Campbell UB, Rubinstein E, Jaksa A, Mattox P, Mo J, et al. The Structured Process to Identify Fit-For-Purpose Data: A Data Feasibility Assessment Framework. Clin Pharmacol Ther. 2022;111(1):122-34.
3. Schneeweiss S, Rassen JA, Brown JS, Rothman KJ, Happe L, Arlett P, Dal Pan G, Goettsch W, Murk W, Wang SV. Graphical Depiction of Longitudinal Study Designs in Health Care Databases. Ann Intern Med. 2019 Mar 19;170(6):398-406).
4. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64.
5. Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, Bartels DB, Kahler KH, Bessette LG, Schneeweiss S. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. BMJ. 2021 Jan 12;372:m4856. doi: 10.1136/bmj.m4856. PMID: 33436424; PMCID: PMC8489282.