17 March 2025
EMA/99865/2025
Committee for Human Medicine Products/Methodology Working Party (CHMP/MWP)

Reflection paper on use of real-world data in non-interventional studies to generate real-world evidence for regulatory purposes

Draft agreed by Methodology Working Party (MWP)	October 2023
Adopted by CHMP PROM for release for consultation	15 April 2024
Start of public consultation	3 May 2024
End of consultation (deadline for comments)	31 August 2024
Agreed by Methodology Working Party (MWP)	March 2025
Adopted by CHMP PROM	17 March 2025

Keywords	Non-interventional study, real-world data, real-world evidence, feasibility assessment, bias, confounding, data quality

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

Table of contents

1. Introduction

This reflection paper discusses methodological aspects of non-interventional studies (NIS) using real-world data (RWD) in order to generate real-world evidence (RWE) for regulatory purposes.

A NIS is a clinical study that does not fulfil any of the conditions defining a clinical trial (CT) in Article 2.2 of Regulation (EU) No 536/2014 on clinical trials on medicinal products for human use, where a clinical trial is defined as follows: (a) the assignment of the subject to a particular therapeutic strategy is decided in advance and does not fall within normal clinical practice of the Member State concerned; (b) the decision to prescribe the investigational medicinal products is taken together with the decision to include the subject in the clinical study; (c) diagnostic or monitoring procedures in addition to normal clinical practice are applied to the subjects.

RWD are data that describe patient characteristics (including treatment utilisation and outcomes) in routine clinical practice. RWE is evidence derived from the analysis of RWD.

CTs generally use randomisation, blinding and a controlled environment, and are the main source of evidence to evaluate the benefits and risks of medicines in marketing authorisation procedures. The use of NIS for assessing the safety and effectiveness of medicines is often limited by the absence of randomisation, uncontrolled conditions, non-standardised treatments and uncertainties regarding data quality and completeness. However, the ability to capture electronic healthcare data and data from registries is now providing new opportunities to use RWD in NIS and generate RWE that reflects clinical practice. As such, NIS using RWD can complement and support data from RCTs by filling gaps in knowledge and reducing uncertainties about a product's safety and effectiveness.

Examples where NIS using RWD have supported regulatory assessment include:

- To characterise disease epidemiology (incidence, prevalence, risk factors and progression).
- To understand the clinical context by describing standards of care, variability in clinical practices, and unmet medical needs.
- To describe patterns of drug utilisation (e.g. indications, characteristics of users of medicines, incidence and prevalence of use, doses, duration, and patterns of use).
- To support the feasibility assessment and the planning of non-interventional post-authorisation safety, effectiveness and drug utilisation studies by measuring outcome incidence, treatment exposure, the duration of available follow-up and the impact of applying different eligibility criteria on sample size.
- To compare patient characteristics of a study population to those of the clinical practice population in a real-world setting.
- To perform post-marketing monitoring, investigate safety concerns and effectiveness, and evaluate the effectiveness of risk minimisation measures.

Given the large amount of information that NIS using RWD can generate for regulatory purposes, it is important to understand their limitations as well as how some of these limitations could be overcome or mitigated to increase the reliability of the evidence. This reflection paper is therefore relevant to all stakeholders involved in the planning, conduct and analysis of NIS using RWD to generate RWE to be submitted for regulatory purposes in the EU, regardless of whether the NIS is conducted in the EU or elsewhere. These stakeholders include Marketing Authorisation Holders (MAHs) and Applicants, regulatory authorities, Health Technology Assessment bodies, payers, academia, RWD holders and healthcare professionals' and patients' associations.

2. Scope

The scope of this reflection paper is the design, conduct and analysis of NIS using RWD to generate RWE for regulatory purposes. The use of RWD in the context of CTs, e.g. pragmatic trials or externally controlled trials, or to serve as a data source to recruit individuals for a CT, is out of scope of this document.

General methodological principles, approaches for the conduct of NIS and epidemiological terms used in this context are described in textbooks and scientific guidelines (1). This reflection paper focuses on methodological principles that are considered important for the conduct and assessment of NIS using RWD and used for regulatory decision-making throughout a medicine's lifecycle.

A large variety of RWD can be used in NIS, such as clinical data, data related to healthcare services utilisation, medical claims, prescribing and dispensing of medicinal products, socio-economic and lifestyle data, patient experience data, data collected with wearable biometric devices and genetic data. A critical aspect when assessing the suitability of RWD for a regulatory purpose is data quality, including data reliability and relevance as described in Chapter 7. , and, depending on the research question, the extent to which the data reflects routine clinical practice. RWD may originate from primary data collection, i.e. data collected from patients, caregivers, healthcare professionals or other persons involved in patient care specifically for the study in question, or from secondary use of data, i.e. use of existing data for a different purpose than the one for which it was originally collected. Most recommendations expressed in this reflection paper apply to both types of data, but specific recommendations applying to primary data collection or secondary use of data are specified in the text.

This reflection paper makes a distinction between NIS having descriptive objectives and NIS having causal objectives. This distinction has important implications for the study design. A study with descriptive objectives is designed to describe patient characteristics without regards to any causal hypothesis, but it may include a measure of association between the distribution of these characteristics and the categories of other variables. A study with causal objectives is designed to investigate the effect, causative or preventive, of an exposure in comparison to what would have happened to the same individuals under non-exposure or another exposure. In the context of this reflection paper, the exposure is generally a medical intervention, and the outcome of interest is generally a measure of its safety or effectiveness. Reference to studies with causal objectives in this document does not imply an expectation that specific analytical methods will be used.

3. Legal obligations and regulatory requirements

The relevance and feasibility of including NIS using RWD in a regulatory procedure should be discussed at an early stage between the MAHs or Applicants and EMA, through e.g. scientific advice and protocol assistance procedures. It is the responsibility of the MAHs or Applicants to include the database holders and any other stakeholders, such as patients' organisations, in the discussion.

NIS using RWD may be proposed to fill knowledge gaps, but its relevance depends on the intended regulatory purpose within the context of a specific application. The regulatory assessment does not mandate a specific study design but requires that the evidence generated can support the regulatory objective. The relevance of a NIS using RWD to generate RWE for a specific application can therefore only be determined on a case-by-case basis.

The legal obligations and regulatory requirements applicable to NIS should be followed. The following documents are particularly relevant in the context of this reflection paper.

Legal obligations

- Regulation (EU) No 536/2014 on clinical trials on medicinal products for human use.
- Regulation (EC) No 726/2004 laying down Community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency.
- Commission Implementing Regulation (EU) No 520/2012 on the performance of pharmacovigilance activities.
- Directive 2001/83/EC on the Community code relating to medicinal products for human use.
- Regulation (EU) 2016/679 (GDPR) on the protection of natural persons with regard to the processing of personal data and on the free movement of such data.
- Regulation (EU) 2018/1725 (EUDPR) on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data

Regulatory requirements

- Guideline on good pharmacovigilance practices. Module VIII -Post-authorisation safety studies.
- Guideline on good pharmacovigilance practices: Module XVI – Risk minimisation measures: selection of tools and effectiveness indicators.
- Scientific guidance on post-authorisation efficacy studies and Post-authorisation efficacy studies: Questions and Answers.
- Guideline on registry-based studies.
- Good Practice Guide for the use of the Metadata Catalogue of Real-World Data Sources V 1.0
- European Medicines Agency. Reflection paper on the use of Artificial Intelligence (AI) in the medicine product life cycle
- European Medicines Agency (EMA) and Heads of Medicines Agencies (HMA) Data Quality Framework for EU medicines regulation.
- European Medicines Agency. Data Quality Framework for EU medicines regulation: application to Real-World Data.
- ICH E9 (R1). Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical principles for Clinical Trial.
- ICH M14. Guideline on general principles on plan, design and analysis of pharmacoepidemiological studies that utilize real-world data for safety assessment of medicines.

4. Study design

4.1. General Considerations

The design of the NIS should be primarily driven by the need to obtain reliable evidence regarding the research question. It is therefore essential that the research question is expressed with sufficient detail and attention to the regulatory question as it forms the basis for the selection of data source(s), study design, analysis approach and discussion regarding the feasibility of the study to meet regulatory objectives. It is the MAH's and Applicant's responsibility to justify that the use of RWD is appropriate

and feasible to meet the pre-defined study objectives. If a research question results in a causal objective, this should be clearly stated in the study protocol and the corresponding methodological requirements should be addressed.

Methodological standards for NIS and use of RWD, such as those described in the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology (1), should be applied. The description of the study design should be supported with a graphical representation (2). The ENCePP Checklist for Study Protocols (3) should be included as an Annex to the protocol of a non-interventional study submitted to a regulatory authority, as a check that important study design components have been addressed.

4.2. Feasibility assessment

A feasibility assessment is recommended prior to writing the study protocol to guide its development and facilitate early discussions with regulatory authorities. It is a preparatory document for the protocol, and it does not replace the submission of the protocol and statistical analysis plan. It should be prepared by the party intending to submit the study results as part of a regulatory application.

The feasibility assessment should include:

- in case of primary data collection, the justification and method for study site(s) and population selection, the method and feasibility of data collection, the need for informed consent and ethical approvals, the enrolment projections and the capacity to identify, collect and report adverse events if applicable;
- in case of secondary data use, a characterisation of the RWD source(s) with an evaluation of their reliability and relevance in order to meet the study objectives (see Chapter 7.); this discussion should address, as appropriate, the feasibility of the planned study design based on the availability of data on exposures, outcomes, and covariates of interest;
- a discussion of the choice of inclusion/exclusion criteria, how they may impact sample size and of whether the available sample size may provide sufficient precision of key estimates, overall and in subgroups of interest (e.g. country-specific estimates in a descriptive drug utilisation study);
- as applicable, data on the incidence of exposure and study outcomes in the source population for the study, as they will inform on precision but also on anticipated timelines for the study to meet the regulatory objectives;
- a discussion on how the above evaluation may impact the milestones for the study;
- as applicable, options to increase feasibility.

When an existing RWD source is proposed to be used, the feasibility assessment should be performed in collaboration with the data source holder to ensure the timely availability of data for regulatory decision-making and set up realistic timelines for the completion of the study. If applicable, exploratory descriptive analyses should be conducted to document the study feasibility.

The recommendations and checklist for the feasibility assessment provided in the Guideline on registry-based studies can be considered and adapted.

4.3. Studies with descriptive objectives

Studies with descriptive objectives may be conducted for different purposes, including risk assessment or prediction and service evaluation. They generally aim to observe and accurately measure patient characteristics at a single time point or over time.

Depending on the research question, it may be essential that the study population is representative of the real-world target population. When some of the characteristics studied may be influenced by the setting in which they are observed, the study should pay attention to the conditions that may influence the results of the study and this information should be addressed in the study design and analysis to help understand their impact, e.g. through sensitivity analyses. The aspects to be considered may include:

- The healthcare setting where the RWD are collected, e.g. primary care, specialist care, hospital care, and the type of RWD, e.g. claims data, data on medicines prescription or dispensing or registries, which may result in selection mechanisms that can influence the study feasibility and results.
- The healthcare system of the country(-ies) where the RWD are collected, which may influence the availability and accessibility of exposure and outcome data, for example data related to specialist care in hospitals, and the possible duplication of data if patients may consult different healthcare professionals without the records being linked.
- Regional differences in clinical practice and healthcare systems management, e.g. diagnostic criteria, prescribing practices, prescribing formularies, coding practices or reimbursement policies; contemporary clinical practice and any change of relevance over time should be highlighted.
- The specificities of coding terminologies for medicinal product exposure and clinical events, use of a common data model and how data quality is assessed and managed (e.g. data quality metrics, data quality controls, misclassification and missingness, benchmarking).

4.4. Studies with causal objectives

The causal interpretation of any treatment effect requires a comparator to quantify the effect. The aim of the study design in studies with causal objectives should be to achieve valid comparisons between exposure groups by mitigating the risk of selection bias, information bias and confounding (see Chapter 5.).

The recommendations provided for studies with descriptive objectives (Chapter 4.3.) should also be considered for studies with causal objectives.

The target trial emulation (TTE) framework should be considered as a strategy that uses existing tools and methods to formalise the design and analysis of NIS using RWD with causal objectives (4, 5, 6). The first step of this framework is to specify the key elements of a hypothetical (target) trial that would answer the research question. The second step is to design a NIS that emulates the hypothetical trial using epidemiological methods.

The TTE framework is considered useful for the following reasons:

- it provides a structured and coherent framework for the design of NIS with a causal objective, with similarities with CTs in terms of terminology, definition of the estimand and analytical approaches;
- it helps the investigators to consider potential bias and adequate methods to address them;

- given the need to explicitly describe the design elements needed to emulate the CT, it provides a high level of transparency on the study design, the assumptions needed to emulate the trial and the definition of causal effects; this level of transparency may facilitate the evaluation and the replication of the study;
- the detailed and structured definition of inclusion/exclusion criteria and allocation of time periods defined by aligning study entry and estimated start of treatment have been shown to reduce bias, such as the prevalent user bias (7) and the immortal-time bias (8).

Although the TTE framework can improve the internal validity of the NIS, the lack of randomisation and blinding still requires attention to the prevention and/or control of selection bias, information bias and confounding described in Chapter 5. .

The estimand framework described in the ICH E9 (R1) Addendum on Estimands and Sensitivity Analysis in Clinical Trials should be considered in the design of the hypothetical trial, such as the attributes of the estimand, intercurrent events and strategies to handle intercurrent events. This will help increase the coherence between definitions of exposures, outcomes and intercurrent events between the NIS and the target trial, The main statistical analysis may also be aligned with the estimand framework, e.g. concerning the approach to missing data handling and sensitivity analyses.

5. Bias, confounding and effect modification

5.1. General considerations

The non-interventional nature of NIS may lead to bias that distorts the measure of association due to processes of selection (selection process in the overall dataset, systematic differences in the selection of study groups or differences in loss-to follow-up between study groups), misclassification (differences in the classification of individuals as regards their exposures, outcomes or covariates, including time-related classification and measurement errors in continuous variables), and confounding (difference in the risk of developing the outcome of interest). If not adequately prevented or controlled, bias and confounding may limit the causal interpretation of the results.

Potential effect modification of interest needs to be addressed to identify and describe study population characteristics that may affect the generalisability of the study results to a defined target population.

5.2. Selection bias

For many of the data sources used, there are potential selection mechanisms that need to be addressed not only in the feasibility assessment but also later in the design of the study. Selection bias is difficult and often impossible to address in the analysis stage. At the design stage, the following aspects should therefore be considered.

- The selection of specific study sites and populations in case of primary data collection and of specific existing data sources in case of secondary use of data should be justified on scientific grounds and discussed in terms of characteristics of the individuals enrolled in the NIS and potential differential loss to follow-up across study groups.
- Any inclusion and exclusion criteria should be adequately defined and justified in the protocol with a description of the diagnostic and/or procedural codes and of any algorithm used to include patients in the analysis. This description should also address the completeness and possible misclassification of the data used to define the inclusion and exclusion criteria. If misclassification may be present, the possible impact on the study results should be addressed.

In case of secondary use of data, it is generally recommended to use wide inclusion criteria when the data are extracted from the original RWD source, if applicable, and apply stricter inclusion/exclusion criteria in the analysis stage. This allows for stratified analyses and sensitivity analyses that can inform the interpretation of results.

- The definition of inclusion and exclusion criteria for the study population should consider implicit selection criteria resulting from the method used to identify the study population and define exposure categories, as these selection criteria may not be balanced between these categories. For example, differences between healthcare seeking behaviours in vaccinated and non-vaccinated persons identified through the healthcare system may be related to their socio-economic status or other factors influencing their probability of being vaccinated and their probability of presenting with the outcome.
- Depending on the research question, a new (incident) drug user design (9) can be considered instead of including both prevalent and incident drug users. Prevalent drug users are patients already taking the study treatment or a similar treatment before the start of follow-up. Including such patients can cause selection bias because patients who experience the outcome of interest early during treatment become underrepresented (i.e. depletion of susceptibles). Bias can also arise if exposure and confounding factors are time-dependent or if variables impacting medicine prescription at study entry (e.g. disease severity) are influenced by use of previous medicine that is part of the study exposure. The use of prevalent new user cohorts can also be used in some situations, e.g. to allow inclusion of initiators of the new medicine who were previously on an older comparator (10).
- Comparisons of study populations from different sites, RWD sources or time periods may introduce bias. The variables influencing the inclusion of individuals in these study populations may vary across time and settings and may not be known or measured, hence they may affect the exposure status and/or the study outcome. This design therefore needs to be justified and the likelihood of bias and confounding should be recognised.

5.3. Information bias

Information bias may arise when key study variables (exposure, outcome, or confounders) are inaccurately measured or classified. Misclassification can arise at many different steps of data collection and extraction: diagnosis, coding, recording, data transformation, data aggregation, summarisation, and analysis. The following is recommended:

- In case of primary data collection, the study design should give considerations to ensuring accuracy of measurements and minimising potential missingness of data.
- In case of secondary use of data, the different steps of data collection or extraction applied in the RWD source should be identified at the stage of study design. Ideally, these steps should be verified to evaluate if the data source(s) contain(s) enough details on exposures (e.g. dose, duration, time periods and indication), outcomes (e.g. diagnostic code, disease severity and date of occurrence), confounders and eligibility criteria to correctly classify the patients.
- Any validation study previously performed should be identified and evaluated. A new validation study may be proposed as part of the feasibility analysis.
- Misclassification is traditionally categorised as differential or non-differential. It is often stated that differential misclassification can lead to biases in any direction, whilst non-differential misclassification generally drives the association between the exposure and the outcome towards the null value. From a regulatory perspective, non-differential misclassification is

often presented as being preferable for superiority studies as it is conservative. At the design stage this reasoning should be avoided, as the assumption that misclassification will be non-differential is difficult to verify or requires additional analyses that may not be done. In studies with causal objectives, the impact of misclassification also largely depends on study objectives that may co-exist in a same study, e.g. non-inferiority or superiority, effectiveness or safety. It is therefore more important to identify potential misclassification of exposures, outcomes and relevant covariates to minimise and measure it during the study if possible.

5.4. Time-related bias

Time-related bias, including immortal time bias (11), is an information bias that may occur in cohort studies where the exposure status may change over time and the allocation of time periods of observations to the non-exposed/exposed person-time is incorrect. It is therefore recommended:

- to define and align at the design stage for all included individuals the timepoint of eligibility, treatment initiation and start of follow-up to prevent the occurrence of time-related bias, to define changes of patient exposure or outcome status, and plan the data collection or extraction of important dates;
- to consider the appropriate assumptions concerning risk windows and ensure that the available length of follow-up is sufficient to include the relevant risk windows;
- to plan sensitivity analyses to evaluate the impact of these assumptions;
- to include in the study protocol graphical representations of the study design and study diagrams (2);
- to consider the target trial emulation framework to decrease the risk of wrongly assigning unexposed person-time as exposed or vice versa (see Section 4.4.).

5.5. Confounding

Confounding should be addressed at the design stage. Confounders may be unknown, or known but inadequately measured. The chosen approach for the identification of known confounders has to be clearly described. They should be systematically identified and clearly stated at the design stage, and the design should attempt to minimise their impact on the results. The following is recommended:

- Potential confounders (risk factors for the outcome of interest) should be identified from various sources (e.g., disease knowledge and previous studies identified through systematic literature search) to plan the data collection or extraction for the variables to be accounted for; some confounders such as age, sex, socio-economic status, and geographic location are common to many studies. Any potential confounders should be identified irrespective of availability of measured confounders in the available RWD. It is particularly important to identify potentially important unmeasured confounders.
- The analytical methods to address potential confounding should be pre-specified in the protocol or analysis plan.
- For studies with causal objectives, use of an active comparator should be preferred to a non-pharmacological comparator or a non-user comparator to increase the similarity in measured patient characteristics between treatment groups and reduce potential confounding by indication, disease severity or unmeasured variables (9). Its use is optimal in the context of the new user design whereby comparison is between patients with the same indication

initiating different treatments. Depending on the study objectives, a justification should be provided if the use of an active comparator was considered not appropriate or not feasible.

- For studies with causal objectives, inclusion of a control exposure and/or control outcome should be considered to help in the interpretation and appraisal of results. A negative control exposure (exposure known not to be causally associated with the outcome of interest) or negative control outcome (outcome known not to be causally associated with the exposure of interest) can help assess the presence of residual confounding by revealing an association where none is expected. A positive control exposure (exposure known to be causally associated with the outcome of interest) or positive control outcome (outcome known to be causally associated with the exposure of interest) can help identify confounding towards the null value by revealing a reduced or null association where a clear departure from the null is expected. The choices between a negative and a positive control and between a control exposure and a control outcome depend primarily on the study objective, taking also into account the available data and the clinical context.

5.6. Effect modification

- In studies with causal objectives, any potential effect modification of interest should be addressed at the design phase in order to verify the availability of relevant data in the data sources planned to be used, determine whether the sample size will be sufficient to allow appropriate analyses to characterise effect modification, and pre-specify the analytical approach in the statistical analysis plan.
- In case the study population characteristics may influence the treatment effect estimates and affect the generalisability of the study results to a defined target population, a comparison should be made between the cohort's key characteristics and those of the target population using other information (e.g. published research or national statistics). The effects should be measured across relevant characteristics through sensitivity analyses. In some studies, such as registry-based studies, it may be possible to compare the characteristics of patients included and excluded in the study, such as age, sex, socio-economic status, disease severity and medication use.

6. Governance and transparency

6.1. Governance

- The study governance should follow the principles described in the ENCePP Code of Conduct (12)and the Good Pharmacoepidemiology Practices (GPP) of the International Society for Pharmacoepidemiology (ISPE) (13).
- The conduct of NIS in the EU needs to follow the applicable EU data protection rules at each step of the processing of personal data, including the option for data sharing/pooling between data source holders and other stakeholders like competent authorities and MAHs and Applicants, i.e. the General Data Protection Regulation (EU) 2016/679 (GDPR), which applies to processing carried out by organisations and bodies operating within the EU, and the Regulation (EU) 2018/1725 (EUDPR), which applies to EU institutions, bodies, offices and agencies.
- In case of primary data collection, a study governance team involving clinicians, patients' representatives and disease experts may help set up the processes for data definition, collection and reporting. Ethical approval of the data elements to be collected at a local or

national level may be required. The ethical and procedural obligations may require to obtain informed consent from individuals before participating in the study. The informed consent should clearly outline areas such as an explanation of the purposes of the study, the expected duration, intended use of their data and cover all data to be accessed and processed as specified in the study protocol (including but not limited to the access for monitoring, auditing or inspections by competent authorities). It should also provide information on how the results of the study will be used and disseminated (14).

- In case of secondary use of the data, knowledge of the governance applied to the RWD used in a study (data source, raw data and results) helps understand any restrictions related to the conditions to access and publish the data, including constraints that may affect the availability of some data. Governance principles applied to the data source should therefore be adequately described and disclosed to regulators. Any requirement for ethical approval and informed consent, if applicable, should be described.

6.2. Transparency

Transparency is essential to share study information and support the evaluation, interpretation, and reproducibility of results. The main tools for the transparency of NIS are the HMA-EMA Catalogues of RWD sources and studies (15).

MAHs, Applicants and concerned stakeholders should:

- register in the HMA-EMA Catalogue of RWD studies all NIS performed by the MAH or Applicant, together with the study protocol and study report;
- ensure that the data sources used in the NIS submitted to regulators are registered in the HMA-EMA Catalogue of RWD sources ; if a data source is already registered, it would be appropriate to update this information if the last update was performed more than 12 months ago; this provision may be included in the contractual agreement between the MAH or Applicant and the data owner, as relevant;
- report study results using or adapting the format recommended for the post-authorisation safety studies.
- make publicly available the codes used for the creation of the analytical dataset and the programming code for the statistical analyses.

7. Data quality

The HMA-EMA Data Quality Framework for EU medicines regulation and its RWD specific recommendations (16) provide guidance for assessing the data quality of RWD, collected by either primary data collection or secondary use of data, with the goal of improving the usefulness of RWE for regulatory purposes.

In case of secondary use of data, RWD are often used in NIS without the ability to influence the way they were collected, coded and recorded. For this reason, it is essential to adequately characterise the dataset and assess the quality of the RWD that will be used for a specific study and present it in the feasibility assessment (see Chapter 4.2.). At least two key dimensions of data quality should be addressed in the study protocol: reliability, which is the property of the data irrespective of their use in any specific study, and relevance, which should be evaluated in relation to specific study objectives and data needs.

In case artificial intelligence technologies are used to collect and process RWD used in a NIS, the methodologies applied to evaluate their performance, the risk of bias and the impact on the results should be detailed in the study protocol (17).

7.1. Reliability

Reliability determines whether data represent the intended underlying medical concepts and are complete, trustworthy, and credible. Different dimensions of reliability can be evaluated and documented by using a data quality framework (see Chapter 7.5.). The following is recommended:

- Adequate information on methods and results of the evaluation of reliability should be made available in the study protocol as an Annex or a linked document. Reference can be made to the description of the RWD source, its data elements and the quantitative information included in the HMA-EMA Catalogue of RWD sources or other publications with more recent information.
- Information regarding the standard data quality management applied to the RWD source should also be made available, such as the steps of data cleaning, extraction and transformation, the data quality checks applied to detect logical inconsistencies and erroneous, missing or out-of-range values, the validation of algorithms to extract and code data, and the remedial actions that are taken at the level of the data source.
- Any validation performed on the data source should be highlighted in the study protocol.
- It is expected that some data quality issues inherent to the data source deemed appropriate for the study objective are difficult or impossible to resolve, and therefore that uncertainties on some data quality aspects will remain. These uncertainties and their possible impact on the study results should be clearly identified in the feasibility assessment and the study protocol.

7.2. Relevance

Relevance determines, for a given research question, whether key data elements (exposure, outcomes, covariates) are available, the size of the study population is adequate, the study population is representative and of sufficient coverage of the target population for the study objective, and the study design is appropriate to fully answer the research question. The following is recommended:

- Relevance is study-specific and should be discussed in the feasibility assessment and the study protocol. The HARmonised Protocol Template to Enhance Reproducibility (18) provides a format for the presentation of information on the study design, exposure, outcome, and covariates and should be considered as a complement to the applicable protocol formats.
- Depending on the research question, summary statistics of important variables (e.g., age, sex, disease severity, medication use) may be presented in comparison to RWD from the literature or other sources to understand the value of the RWD source to fulfil the study objectives.
- The evaluation of the relevance of the proposed RWD source and study design should lead to a justification in the feasibility analysis and study protocol of why they are deemed fit-for-purpose to answer the research question.

7.3. Multi-database studies

The number of RWD sources available for secondary use is increasing, and this increase is associated with an increasing number of multi-database studies (19).

If several data sources are used in a study, the following is recommended:

- A federated approach, possibly based on a common data model, could be used whereby the data stay local and are analysed locally, and only aggregated data are shared externally. Alternatively, a single database could be constructed from the multiple data sources.
- The number of data sources and the associated increase in sample size should not reduce the quality requirements, and information on the reliability and relevance of each of the RWD source should be presented.
- Whilst consistency of conclusions across different sources may provide stronger confidence in the evidence, any heterogeneity between results generated from different RWD sources based on a same protocol is a source of uncertainties and needs to be addressed in the final study report. Qualitative and quantitative assessments of the sources of heterogeneity should be provided to better understand potential differences in the results across data sources (see Section 8.7.).

7.4. Data linkage

Data sources may need to be linked to combine individual-level RWD on exposure, outcomes, and covariates from different sources, such as genetic data, mother-child data or data from different registries using a unique patient identifier. For these studies, the study protocol should describe:

- The data elements used to link the data.
- The linkage methodology, including a description of how the performance of the matching will be measured in case of probabilistic matching; the impact of an imperfect matching on the study results needs to be evaluated and discussed in relation to possible bias.

7.5. Data quality frameworks

Several data quality frameworks provide a set of characteristics determining the fitness-for-use of data (20). While they differ by the number of dimensions and the names given to these dimensions, they do not diverge substantially in terms of definitions. It is therefore recommended:

- To follow the HMA-EMA Data Quality Framework for EU medicines regulation document (16) or a different data quality framework appropriate to the data source and make the results available to regulators
- To develop expertise for the implementation, analysis, and interpretation of at least one data quality framework

8. Statistical analyses

The statistical analyses should be performed according to a pre-defined statistical analysis plan (written as part of the protocol or as a separate document) and this plan should be developed before the preparation of the analysis dataset. The following aspects deserve additional attention.

8.1. Model specification

The assumptions of the analytic approach (specification of the statistical model, key variables, censoring assumptions, etc.) and the model diagnostics are important considerations to be presented in the statistical analysis plan. These assumptions, and their possible impact on the results, should be addressed through sensitivity analyses and assessed in the final study report.

If Bayesian models are used, graphs of the full posterior distribution of key model parameters as well as relevant summaries of these posteriors should be presented (e.g. mean and/or median, 95% credible intervals and relevant posterior probability statements).

8.2. Estimation and precision

For NIS, rather than focusing on hypothesis testing, use of estimates quantifying the magnitude of the effect and of confidence intervals describing the precision of these estimates, both overall and in important subgroups, is essential to support decision making derived from the data (21).

For NIS based on large RWD sources, statistical analyses may produce statistically significant results and narrow confidence intervals that may not be clinically relevant and may be subject to bias. An integrated evaluation is therefore essential for any conclusion based on results, i.e. to estimate the magnitude of the effect and its clinical relevance, to provide an appropriate description of the precision of estimates with confidence intervals, and to supplement this estimation with an assessment of the impact of selection bias, information bias and confounding.

8.3. Time-dependent analyses

Longitudinal follow-up in cohort studies require accurate accounting of time periods in each exposure category from study entry until end of follow-up.

- A time-dependent analysis should be planned when appropriate to the research question in the cohort studies where events occur at different time points to account for time-dependent exposure and confounding.
- Treatment switches are common in longitudinal pharmacoepidemiological studies. The methods used to handle treatment switches is dependent of the specific outcome studied. Approaches to address the methodological challenges posed by treatment switches should be pre-planned.
- Varying exposure may be related to intercurrent events and strategies to handle them should be aligned to the analytic approaches (see Chapter 4.4.).

8.4. Stratified analyses

Stratified analyses may provide further insights into the results and can fulfil several objectives: to provide results of the analysis in sub-groups of interest as part of the research question, to measure an effect estimate within relevant categories of a confounder, and to perform analyses in sub-groups defined by potential effect modifiers.

- Stratified analyses should be pre-specified and may be planned as supplementary analyses.
- In multi-database studies, stratified analyses may also help assess the robustness of the results across important subgroups/datasets; they should be performed by country and setting (e.g. primary care or hospital setting) in addition, for example, to age, gender and disease severity.
- Additional stratified analyses may be requested by regulatory assessors. If formulated at an early stage of the discussions with MAHs and Applicants, they should be addressed with the research question or, at the latest, in the development of the study protocol or statistical analysis plan.

8.5. Sensitivity and supplementary analyses

Sensitivity analyses should be considered to assess the impact on the results of assumptions made in the primary analysis, biases or limitations in the data. Supplementary analyses may be considered to provide additional contextual information, e.g. to better understand the study data or the impact of choices in the study design.

- Sensitivity and supplementary analyses should be pre-specified in the statistical analysis plan. They should be discussed and aligned with regulators at an early stage if they affect the study design and the choice and definition of study variables.
- In studies with causal objectives, the sensitivity analyses should consider the recommendations of the ICH E9(R1) Addendum on Estimands and Sensitivity Analysis in Clinical Trials. Primary and sensitivity analyses should address the same research question. For example, analyses changing the definition of exposure or outcome may target a different research question and should not be referred to as sensitivity analyses.

8.6. Missing data

Missing data can lead to bias and confounding, and the following is recommended:

- to describe the management of missing data in the study protocol and the statistical analysis plan;
- to provide a thorough justification for the assumptions made regarding missing data and the appropriateness of the method chosen to handle them in the analysis;
- to consider sensitivity analyses to missing data assumptions made in the main analysis to understand their impact on the results;
- to follow the recommendations of the ICH E9(R1) Addendum on Estimands and Sensitivity Analysis in Clinical Trials.

8.7. Heterogeneity

In multi-database studies, different estimates may be found even when the same protocol is applied across all data sources (22). Many factors may explain heterogeneity in addition to random variation: different healthcare systems and reimbursement policies for healthcare, different populations covered by the data source, different practices for data collection, coding systems, and recording. It is recommended:

- to anticipate differences between data sources and study populations in the study protocol in light of the study objectives and to describe how they will be managed;
- to discuss in the statistical analysis plan quantitative management of heterogeneity through appropriate methods for evidence synthesis;
- to consider use of forest plots and other visualisation methods of the results (point estimates and 95% confidence intervals) to help evaluate the results.

9. References

1. ENCePP. ENCePP Guide on Methodological Standards in Pharmacoepidemiology. (Revision 11), 2023.
2. Schneeweiss S, Rassen JA, Brown JS et al. Graphical Depiction of Longitudinal Study Designs in Health Care Databases. Ann Intern Med. 2019;170(6), 398–406.
3. ENCePP. ENCePP Checklist for Study Protocols (Revision 4), 2018.
4. ENCePP. Target trial emulation. ENCePP Guide on Methodological Standards in Pharmacoepidemiology. Chapter 4.4.2. Target trial emulation.
5. Hernán, M. A., & Robins, J. M. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8),758–64.
6. Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC). Chapters 3.6 (The target trial) and 22 (Target trial emulation).
7. Hernan MA, Alonso A, Logan R et al. Observational studies analysed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology 2008;19(6):766-79.
8. Dickerman BA, García-Albéniz X, Logan RW et al. Avoidable flaws in observational analyses: an application to statins and cancer. Nat Med. 2019;25(10):1601-6.
9. Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Curr Epidemiol Rep. 2015;2(4):221-8.
10. Webster-Clark M, Mavros P, Garry EM, et al. Alternative analytic and matching approaches for the prevalent new-user design: A simulation study. Pharmacoepidemiol Drug Saf. 2022;31(7):796-803.
11. Suissa S. Immortal time bias in pharmaco-epidemiology. Am J Epidemiol. 2008;167(4):492-9.
12. ENCePP. The ENCePP Code of Conduct. Rev. 4, 2018.
13. International Society for Pharmacoepidemiology (ISPE). Guidelines for Good Pharmacoepidemiology Practices (GPP). Revision 3. 2015.
14. European Commission. Informed consent guidance.
15. EMA. Good Practice Guide for the use of the HMA-EMA Catalogues of real-world data sources and studies. 2025.
16. HMA-EMA. Data Quality Framework for EU medicines regulation. 2023.
17. EMA. Reflection paper on the use of Artificial Intelligence (AI) in the medicine product life cycle. 2024.
18. Wang S. 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. Pharmacol Drug Saf. 2022.
19. ENCePP. Chapter 9.2.Models of studies using multiple data sources. ENCePP Guide on Methodological Standards in Pharmacoepidemiology (Revision 11), 2023.
20. ENCePP. Chapter 13.2 Data Quality Frameworks. ENCePP Guide on Methodological Research Standards in Pharmacoepidemiology, (Revision 11), 2023.
21. Wasserstein RL, Lazar NA. The ASA Statement on *p*-Values: Context, Process, and Purpose. The American Statistician 2016;70(2):129-33
22. Russek M, Quinten C, de Jong VMT et al. Assessing heterogeneity of electronic health-care databases: A case study of background incidence rates of venous thromboembolism. Pharmacoepidemiol Drug Saf. 2023 ; 32(9):1032-48.