EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

22 October 2021
EMA/426390/2021
Committee for Human Medicinal Products (CHMP)

Guideline on registry-based studies

Draft approved by the Cross-Committee Task Force on Registries	25 May 2020
Draft sent to the EU Regulatory Network for consultation including EMA committees, Patients' and Consumers' Working Party and Healthcare Professionals' Working Party	9 July 2020
Start of public consultation	24 September 2020
End of consultation	31 December 2020
Final guideline agreed by the Cross-Committee Task Force on Registries	7 September 2021
Final guideline adopted by CHMP	16 September 2021

Keywords	**Patient registry, Real World Evidence, Real Word Data, registry-based study, feasibility analysis**

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

© European Medicines Agency, 2021. Reproduction is authorised provided the source is acknowledged.

Table of contents

Abbreviations **3**
Glossary **4**
1. Introduction **5**
2. Scope and objective **5**
3. Methods and processes **6**
3.1. Differences between a registry-based study and a patient registry 6
3.2. Use of registry-based studies for evidence generation 7
3.3. Planning a registry-based study 8
3.4. Study protocol 9
3.5. Study population 10
3.5.1. Choice of study population 10
3.5.2. Informed consent 11
3.5.3. Data protection 11
3.6. Data collection 12
3.7. Data quality management 12
3.8. Data analysis 13
3.9. Data reporting 14
4. Legal obligations and regulatory requirements **15**
Annex: Considerations on patient registries **20**
A.1. Introduction 20
A.2. Registry population 20
A.3. Data elements 21
A.4. Quality management in patient registries 23
A.5. Governance 25
A.6. Data sharing outside the context of registry-based studies 26
References **27**
Appendices **31**
Appendix 1. Checklist for evaluating the suitability of registries for registry-based studies 31
Appendix 2. Safety reporting 33
Appendix 3. Examples of recommended international terminologies for data elements 34

Abbreviations

ADVANCE	Accelerated Development of VAcciNe benefit-risk Collaboration in Europe (a project of the Innovative Medicines Initiative (IMI))
AESI	Adverse Event of Special Interest
AHRQ	US Agency for Healthcare Research and Quality
ATMP	Advanced Therapy Medicinal Product
CHMP	Committee for Medicinal Products for Human Use (at EMA)
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EUnetHTA	European Network for Health Technology Assessment (Joint Action on Health Technology Assessment)
EU PAS Register	European Union Electronic Register of Post-Authorisation Studies
EU RD Platform	European Platform on Rare Diseases
GCP	Good Clinical Practice
GDPR	General Data Protection Regulation
GVP	Good Pharmacovigilance Practices
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ITF	Innovation Task Force (at EMA)
HTA	Health Technology Assessment
MAA	Marketing Authorisation Applicant
MAH	Marketing Authorisation Holder
MedDRA	Medical Dictionary for Regulatory Activities
NCA	National Competent Authority
OMOP	Observational Medical Outcomes Partnership
PAES	Post-Authorisation Efficacy Study
PARENT	Cross border PAtient REgistries INiTiative (Joint Action under the EU's Health Programme 2008-2013)
PASS	Post-Authorisation Safety Study
PRAC	Pharmacovigilance Risk Assessment Committee (at EMA)
PRIME	PRIority MEdicines (at EMA)
QPPV	Qualified Person Responsible for Pharmacovigilance
RCT	Randomised Clinical Trial
REQueST	Registry Evaluation and Quality Standard Tool (developed by EUnetHTA)
RMP	Risk Management Plan
RRCT	Registry-based Randomised Clinical Trial
SAP	Statistical Analysis Plan
SAR	Serious Adverse Reaction
SmPC	Summary of Product Characteristics
SUSAR	Suspected Unexpected Serious Adverse Reaction
US	United States

Glossary

Patient registry (synonym: registry): Organised system that collects uniform data (clinical and other) to identify specified outcomes for a population defined by a particular disease, condition or exposure (1). The term 'patient' highlights the focus of the registry on health information. It is broadly defined and may include patients with a certain disease, pregnant or lactating women or individuals presenting with another condition such as a birth defect or a molecular or genomic feature.

Disease registry*:* Patient registry whose members are defined by a particular disease or disease-related patient characteristic regardless of exposure to any medicinal product, other treatment or particular health service.

Registry-based study: Investigation of a research question using the data collection infrastructure or patient population of one or several patient registries.

A registry-based study is either a clinical trial or a non-interventional study as defined in Article 2 of Regulation (EU) No 536/2014. The table in Annex I of the Questions & Answers Document - Regulation (EU) 536/2014 provides the difference between non-interventional studies and interventional trials.

A registry-based study may apply primary data collection in addition to secondary use of the existing data in the registry.

Registry-based randomised clinical trial: Randomised clinical trial embedded in the data collection infrastructure of one or several patient registries (e.g. randomisation, data collection or follow-up).

Registry database (synonym: register): Database derived from one or several registries.

Primary data collection: Collection of data directly from patients, caregivers, healthcare professionals or other persons involved in patient care.

Secondary use of data: Use of existing data for a different purpose than the one for which it was originally collected.

Harmonised or mapped data elements*:* Data elements that have been harmonised or mapped across data sources to facilitate the implementation of a common data quality system, data exchange, data analysis and/or the interpretation of results from a study.

Competent authority: This term should be understood in its generic meaning of an authority regulating medicinal products and/or an authority appointed at national level for being in charge of all or individual pharmacovigilance processes. For the purpose of this guideline, the term "competent authority" covers the competent authorities in Member States (or National Competent Authorities - NCAs) and the Agency.

1. Introduction

The European Medicines Agency (EMA) Patient Registry Initiative and the EMA Cross-Committee Task Force on Registries (2) have explored ways to improve the use of patient registries to support regulatory decision-making on medicinal products within the European Union (EU). Recommendations on aspects to be addressed for registry-based studies were issued in five workshops on specific patient registries (3) and in the Committee for Medicinal Products for Human Use (CHMP) Qualification Opinions for two networks of registries via the EMA Scientific Advice Working Party (4) (5). The EMA's Cross-Committee Task Force on Registries also published for consultation a discussion paper on methodological and operational aspects of the use of patient registries for regulatory purposes. The information gained in these activities has been integrated in this new Guideline on registry-based studies, which also uses recommendations from the PARENT Joint Action Methodological Guidance (6), the EUnetHTA's Registry Evaluation and Quality Standards Tool (REQueST) (7), the US Agency for Healthcare Research and Quality (AHRQ)'s Users' Guide on registries (1), and the European Reference Network Patient Registries platform (8).

2. Scope and objective

The objective of this Guideline is to provide recommendations on key methodological aspects that are specific to the use of patient registries by marketing authorisation applicants and holders (MAAs/MAHs) planning to conduct registry-based studies. To support these recommendations, considerations and aspects of patient registries that NCAs and EMA view important as good regulatory practice in registry-based studies are included in the Annex. The relevant legal basis and regulatory requirements that apply to these studies are listed in Chapter 4.

Patient registries may have several purposes, such as to monitor the clinical status, quality of life, comorbidities and treatments of patients over time or to monitor and improve overall quality of care. They are a source of data on the presence or occurrence of a particular disease or health-related individual characteristic(s), such as a set of signs or symptoms, or a specific condition, such as pregnancy, breast-feeding, a birth defect or a molecular or genomic feature. They are therefore an important source of data for registry-based studies on healthcare practices, utilisation of medicines and medical devices, and outcomes of treatments. They may, in particular, represent an important source of data on rare diseases and patients treated with advanced therapy medicinal products (ATMP) (9), including gene therapy (10).

In some countries, datasets created by a comprehensive registration of administrative and healthcare data of the population at the regional or national level are called registries or registers. Such registries or registers collect healthcare data at population level and therefore many recommendations included in this Guideline are less relevant, for example on possible selection bias or concerns about generalisability of study results using such registries.

The term *product registry* is sometimes used to indicate a system of data collection by MAAs/MAHs targeting patients exposed to a specific medicinal product or substance. From a regulatory perspective, recruitment and follow-up of these patients with the aim to evaluate the use, safety, effectiveness or another outcome of this exposure typically falls outside of normal routine follow-up of patients and therefore corresponds to a clinical trial or non-interventional study in the targeted population. It is therefore preferable to avoid using the term "product registry" in this situation and directly refer to the appropriate terminology instead (clinical trial or non-interventional study).

Details on procedural aspects related to the interactions with NCAs and EMA on registry-based study protocols and results are not within the scope of this Guideline. These can be found in the relevant

guidance documents published on the EMA website, and references are included throughout this document.

Although this Guideline is primarily targeted to MAAs/MAHs and others who wish to undertake registry-based studies with a possible regulatory purpose, it is also relevant to registry holders, patients and other persons involved in the funding, creation and management of patient registries, and those participating in the collection and analysis of registry data.

Legal requirements are identifiable by the modal verb "shall". Recommendations that are not legal requirements are provided using the modal verb "should".

3. Methods and processes

3.1. Differences between a registry-based study and a patient registry

Important methodological differences between a registry-based study and a registry are summarised in the Table below. The principles outlined in the Table are further explained in Chapters 3.3 to 3.9 for the registry-based studies and in the Annex for the patient registries.

Topics	Registry-based study	Patient registry
1. Definition	Investigation of a research question using the data collection infrastructure or patient population of one or more patient registries.	Organised system that collects uniform data (clinical and other) to identify specified outcomes for a population defined by a particular disease, condition or exposure
2. Duration of follow-up	Timelines driven by the study objectives, the collection/extraction and analysis of the relevant study data.	Timelines driven by schedules for data collection and any anticipated data analyses which prompted the registry.
3. Patient enrolment	Defined by research objective(s) and may include a subset of a registry population; in case of a clinical trial, allocation to treatment arm (e.g. with randomisation) is to be documented; generalisability of the study results to be documented.	Aimed at enrolment of all patients with the particular disease or condition; generalisability of registry data to be documented.
4. Data collection	Restricted to what is needed by the research question including data on potential confounders and effect modifiers; collection of additional data not routinely collected in the registry may be required; if such additional data includes subject monitoring outside the terms of the SmPC and normal clinical practice, the legislation for clinical trials may apply; study may involve primary data collection in addition to secondary use of data.	Data collected based on the purpose of the registry; agreed core set of data elements to be collected with documented definitions, coding system and data entry procedures; data collected for the purpose of a registry can involve primary collection of data or secondary use of data.
5. Analysis plan	Detailed statistical considerations most commonly defined in separate document in addition to study protocol and to registry protocol; descriptive or hypothesis driven statistical analysis plan.	Statistical analysis plan with analyses often performed routinely at intervals based on patient accrual or analyses of pre-defined outcomes at time points described in the registry protocol.
6. Data quality management	Study-specific data quality management to be prospectively defined and implemented with a risk-based approach.	Quality management applied routinely to data and processes with a focus on core set of data elements; data systems to ensure data integrity, completeness and security;

		data quality management to be prospectively defined and documented.

3.2. Use of registry-based studies for evidence generation

The acceptability of registry-based studies as a source of evidence for regulatory purposes depends on several factors related to the specific regulatory assessment procedure for the concerned medicinal product, the characteristics of the concerned registry (see Annex) and the objectives, design and analytical plan of the proposed study (11). Early consultation with national competent authorities (NCAs), where applicable, and with EMA (e.g. the procedure for Scientific Advice and Protocol Assistance) is recommended when a registry-based study is proposed to be used and study protocols should be published (12). Examples where registry-based studies have been used for evidence generation are presented below:

- To complement the evidence generated in the pre-authorisation phase

  Pre-clinical studies and clinical trials (CT) are at the core of the scientific evaluation of the efficacy and safety of medicines prior to granting a marketing authorisation. In some circumstances, pre-authorisation evaluation may be supported by observational evidence derived from patient registries. Examples of such evidence may include information on standards or real-world practice of care for the disease, incidence, prevalence and determinants of disease outcomes in clinical practice, or the characteristics of the registry population.

  Studies based on patient registries may also contextualise the results of uncontrolled trials, and patient registries have been used to support registry-based randomised controlled trials (RRCTs) for patient recruitment (e.g., to identify patients meeting inclusion/exclusion criteria), randomisation allocation, sample size calculation, endpoints identification, data collection and study follow-up (13) (14) (15) (16). Open questions remain regarding the validity and relevance of RRCTs (13), (17). It is therefore recommended to obtain Scientific Advice from EMA and, where applicable, from the concerned NCAs, health technology assessment (HTA) bodies and health insurance schemes as payers on the acceptability of the chosen approach for evidence generation in case deviations from a traditional randomised clinical trial (RCT) design are considered.

- To provide evidence in the post-authorisation phase

  Patient registries can be the basis for recruitment and randomisation for RCTs and non-interventional studies, post-authorisation efficacy studies (PAES) (18) and post-authorisation safety studies (PASS) (19) performed after marketing authorisation. They may allow linkage of patient records with other data sources such as biobank data, census data, or demographic data. In the context of medicinal products with efficacy previously demonstrated in RCTs, registry-based studies may help, for example, to assess the effectiveness of adapted dosing schemes applied in clinical practice and understand effectiveness and safety of products in a broader clinical disease-related context and a more heterogenous patient population. With adequate sample size and appropriate study design, registry-based PASS can provide data to quantify and characterise risks, to identify risk factors for the occurrence of adverse reactions, to evaluate the safety profile of a medicinal product in long-term use, to assess patterns of medicines utilisation that add to knowledge on the benefit-risk profile of the medicinal product, or to measure the effectiveness of a risk minimisation measure (e.g. by estimating its public health impact) (20).

  Products intended for rare diseases are often studied in uncontrolled trials and the size of the safety and efficacy datasets at time of marketing authorisation application is small. In these cases, follow-up for efficacy and safety may be needed, and PAES and PASS are often imposed for post-

authorisation evidence generation. These are frequently and preferentially performed on the basis of existing patient registries.

- To evaluate the effects of medicinal products used during pregnancy and breast feeding

  Pregnancy registries include pregnant women exposed or not to different treatments and followed up to collect information on outcomes of pregnancy and in the offspring for a given medicinal product. Despite the challenges of such studies related to the completeness of information on pregnancy outcomes, the ascertainment of the exposure window/trimester, teratology information services or electronic healthcare records where mother-child linkage is possible, pregnancy registries may also provide valuable data on the benefit-risk balance of medicinal products in breastfeeding (21).

3.3. Planning a registry-based study

The first step is to identify the scientific question(s) to be addressed by the study and critically consider if a registry-based study is appropriate to provide the desired answers.

Planning a registry-based study may require to identify one or several suitable registry(-ies), to obtain from each registry and each individual centre (if no central registry coordination exists) an agreement to collaborate, to set up a database, a data extraction process and quality control activities, and to identify whether a new informed consent is needed for using data from the registry (see Chapter 3.4). As appropriate, data stewardship expertise should be considered for drafting and executing a research data management plan addressing the data source, the nature of expected study results, and how the data will be findable, accessible, interoperable, and reusable (22).

MAAs/MAHs should therefore discuss early with NCAs and EMA (including the concerned Rapporteurs or Lead Member States and concerned EMA Committees) through e.g. scientific advice and protocol assistance procedures, as well as with registry holders, HTA bodies and health insurance schemes/payers if relevant the feasibility of using registries to meet regulatory needs. It is the responsibility of the MAAs/MAHs to include the registry holders and any other stakeholders in the discussion. The EMA Innovation Task Force (ITF) (23), the EMA PRIority MEdicines (PRIME) procedure (24), if applicable, and pre-submission meetings can also be used in the pre-authorisation phase. A strategy for post-authorisation activities should be developed in the pre-authorisation phase and discussed in Scientific Advice and PRIME procedures if applicable.

MAAs/MAHs proposing a registry-based study should provide adequate information regarding the availability of data, the quality management procedures applied and the need and feasibility of introducing any study-specific additional data collection and quality control measures. In case of study-specific primary data collection, adequate measures may be needed to detect averse events and promptly report suspected adverse reactions in accordance with the study protocol. A feasibility analysis should therefore be considered by the MAA/MAH or research organisation initiating the study prior to writing the study protocol, to guide its development and facilitate the discussion with NCAs, EMA, HTA bodies and other parties. The feasibility analysis should be performed in collaboration with registry holders and include the following information, as applicable:

- General description of the registry(-ies) or network of registries; the Checklist for evaluating the suitability of registries for registry-based studies (see Appendix 1) can be used to prepare this description; the epidemiology of the disease, this is more precise, medicines use and standards of care applied in the country or registry setting should be described if relevant for the specific study;
- Analysis of the availability in the registry of the core data elements needed for the planned study period (as availability of data elements may vary over time), including relevant confounding and

effect-modifying variables, whether they are mapped to any standard terminologies (e.g. MedDRA, OMOP common data model), the frequency of their recording and the capacity to collect any additional data elements or introduce additional data collection methods if necessary (see Chapter 3.6);

- Analysis of the quality, completeness and timeliness of the available data elements needed for the study, including information on missing data and possible data imputations, risk of duplicate data for the same patient, results of any verification or validation performed (e.g. through an audit), analysis of the differences between several registries available in the network and their possible impact on data integration, description of the methods applied for data linkage as applicable, and possible interoperability measures that can be adopted (see Chapter A.3);
- Description of processes in place for the identification of adverse events and prompt reporting of suspected adverse reactions occurring in the course of treatments, and capacity to introduce additional processes for their collection and reporting if needed (see Chapter A.3);
- Study size estimation and analysis of the time needed to complete patient recruitment for the clinical study by providing available data on the number of centres involved in the registry(-ies), numbers of registered patients and active patients, number of new patients enrolled per month/year, number of patients exposed to the medicinal product(s) of interest, duration of follow-up, missing data and losses to follow-up, need and possibility to obtain informed consent (see Chapter 3.4);
- Evaluation of any potential information bias, selection bias due to the inclusion/exclusion criteria of centres (e.g. primary, secondary or tertiary care) and patients, potential time-related bias between and within registry(-ies), and potential bias due to loss to follow-up (see Chapter 3.8);
- Evaluation of any potential confounding that may arise, especially if some data elements cannot be collected or measured (see Chapter 3.8);
- Analytical issues that may arise based on the data characteristics and the study design (see Chapter 3.8);
- Any data privacy issues, possible limitations in relation to informed consent and governance-related issues such as data access, data sharing and funding source (see Chapter A.5);
- Overall evaluation of the suitability of the registry for the specific study, taking into account any missing information on the above-mentioned aspects.

The final report of the feasibility analysis may be submitted either separately or as part of the proposed protocol for a registry-based study. In order to inform the feasibility of other studies in the same registry and reduce duplication of work, the feasibility analysis should be published with the study protocol in the EU PAS Register in agreement with the registry holder. Any confidential information may be redacted if needed (25).

For regulatory studies addressing a class of products where several MAHs have the same obligation to perform a study, MAHs are encouraged to design a joint registry-based study or to join an already existing study on the same topic. When the study design falls into the clinical trial category, a clinical trial could be performed through joint trial sponsorship as provided for in Article 72 of Regulation (EU) No 536/2014.

3.4. Study protocol

The study protocol should describe how the registry infrastructure and population will be used to address the research question of interest, how the study will be conducted and how the validity (both

internal and external) of the results will be ensured. To verify whether a registry-based study is a clinical trial or a non-interventional study, reference should be made to Regulation (EU) 536/2014 and Annex I to the Questions & Answers Document - Regulation (EU) 536/2014 published by the European Commission.

The structure and content of the study protocol should follow the existing regulatory requirements. Clinical trials should follow the ICH E6 (26), ICH E8 (27) and ICH E9 (28) guidelines and the ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials (29). Protocols for non-interventional studies should follow the guidance on the format and content of the protocol for PASS (19)(30) or the Scientific Guidance on PAES (18). They should apply the best methodological standards, including if applicable those described by the ENCePP Guide on Methodological Standards in Pharmacoepidemiology (31). The ENCePP Checklist for Study Protocols (32) identifies important points to be addressed when designing a non-interventional study and writing the study protocol.

Where the registry-based study entails secondary use of data, the study protocol should specify the events of interest that are already collected in the registry and discuss the risks of bias and unmeasured confounding. Dedicated and complete search strategies, coding lists or adjudication should be used to accurately define the outcomes of interest.

The protocol should specify agreements made with the registry holder on the additional variables that can be collected, with timelines for data availability.

The protocol should provide an estimation of the study size needed to answer the research question. The feasibility of attaining this study size within the registry should be assessed using realistic assumptions, both in terms of number of patients (based on inclusion and exclusion criteria) and in terms of duration of follow-up. This should include considerations regarding the estimand and intercurrent events as well as missing data, the need for imputation, and consequent considerations on effect and sample size [ICH E9 (R1)] (29). Where there are doubts about the feasibility of achieving the required study size, possible extension of the study population by recruiting from (an)other registry(-ies) could be considered, weighing the strengths and limitations of using a single registry versus combining datasets of patients with the same disease across multiple registries. Formal powering of the study for the sample size may not be necessary in descriptive studies, but considerations should be made whether a foreseen sample size allows a sufficient precision to inform the study objective and this should be described and justified in the protocol.

If a registry-based study is to be conducted across multiple registries, a common study protocol should be developed based on core data elements available in the registry (see Chapter A.3) and a common design, even if some aspects of the study may vary according to the characteristics of each registry and not all outcomes may be assessed in all registries. Nevertheless, the protocol should also describe differences between registries, assess the resulting heterogeneity of data and critically discuss its potential impact on study results. The protocol or statistical analysis plan (SAP) should propose sensitivity analyses addressing this heterogeneity.

Where several registries are suitable for a study but not all of them are intended to be involved, the study protocol should provide the justification of the choice, i.e. inclusion and exclusion criteria, and discuss the potential impact of selection and interpretability of datasets and findings.

3.5. Study population

3.5.1. Choice of study population

The registry population serves as the source population for the registry-based study. The choice of the study population should be driven by the study objectives and may represent the totality of the

registry population or only a subset with pre-defined characteristics. For example, when studying a medicine of interest, the potential study population may include various groups of patients: newly diagnosed patients entering the registry and receiving a first prescription of the medicine of interest, and registry patients already diagnosed with the disease and who are switched from another treatment, receive the medicine of interest as add-on therapy or have received the medicine of interest only in the past. In such situations, it is useful to collect the data needed to describe all patients receiving the medicine of interest and assess the heterogeneity between subsets of these patients.

In case of study-specific primary data collection within an existing registry, it is critical that procedures are in place to support complete data collection on all eligible patients enrolled in the registry.

3.5.2. Informed consent

Informed consent such as required by Regulation (EU) No 536/2014 serves as an ethical standard and procedural obligation. It provides the fundamental condition under which a person can be included into a study. It is not conceived as a legal basis but should be seen as a safeguard for data processing compliance. Therefore, it is important to distinguish between the requirement for consent for a subject to participate in a study and the requirements for a lawful processing of personal data under the GDPR (33).

In the context of a registry-based study, the ethical and procedural obligation requires to obtain informed consent from patients to participate in the study in addition to the consent already given for participating in the registry, as applicable. It 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 about what will happen to the results of the study. (34) (35)

3.5.3. Data protection

The conduct of registry-based studies needs to respect the following applicable Union data protection rules at each step of the processing of personal data, including the option for data sharing/pooling between registries and other stakeholders like competent authorities and MAAs/MAHs:

- the General Data Protection Regulation (EU) 2016/679 (GDPR), which applies to processing carried out by organisations and bodies operating within the EU, and
- Regulation (EU) 2018/1725 (EUDPR), which applies to Union institutions, bodies, offices and agencies.

Personal data is information that relates to an identified or identifiable individual. An identifiable individual is one who can be identified, directly or indirectly. For this, all the information being processed should be taken into account together with all the means reasonably likely to be used to identify that individual.

When conducting registry-based studies, the legal basis of the personal data processing needs to be established. Specific considerations may be required in case of processing of special categories of personal data such as sensitive (health) information. It should be noted that Member States are allowed to maintain or introduce further conditions, including limitations with regard to the processing of genetic data, biometric data or data concerning health.

According to the principle of accountability, it is the obligation of the data controller (e.g. a registry holder, MAA/MAH, investigator) to implement appropriate technical and organisational measures to ensure and be able to demonstrate that the personal data are processed in accordance with data protection requirements.

The data protection authorities of the Member States are competent for monitoring and enforcing the application of GDPR and are natural interlocutors and first point of contact for the public, businesses and public administrations for questions regarding the GDPR.

3.6. Data collection

A registry-based study may be based entirely on data already collected in the registry (secondary use of data) or may also require additional study-specific primary data collection. For the secondary use of data, only the data needed to ensure the validity and usefulness of the results should be extracted from the registry database, for example, data on exposures, outcomes, confounding and effect modifying variables and variables describing the patient population or the setting from which the data were collected.

Additional study-specific primary data collection may add complexity to the registry-based study. The data collection method applied should clearly be described in the study protocol as it has implications with regards to potential sources of bias and confounding, adequate retrieval of missing data and safety reporting requirements. If the study is classified as a clinical trial, the study-specific data collection must be handled according to ICH E6 (26). Additional study-specific data collection may also affect the ongoing registries' data collection and maintenance and require audit and validation.

3.7. Data quality management

Data quality management for a registry-based study depends on various factors, including the planned use of the study results and whether the study makes use of primary data collection or secondary use of registry data. While data quality management of the registry, as described in Chapter A.4, is the responsibility of the registry holder, it is the MAA/MAH's responsibility to manage the data quality of the registry-based study and interpret the results based on findings on data quality. Specific details on level of data verification and actions to be taken if there are relevant findings, including possible internal or external audits, should be described in a specific data management plan. This plan should be discussed and agreed upon by the MAA/MAH and the registry holder.

Methods and specific measures should be guided by the feasibility analysis and be selected with a view to minimise risk of invalid study results:

- The validity of any data cleaning, extraction and transformation processes should be verified and monitored. This may be specifically relevant in studies using a network of registries where the transformation is performed locally. A risk-based approach requires the identification of data that are critical for data protection and the reliability of the study results.
- Quality checks of the data used in the study should be performed to alert on erroneous, missing or out-of-range values and logical inconsistencies, and trigger prompt data verification and remedial measures if needed.
- In studies with primary data collection, the various factors (e.g. limited human or material resources or inadequate training) influencing quality should be identified and addressed to preserve the integrity of the study. Possible measures include random source data verification, on-site review of processes and computerised systems used for data collection and management. The collected information per time interval for the main outcome parameters can be compared to the amount expected.

The European Commission's risk-proportionate approaches in clinical trials (36), the EMA Reflection Paper on risk-based quality management in clinical trials (37), the GVP Module III on pharmacovigilance inspections (38) and national regulations should be consulted on these aspects.

3.8. Data analysis

The analytical approach to the outcomes of interest should be pre-specified in the registry-based study protocol and the SAP as applicable. Changes to the pre-specified statistical analysis should be reflected by an amendment to the study protocol and/or by an amendment to the SAP. All changes should be presented in the study report.

The ICH E9 (R1) addendum (29) should be considered when planning data analysis by aligning the estimand(s) of interest with (an) adequate estimation and testing method(s). Sensitivity analyses should explore the robustness of estimates on the primary estimand of interest to deviations from underlying assumptions and limitations in the data. For non-interventional studies, the ENCePP Guide on Methodological Standards in Pharmacoepidemiology (31) presents methods to address bias and adjust for confounding.

Depending on the objectives of the registry-based study, the data analysis may need to include an evaluation of the representativeness of the study population in relation to the source population, as it may influence the external validity of the registry-based study (see also Chapter A.2). In case of primary data collection, a comparison of available data between eligible registry patients who are recruited, who decline recruitment or who withdraw from the study and between patients randomised and not randomised in the study, should be performed. If possible, this should be supplemented by a comparison of the study population with a similar population identified from scientific literature data, available electronic healthcare databases, other registries deemed suitable for the study but not used for data collection as justified in the study protocol, or other population-based data sources.

Missing data may lead to bias and confounding, and their handling should be carefully described in the study protocol and the SAP. A thorough justification should be provided for the assumptions about their distribution, causes and timing. The ICH E9 (R1) addendum (29), the EMA Guideline on Missing Data in Confirmatory Clinical Trials (39) and the ENCePP Guide on Methodological Standards in Pharmacoepidemiology (31) provide guidance on how to handle missing data.

In the absence of randomised treatment allocation in registry-based non-interventional studies, some common analytical issues should be addressed:

- The characteristics of patient groups prescribed different treatments are likely to differ. Treatment decisions may be influenced by various factors that may also be associated with the risk of occurrence of the outcome of interest, such as disease severity or the monitoring practice of patients. While methods for addressing this underlying problem have been proposed, these do not provide a unique solution and several sensitivity analyses using different approaches should be performed. In addition, ascertainment of marginal treatment effects over time and factors underlying treatment trajectories may require complete collection of information over the course of the study.
- Registries and registry-based studies may involve different time points for patient inclusion and follow-up, initiation of treatments of interest and ascertainment of events and other variables. The probability of occurrence of events of interest may also be time-dependent. These time points are important to consider as they affect the comparability between treatment groups. Graphical representation of the analysis plan should be used to help understand the various time components of the study and the registry (40). When investigating a treatment effect, immortal time bias can occur when the follow-up period for the study starts before initiation of the treatment under study and the period between start of follow-up and start of treatment is misclassified as exposed.
- Selection bias, information bias and time-related bias may also occur in comparisons to historical control groups. The clinical context may have changed with regard to e.g. treatment options,

diagnosis, medical practice in choice of treatments according to severity of disease, patient care, secular trends in the occurrence of important events, completeness of data collection or other uncollected or unknown factors. These sources of bias should be identified and the impact on the validity of the results assessed.

- A comparative non-exposed control group may be selected from outside the registry, for example from another registry or electronic healthcare records in a country/region where the medicine has not yet been marketed. In this situation, one should ensure that underlying differences between the two populations influencing the risk of outcome occurrence are adequately measured and accounted for in the analysis. Since it may not be possible to identify all underlying differences between populations and completeness of data collection may differ, such comparisons need to be interpreted cautiously.
- Registries offer the opportunity to compare patients prescribed a treatment of interest with patients who are untreated or who have received a different treatment(s) over a long period of time. Inclusion of prevalent medicine users (i.e. patients already treated for some time before study follow-up begins) can introduce two types of bias. Firstly, prevalent medicine users are "survivors" of the early period of treatment, which can introduce substantial (selection) bias if the risk for adverse reactions varies with time (e.g. if treatments carry a risk of hypersensitivity reactions or affect cardiovascular risk). Secondly, covariates influencing medicine prescription at study entry (e.g. disease severity) may be affected by previous medicine use, or patients may differ regarding health-related behaviours (e.g. healthy user effect). A new user design reduces these biases by restricting the analysis to incident medicine users, i.e., patients who enter the study cohort only at the start of the first course of the treatment(s) of interest during the study period. The disadvantages of a new-user design may be a lower sample size and a lower number of patients with long-term exposure, which may then require to extend the duration of the study.
- In the context of the new user design, use of an active comparator may reduce confounding by indication or disease severity as a comparison is made between patients with the same indication initiating different treatments. With newly marketed medicines, however, an active comparator with ideal comparability of patients' characteristics may often be unavailable because newly marketed medicines are often strictly prescribed according to patients' prognostic characteristics and reimbursement considerations, which leads to channelling bias.

3.9. Data reporting

National and EU obligations and reporting requirements for clinical trials and non-interventional studies should be followed (see Chapter 4).

The methods used in the study should be published with sufficient details, while protecting patient privacy, to allow for replication using the same registry database or using a database derived from another registry collecting similar data. Relevant guidelines on reporting of results from clinical trials and non-interventional studies are presented in the EMA Policy on publication of clinical data for medicinal products for human use (41), the Good Pharmacovigilance Practices Module VIII (19) and the ENCePP Guide on Methodological Standards in Pharmacoepidemiology (31). Post-authorisation registry-based non-interventional studies should be registered in the EU PAS Register (25) with the study protocol, the SAP if applicable and the final study report. The final report must contain all study results derived from the analyses prespecified in the study protocol and SAP, whether favourable or unfavourable. The analytical code as well as any prior feasibility analyses (see Chapter 3.3) are ideally also made available. A summary in lay language of the main results and conclusions of the final study report should be prepared and distributed to the registry participants in collaboration with the registry holder.

For non-interventional studies, the principles of scientific independence and transparency for reporting study results described in the ENCePP Code of Conduct (42) and the ADVANCE Code of Conduct for vaccines (43) should be followed. The responsibility for preparing the final study report lies at the appropriate level of study governance, e.g. medical/scientific advisory board, principal investigator and local registry investigators in studies based on multiple registries. For studies funded by a MAA/MAH and requested by a regulatory authority, all parties involved should be responsible for ensuring that the study meets the regulatory requirements of the competent authority and the MAA/MAH should be able to comment on the study results and their interpretation as well as on the format of the report. Requests by the MAA/MAH that interpretation of the results or their presentation be changed should be based on sound scientific reasons or documented regulatory requirements.

Following the submission of the final study report, the competent authority may request additional information and clarifications from the MAA/MAH or may initiate an inspection. Therefore, if a research contract is signed between the MAA/MAH and the registry holder, the contract should include a requirement for the registry holder to address the scientific aspects of the request, with the possibility for the MAA/MAH to provide comments, as well as a requirement to allow a possible regulatory inspection of the registry-based study.

4. Legal obligations and regulatory requirements

The following table summarises the legal basis and regulatory requirements applicable to MAAs/MAHs for different activities related to registry-based studies.

Activities	Legal obligations and regulatory requirements
All activities related to the planning, data collection, data management, data analysis, and data reporting	All activities should be clearly set out in the study protocol and agreed with all involved parties including registry holders, NCAs and EMA where applicable. To verify whether a registry-based study is a clinical trial or a non-interventional study, reference should be made to table in Annex I of the Questions & Answers Document - Regulation (EU) 536/2014 published by the European Commission.

	REFERENCES
	For a clinical trial:
	• Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC ICH E6 - Good clinical practice latest revision
	• ICH E8 - General considerations for clinical studies latest revision
	• ICH E9 - Statistical principles for clinical trials
	• ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials
	For a non-interventional study:
	• Directive 2001/83/EC (the Medicinal Products Directive), Regulation (EC) No 726/2004 (as amended), Commission Implementing Regulation (EU) No 520/2012 on the Performance of Pharmacovigilance Activities Provided for in Regulation (EC) No 726/2004 and Directive 2001/83/EC
	• EMA Scientific Guidance on Post-Authorisation Efficacy Studies and the Guidance in Post-Authorisation Efficacy Studies: Questions and Answers on PAES
	• GVP Module VIII - Post-authorisation safety studies
	• Others to consider:
	  ○ GVP Module III - Pharmacovigilance inspections
	  ○ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the

Activities	Legal obligations and regulatory requirements
	processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation (GDPR)) ○ Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 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 ○ The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology ○ The ENCePP Checklist for Study Protocols ○ The ENCePP Code of Conduct ○ The ADVANCE Code of Conduct for collaborative vaccines studies ○ The Registry Evaluation and Quality Standards Tool (REQueST) • Relevant national legislation and requirements in the country where the study is conducted
Scientific Advice procedures	The MAA/MAH, an organisation subcontracted by the MAA/MAH or an organisation acting independently from any MAA/MAH may ask the EMA or NCAs as applicable for Scientific Advice on the most suitable methods and study designs to generate evidence for the development or maintenance of a medicine. Consultation in parallel with another regulatory authority(ies) or HTA body(ies) is facilitated through EMA procedures.
	REFERENCES • EMA Guidance for Applicants seeking scientific advice and protocol assistance • National Scientific Advice procedures
Safety monitoring and reporting of adverse events and suspected adverse reactions	For non-interventional registry-based studies initiated, managed or financed by a MAA/MAH, collection, processing and reporting of adverse events and suspected adverse reactions should be described in the study protocol. Appendix 2 provides a decision tree concerning MAA/MAH responsibilities where a registry-based study fulfils the definition of a non-interventional study.
	REFERENCES *For a clinical trial:* • Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC ICH E6 - Good Clinical Practice • Communication from the Commission - Detailed Guidance on the Collection, Verification and Presentation of Adverse Event/Reaction Reports Arising from Clinical Trials on Medicinal Products for Human Use ('CT-3') • ICH E2F Development Safety Update Report (DSUR) (safety information to be periodically summarised and assessed on aggregate level in DSURs) • GVP Module VI - Collection, management and submission of reports of suspected adverse reactions to medicinal products *For a non-interventional study:* • Directive 2001/83/EC (the Medicinal Products Directive) and Regulation (EC) No 726/2004 (as amended) • Scientific Guidance on Post-Authorisation Efficacy Studies and the Guidance on Post-Authorisation Efficacy Studies: Questions and Answers on PAES

Activities	Legal obligations and regulatory requirements
	• GVP Module VI - Collection, management and submission of reports of suspected adverse reactions to medicinal products (sections referring to the management of solicited reports or spontaneous reports as applicable, to the management of adverse events in studies with a design based on primary data collection or based on secondary use of data) • GVP VII - Periodic safety update report (safety information to be summarised in PSURs and other periodic and regulatory reports) • GVP Module VIII - Post-authorisation safety studies (summary of all collected adverse events/adverse reactions in the interim and final study report as applicable) • Other GVP Modules: V - Risk management system, IX - Signal management (including section on emerging safety issues (ESIs))
Transparency, study registration	The registration requirements depend on the type of registry-based studies: • Clinical trials shall be entered in the European database (EudraCT or CTIS when the Clinical Trials Regulation becomes applicable on 31 January 2022) with their protocol upon receipt of the Clinical Trial Application (CTA). Results shall be published in the European database by the sponsor. • For non-interventional PASS: Imposed studies initiated, managed or financed by an MAH shall be registered by the MAH in the EU PAS Register. Non-imposed studies required in the RMP or conducted voluntarily in the EU should also be registered in the EU PAS Register. Registration should include the study protocol and the study report. • For non-interventional PAES: Studies initiated, managed or financed by an MAH should be registered in the EU PAS Register, independently from whether they are imposed or not. • All non-interventional PASS/PAES initiated, managed or financed by other parties than an MAH should also be registered in the EU PAS Register together with their protocols and studies results when available. Making this information available will help increase transparency, reduce publication bias and support collaborations between centres and any other parties.
	REFERENCES *For a clinical trial:* • Eudralex Vol 10 – Clinical Trials Guidelines: Chapter I Application and application form • Commission Guideline - Guidance on posting and publication of result-related information on clinical trials in relation to the implementation of Article 57(2) of Regulation (EC) No 726/2004 and Article 41(2) of Regulation (EC) No 1901/2006 • EudraCT website, Clinical Trials Information System (CTIS) • Commission Implementation Regulation (EU) 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 (Annex III - Protocols, abstracts and final study reports for post-authorisation safety studies) *For a non-interventional study:* • GVP Module VIII - Post-authorisation safety studies • ENCePP Guide on Methodological Standards in Pharmacoepidemiology / EU PAS Register website • EMA Scientific Guidance on Post-Authorisation Efficacy Studies; Guidance on Post-Authorisation Efficacy Studies: Questions and Answers • National requirements

Activities	Legal obligations and regulatory requirements
Record keeping	Retention times depend on the applicable legislation at the time of approval of the clinical study and on the type of clinical study.
	REFERENCES *For a clinical trial:* • Commission Directive 2005/28/EC of 8 April 2005 laying down principles and detailed guidelines for good clinical practice as regards investigational medicinal products for human use, as well as the requirements for authorisation of the manufacturing or importation of such products • Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC EMA Guideline on the content, management and archiving of the clinical trial master file (paper and/or electronic) • National legislation *For a non-interventional study:* • Commission Implementation Regulation (EU) 520/2012 • GVP Modules I – Pharmacovigilance systems and their quality systems • GVP Module VIII - Post-authorisation safety studies • EMA Scientific Guidance on Post-Authorisation Efficacy Studies; Guidance on Post-Authorisation Efficacy Studies: Questions and Answers
Personal data protection	The conduct of registry-based studies needs to respect applicable Union data protection rules
	REFERENCES • Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) OJ L 119, 4.5.2016, p. 1–88 • For more details regarding the territorial scope of the GDPR, please see EDPB Guidelines 3/2018 on the territorial scope of the GDPR (Article 3), 12 November 2019 • Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 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, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC (Text with EEA relevance) PE/31/2018/REV/1 OJ L 295, 21.11.2018, p. 39–98 • Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC • European Commission's Question and Answers on the interplay between the Clinical Trials Regulation and the General Data Protection Regulation • EDPS Preliminary Opinion on data protection and scientific research, 6 January 2020 • EDPB Guidelines 05/2020 on consent under Regulation 2016/679, 4 May 2020 • National data protection authorities of the competent Member States for the monitoring of and enforcing the application of the GDPR and other national data protection legislations that may be applicable in their territories (44)

Activities	Legal obligations and regulatory requirements
Relationship between financing body and subcontractors for registry-based study	Where the MAA/MAH finances and subcontracts a registry-based study imposed by a competent authority to another organisation, it remains responsible to the competent authorities for all legal obligations. Pharmacovigilance responsibilities and obligations apply to MAA/MAH also for voluntary registry-based studies. The contractual arrangement between the MAA/MAH and the other organisation should be detailed, up-to-date and clearly describe the responsibilities of each party. Where the MAA/MAH has subcontracted some of its pharmacovigilance tasks, it shall retain responsibility for ensuring that an effective quality system is applied in relation to those tasks. The other organisation can be another MAH, as different MAHs in the EU can collaborate in initiating, managing and financing a registry-based studies. This must likewise be subject to contractual arrangements.
	REFERENCES • Commission Implementing Regulation (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 • GVP Modules I - Pharmacovigilance systems and their quality systems • GVP Module II - Pharmacovigilance system master file
Information on responsible QPPV	The MAH shall ensure that its qualified person responsible for pharmacovigilance in the EU (QPPV) has sufficient authority to influence the performance of the quality system and the pharmacovigilance activities of the MAH which may include registry-based studies.
	REFERENCES • Commission Implementing Regulation (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 • GVP Module I - Pharmacovigilance systems and their quality systems

Annex: Considerations on patient registries

A.1. Introduction

This Annex proposes aspects of good practice in the establishment and management of patient registries considered relevant to their use for registry-based studies and other possible regulatory purposes.

A.2. Registry population

Use of a patient registry for regulatory purposes generally requires that the data generated from the registry entirely cover or are representative of the population with the particular disease, condition or exposure. Selection bias can affect the validity of the data derived from the registry and can occur at the level of the selection of centres recruiting patients for the registry (i.e. if centres with a non-representative population are preferably included), patient enrolment (i.e. if not all patients are enrolled or patients enrolled are not representative of the patient population) and patient loss to follow-up. These selection biases may be influenced by many factors, including clinical, demographic and socio-economic factors.

The following steps should be considered prior to the enrolment of a registry population:

1. To clearly define the purpose of the registry and the corresponding population with the particular disease, condition or exposure.
2. To translate the population definition into a detailed description of when, where and how patients will be enrolled in practice, for example all patients diagnosed with a certain disease by all hospital specialists managing that disease. It may include exclusion criteria, whose rationale should be justified and documented.
3. To establish processes allowing for enrolment of all eligible patients fulfilling the population definition. This critical step should include prospective enrolment of all newly consenting eligible patients fulfilling the definition and enrolment of eligible existing patients (having already consented for their data to be used or providing a new consent) by methods ensuring representativeness and minimising selection bias, for example by using a pre-existing listing of patients. This step can be facilitated by supporting patient engagement e.g. through patient organisations and the provision of information about the registry to patients prior to enrolment. Completeness of recruitment into registries should be monitored and reported as part of the registry project.
4. To create a system that minimises loss to follow-up and maximises the completeness and accuracy of key information collected on each enrolled patient, including variables representing potential confounders and effect modifiers in future registry-based studies. Completeness of follow-up should be monitored and reported, and deviations from expectations explained. When using the data for regulatory purpose, sensitivity analyses on the effects of incomplete follow-up may be needed.

The level of enrolment and follow-up of patients may depend on the specific disease. Children and other populations (e.g. affected by rare diseases or presenting co-morbidities) may present specific challenges.

Anticipation of incomplete enrolment may require specific solutions to support the registry enrolment strategy and assess the representativeness of the registry population, such as e.g.:

- where possible, comparison of the actual registry population and relevant data elements with another data source covering the same population (e.g. electronic health care records);
- collection of minimum information (where locally allowed) at baseline on patients asked to join the registry but not included in order to compare their characteristics with those of included patients in the region or country; this information may include: age and sex, reason for non-inclusion, indicator of socio-economic status (such as educational level) and disease-associated variables such as severity and treatment.

A.3. Data elements

A.3.1. Identification of data elements to be routinely collected

Data elements from routine clinical care to be collected in a new disease registry should be defined in a multidisciplinary approach with clinicians, patients' organisations and experts of the disease as well as regulators (NCAs, EMA), HTA bodies, payers' organisations and other potential users of registry information, as applicable. Ethics approval of the data elements at a local or national level may also be required.

Definitions should be in line with existing general and disease-specific guidelines for validated outcomes and laboratory tests (e.g. clinical trial guidelines) (45). Definitions, lag times for data availability and data dictionaries should be included in the registry documentation and published or made available in a standard and machine-readable format. It should be clear whether data elements originate from patient self-reports, medical reports or a third-party, as this distinction may have an impact on quality management and data analysis and interpretation. Processes should be put in place to allow the modification or expansion of the set of data elements to meet the potential needs of future registry-based studies.

A.3.2. "Core" versus "optional" data elements

"Core" data elements are those that are considered essential for the purpose of the registry or the network of registries. They should be collected from all patients in all concerned registries and are those on which greater amounts of resources should be allocated to ensure data quality.

"Optional" data elements are those considered of interest and useful to some stakeholders, but not essential to all. The distinction between core and optional data elements may vary according to the scope of potential registry-based studies planned to be performed and the capacity of centres to collect and report data in routine clinical care, e.g. lifestyle factors such as smoking or alcohol use.

The dataset should normally contain the core data elements listed below but the list can be adapted to each situation, for example as regards data elements that remain constant and those that might need to change as time progresses, treatments considered "current" or "concomitant" or diagnoses that may change over time:

- Administrative information: name of centre, availability of informed consent if applicable; registry entry date (for example, date of first contact or date of initial diagnosis); registry exit date and reason for exit (e.g. due to death, move outside the catchment area or other reason); dates of encounters in clinical practice;
- Patient data: age or birthdate (where permitted), gender;
- Disease: diagnosis (dates of initial diagnosis and of final diagnosis if relevant, laboratory tests and results; for diseases where the date of a clinical diagnosis is difficult to determine, date of first consultation, duration of disease or other appropriate information may be used), grade/severity/stage of disease, genomic information if important for the disease, relevant

prognostic factors, relevant milestones in disease monitoring (e.g. laboratory tests, imaging) and core disease outcomes (e.g. remission, relapse, disabilities, functional status, hospitalisation, cause of death);

- Co-morbidities: relevant co-morbidities differentiating past and current ones; co-morbidities to be included in a relevant validated co-morbidity index score may be considered;
- Disease-related treatments: substance, brand name, start and end dates (dates of prescription or dispensing), dose, route, schedule; reason for discontinuation (e.g. progression, adverse event, toxicity, completed planned treatment);
- Relevant concomitant therapies: substance, brand name, indication, dose, route, schedule, start and end dates (dates of prescription or dispensing), surgical and other therapies as applicable;
- Adverse events of special interest (AESIs) and serious adverse events occurring in the course of treatment; selection of AESI should be based on clinical safety information available for the study population, disease, condition or medicinal product;
- Current pregnancy: pregnancy status, pregnancy outcome;
- Patient-reported outcomes collected in clinical practice;
- Additional core data elements defined in disease-specific regulatory guidelines.

A date for important events, exposure and outcomes allows computation of precise time periods critical to the valid analysis of the data of a registry-based study, such as time between entry into the registry and treatment start, time under different treatments, time of onset of AESIs, time to remission of disease, or duration of follow-up. Knowledge of the person-time at risk of an event is also needed to calculate key epidemiological indicators such as incidence rates and perform time-dependent analyses. Consideration should also be given to collection of information referring to the period prior to initial registry enrolment.

Examples of lists of data elements to be collected for disease registries have been published in EMA guidelines (for example, the EMA guidelines for the clinical investigation of recombinant and human plasma-derived factor VIII and factor IX products (46) (47), the reports of EMA workshops on registries for cystic fibrosis (48), multiple sclerosis (49), diseases for which CAR-T cell products are indicated (50), haemophilia (51), and on registries for cancers which therapies are based on the tumours' genetic and molecular features (52). The EU RD Platform has developed a "Set of Common Data Elements for Rare Diseases Registration" (8). This set is aimed to the European Reference Network's registries, to all other rare disease registries at national, regional, and local level in the Member States, to researchers and to patient organisations. Other examples exist in the medical literature. The ENCePP Resources Database (53) contains information on disease registries that may be consulted when developing a list of data elements. Appendix I of GVP Module P III (21) provides possible parental and neonatal data elements from which relevant items can be selected when establishing a questionnaire of pregnancy exposure to medicinal products. These data-elements can also be relevant for setting up a pregnancy registry.

A.3.3. Standardisation of data elements

Data elements collected from patients should ideally be harmonised to international standards across all centres participating in a registry and all registries participating in a network of registries. Such harmonisation supports implementation of a common data quality system (e.g. automated data entry control and check for data consistency), data exchange, identical data analysis with the same programming codes, pooling of data and interpretation of results. Lack of harmonisation may require a mapping of data elements representing the same concept but implemented with different definitions

and terminologies. As this mapping process may be time-consuming and resource intensive, core data elements and formats should be preferably implemented at the design stage of registries. Where the harmonisation of data elements is not (yet) implemented, interim solutions should support use of registry data for regulatory purposes by mapping core data elements to the same terminologies. When a terminology has to be used in line with local requirements, this should be made clear.

Appendix 3 provides examples of recommended international terminologies for different data elements.

A.4. Quality management in patient registries

A.4.1. Framework for quality management

Uncertainties about the quality of the data collected in registries may undermine the confidence in the validity and reliability of the evidence generated from registry data in registry-based studies. The Commission Implementing Regulation (EU) No 520/2012 (54) and GVP Module I (55) provide a quality framework for MAHs, competent authorities of Member States and the EMA. Measurable quality requirements can be achieved by:

- Quality planning: establishing structures (including validated computerised systems) and planning integrated and consistent processes;
- Quality assurance and control: monitoring and evaluating how effectively the structures and processes have been established and how effectively the processes are being carried out;
- Quality improvement: correcting and improving the structures and processes where necessary.

These quality management activities ("plan, do, check, act") should be done in a continuous manner throughout the lifetime of the registry and be regularly assessed. They should be made available to patients, health care professionals and potential users of the registry data to provide confidence that quality management is adequately performed. Responsibilities should be clearly defined to enable sustainability of the quality management system. For registry-based clinical trials, GCP and clinical trial legal requirements on data quality should be met.

Data security should be part of quality management. Use of an existing patient registry for a new purpose, such as a registry-based study, may require availability of predefined data elements for specific users (e.g. users who perform data entry, management, quality control, extraction or analysis) but not necessarily all registry data. Specific measures (e.g., fire walls, log-in codes or access rights) may therefore need to be in place or introduced in the registry system when needed for some categories of users. Traceability (i.e., the possibility to trace changes made to patient data in the registry and who made these changes) should be part of the data security measures.

A.4.2. Requirements for data quality

In this context, data quality includes four main components.

- Consistency: the formats and definitions of the variables are consistent over time, across all centres within a registry and across all registries within a network of registries;
- Completeness: patient enrolment is maximised, patient attrition is minimised and complete information on a core data set is recorded for all eligible patients with minimisation of missing data;
- Accuracy: the data available in the registry is a correct representation of patient information available to the health care professional, e.g. data available in medical charts or laboratory test results; where the registry data are a compilation or duplication of electronic medical records at the point of care, accuracy should rely on a check of the extraction and uploading procedure;

- Timeliness: there is a timely recording and reporting of data and data updates, based on their intended use in compliance with an agreed procedure.

Requirements of data quality may be difficult to achieve concomitantly in all centres within a registry or within all registries of a network of registries; implementation of the same data elements, terminologies, data entry procedures and data control software may not be feasible simultaneously in all centres. Intermediate solutions may be adopted focussing on a core data set and mapping procedures. Centres may progressively implement components of data quality and be included in the registry or network of registries once they have achieved an adequate level of data quality as agreed between the concerned parties according to the data needs.

A comprehensive set of methods for assessing the quality of registries, for recruiting and retaining participants in a registry and for data collection and quality assurance is presented in the AHRQ Users' guide on registries for evaluating patient outcomes (1). Examples of practical aspects and techniques for addressing data quality in patient registries exist in the medical literature (56).

A.4.3. Key performance indicators of data quality

Registries should use performance indicators to assess and drive improvement of data quality. Such indicators should be measurable and associated with remedial measures if acceptable levels of quality are not found. Their definition depends on the disease, governance, infrastructure, local health system and processes in place within the registry or network of registries. They should therefore be defined in a multi-disciplinary approach with all concerned parties. Examples of agreed key performance indicators of data quality are presented in the reports of the EMA workshops on cystic fibrosis registries (48), multiple sclerosis registries (49) and CAR T-cell Therapy Registries (50).

A.4.4. Data quality management activities

Quality management can be supported by the activities described below. These activities should take into account appropriate technical and organisational measures to be implemented to ensure a sufficient level of security when personal data and more specifically health data is processed. Such measures should at least consist of pseudonymisation, encryption, non-disclosure agreements, strict access role distribution, access role restrictions as well as access logs. National provisions, which may stipulate specific technical requirements or other safeguards such as adherence to professional secrecy rules should be also taken into account. Given the variety in the organisation and infrastructure of registries, these recommendations should be adapted to each situation.

- Data quality management activities should be documented, communicated, maintained and updated as necessary, and all relevant source documents should be kept, managed and made available for auditing purposes in a timely manner, including:
  - standard operating procedures, steps of data quality management from data planning to reporting, with data management responsibilities;
  - key performance indicators of data quality, planned and performed data checks (manual or automated) and cleaning processes including query management and on-site monitoring.
- Support tools should be developed and provided, e.g., data collection and reporting software, support function (helpdesk), training material and training sessions. A centralised remote electronic quality control could be set-up to limit on-site visits to be done according to a pre-defined risk approach.
- Appropriate qualification and training of data managers and other persons involved in the data collection process should be ensured, with knowledge about the disease, exposures and outcomes captured in the registry.

- In case of a local data extraction process or manual data entry, routine data quality checks should be performed to alert on erroneous, missing or out-of-range values and logical inconsistencies, and trigger prompt data verification and remedial measure if needed. The validity of any data cleaning, extraction and transformation processes should be documented, especially if it involves mapping of data to a common terminology.
- Internal or external audits with on-site review of processes and data audits should be performed according to a risk-based approach; remote quality control measures, targeted visits and targeted source data verification should be triggered by pre-defined thresholds of data quality measures. The minimum amount of data verification required may depend on the amount of data collected and should ideally take into account critical aspects of data collection where differences may occur, e.g. between individual centres or between persons within individual centres.
- Aggregated registry data should ideally be compared to literature data or data from external data sources such as electronic health records or insurance claims databases as regards the distribution of categories of important variables such as age, gender, factors associated with disease occurrence or severity, or drug exposure.
- Feedback on findings on data quality issues should be given systematically to data providers so that escalation and remedial action can be taken at the level of the data source.
- When considering implementation of corrective and preventive activities, additional workload for data collection and data entry should be addressed, as a cumbersome data entry process may increase the amount of missing data and decrease data quality.

A.5. Governance

Registries generally operate under governance principles influenced by their purpose, operating procedures, legal environment or funding sources (55). Different parties may potentially also have divergent priorities, such as scientific independence, fulfilment of regulatory commitments, transparency or intellectual property rights. Clear governance principles supporting effective collaborations between all parties for regulatory use of registries, including data sharing between stakeholders, are therefore useful. Useful guidance can be found in the ENCePP Code of Conduct (57), which provides principles of scientific independence and transparency for pharmacoepidemiological research, and the ADVANCE Code of Conduct (43) for collaborative vaccine studies. The AHRQ User's Guide on patient registries provides a complementary source of recommendations on the governance of registries (1). EUnetHTA's Registry Evaluation and Quality Standards Tool (REQueST) provides essential standards that support the assessment of registry governance to assure general data quality and personal dana protection (7).

Registry holders should consider the following aspects to ensure transparency, best use and sustainability of their registry:

- To publish documentation of key registry characteristics, such as purpose of the registry, inclusion and exclusion criteria for participating centres and enrolment of patients, core and optional data sets collected (with timelines and frequency of data uploads), quality management process and experience of previous collaborations; the registry should be registered in the ENCePP Resources Databases (53).
- To establish a governance structure for the management of the registry and registry-based studies, with a steering committee, ethics committee and scientific advisory board.
- To establish a single contact point within the registry or network of registries for requesting information on available data and data access conditions.

- To publish a policy for collaborations with external organisations, including information on the scope and decision-making process for participating in collaborations, policy for data sharing and data analysis (explaining possible options for data transfer and analysis based on data privacy rules in place), possible involvement of a third-party, publication policy, and principles for private and public funding.
- To provide a supportive scientific and technical function for collaborations, which may include support for the development of the study protocol, interoperability between registries, amendments to the scope, schedule or methods of data collection or extraction, data management and analysis; the support provided may vary according to the approach of collaboration for using multiple data sources (see the ENCePP Guide on Methodological Standards in Pharmacoepidemiology) (31), resources available in the registry and the contractual agreements proposed.
- To develop a template for research contracts between the registry and external organisations, in line with those recommended by the ENCePP Code of Conduct (57) or the ADVANCE Code of Conduct (43).

A.6. Data sharing outside the context of registry-based studies

There may be situations where registry data could be shared outside the context of formal registry-based studies in the format of counts, aggregated data or statistical reports with NCAs, EMA, MAAs/MAHs, HTA bodies, payer organisations or other parties for clinical development planning or the evaluation or monitoring of medicinal products. These data may concern, for example:

- Disease epidemiology in terms of prevalence, incidence, outcomes, prognostic factors, potential confounding variables for defined outcomes;
- Size and distribution of the population with a specific disease, condition or exposure for a planned clinical trial or non-interventional study according to demographics, co-morbidities or medication use;
- Drug utilisation, with number of prescriptions for specific medicinal products (or other indicator of intensity of exposure), indications, dose, route of administration, schedule, duration of use, co-medications or use in specific population groups such as extent of paediatric use;
- Medical device utilisation, with number, types, indications and dates for specific implanted products;
- Surgical procedures with numbers, types, indications, dates and any other relevant details;
- Safety information on medicinal products, for example summary tables of adverse events recorded for specific medicinal products, aggregated data or anonymised line listings of patients presenting AESIs, or outcomes of exposed pregnancies;
- Utilisation of health care resources such as number of visits, hospitalisations, or laboratory tests performed.

This information may require capacity for sound analysis within the registry or, if allowed by the registry governance and patient consent, transfer of an anonymised dataset with selected variables to the requester or a third-party performing the analysis on behalf of the registry or the requester. Data sharing may require a contractual agreement between the registry or network of registries and the other concerned parties.

References

1. Agency for Healthcare Research and Quality. Registries for Evaluating Patient Outcomes: A User's Guide: 3rd Edition | Effective Health Care Program [Internet]. Available from: https://effectivehealthcare.ahrq.gov/products/registries-guide-3rd-edition/research

2. European Medicines Agency. Patient registries [Internet]. Available from: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/patient-registries

3. Olmo CA, McGettigan P, Kurz X. Barriers and Opportunities for Use of Patient Registries in Medicines Regulation. Clinical Pharmacology & Therapeutics. 2019 Jul 1;106(1):39–42. Available from: https://ascpt.onlinelibrary.wiley.com/doi/full/10.1002/cpt.1414

4. European Medicines Agency. Qualification Opinion - European Cystic Fibrosis Society Patient-Registry (ECFSPR) [Internet]. Available from: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-european-cystic-fibrosis-society-patient-registry-ecfspr_en.pdf

5. European Medicines Agency. Draft qualification opinion on Cellular therapy module of the European Society for Blood & Marrow Transplantation (EBMT) Registry [Internet]. Available from: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/draft-qualification-opinion-cellular-therapy-module-european-society-blood-marrow-transplantation_en.pdf

6. Zaletel M, Kralj M, Magajne M, Doupi P. Methodological Guidelines and Recommendations for Efficient and Rationale Governance of Patient Registries: Eur J Public Health [Internet]. 2015 Oct 1 25(suppl_3). Available from: https://academic.oup.com/eurpub/article/25/suppl_3/ckv169.006/2578054

7. EUnetHTA. REQueST® Tool and its vision paper [Internet]. Available from: https://eunethta.eu/request-tool-and-its-vision-paper/

8. E-Rare. JRC EU RD Platform releases the Set of Common Data Elements for RD Registration | ERA-Net E-Rare [Internet]. Available from: https://vascern.eu/what-we-do/patient-registries/

9. European Medicines Agency. Guidelines relevant for advanced therapy medicinal products [Internet]. Available from: https://www.ema.europa.eu/en/human-regulatory/research-development/advanced-therapies/guidelines-relevant-advanced-therapy-medicinal-products

10. Follow-up of patients administered with gene therapy medicinal products [Internet]. European Medicines Agency. Available from: https://www.ema.europa.eu/en/follow-patients-administered-gene-therapy-medicinal-products

11. McGettigan P, Alonso Olmo C, Plueschke K, Castillon M, Nogueras Zondag D, Bahri P, Kurz X, Mol PGM. Patient Registries: An Underused Resource for Medicines Evaluation: Operational proposals for increasing the use of patient registries in regulatory assessments. Drug Saf. 2019;42(11):1343-1351. Available from: https://link.springer.com/article/10.1007/s40264-019-00848-9

12. European Medicines Agency. Scientific advice and protocol assistance [Internet]. 2018. Available from: https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-advice-protocol-assistance

13. Lauer MS, D'Agostino RB. The Randomized Registry Trial — The Next Disruptive Technology in Clinical Research? New England Journal of Medicine. 2013 Oct 24;369(17):1579–81. Available from: https://www.nejm.org/doi/full/10.1056/NEJMp1310102

14. Gouwei et al. Registry-based randomized controlled trials- what are the advantages, challenges, and areas for future research? Journal of Clinical Epidemiology [Internet]. 2016 Aug 21. Available from: https://www.jclinepi.com/article/S0895-4356(16)30350-X/fulltext#secsectitle0010

15. Jansen-van der Weide M et al. Rare disease registries: potential applications towards impact on development of new drug treatments. 2018. Sep 5; Available from: https://doi.org/10.1186/s13023-018-0836-0

16. European Commission. EUROPEAN PLATFORM ON RARE DISEASE REGISTRATION (EU RD Platform), SET OF COMMON DATA ELEMENTS FOR RARE DISEASES REGISTRATION. Available from: https://eu-rd-platform.jrc.ec.europa.eu/sites/default/files/CDS/EU_RD_Platform_CDS_Final.pdf

17. Karanatsios B, Prang K-H, Verbunt E et al. Defining key design elements of registry-based randomised controlled trials: a scoping review. Trials 2020;21(520):1-22. Available from: https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-020-04459-z

18. European Medicines Agency - Scientific Guidance on post-authorisation efficacy studies [Internet]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/scientific-guidance-post-authorisation-efficacy-studies-first-version_en.pdf

19. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) - Module VIII - Post-authorisation safety studies [Internet]. Available from: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

20. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) - Module XVI - Risk minimisation measures: selection of tools and effectiveness indicators (Rev 2) [Internet]. Available from: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

21. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) - Product- or Population-Specific Considerations III: Pregnant and breastfeeding women [Internet]. Available from: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

22. Jansen P, van den Berg L, van Overveld P et al. Research Data Stewardship for Healthcare Professionals. In: Fundamentals of Clinical Data Science [Internet]. Cham (CH): Springer; 2019. Chapter 4. 2018. Available from: https://pubmed.ncbi.nlm.nih.gov/31314246/

23. European Medicines Agency. Innovation in medicines | European Medicines Agency. Available from: https://www.ema.europa.eu/en/human-regulatory/research-development/innovation-medicines

24. European Medicines Agency. PRIME: priority medicines. Available from: https://www.ema.europa.eu/en/human-regulatory/research-development/prime-priority-medicines

25. EU PAS Register [Internet]. ENCePP. Available from: http://www.encepp.eu/encepp/studiesDatabase.jsp

26. European Medicines Agency. ICH E6 (R2) Good clinical practice. Available from: https://www.ema.europa.eu/en/ich-e6-r2-good-clinical-practice

27. European Medicines Agency. ICH E8 General considerations for clinical studies. Available from: https://www.ema.europa.eu/en/ich-e8-general-considerations-clinical-studies

28. European Medicines Agency. ICH E9 statistical principles for clinical trials. Available from: https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials

29. ICH harmonised guideline - Addendum on estimands and sensitivity analysis in clinical trials - E9(R1) [Internet]. Available from: https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf

30. European Medicines Agency. Guidance for the format and content of the protocol of non-interventional post-authorisation safety studies [Internet]. Available from: https://www.ema.europa.eu/en/documents/other/guidance-format-content-protocol-non-interventional-post-authorisation-safety-studies_en.pdf

31. ENCePP Guide on Methodological Standards in Pharmacoepidemiology [Internet]. ENCePP. Available from: http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml

32. ENCePP Checklist for Study Protocols [Internet]. ENCePP. Available from: http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml

33. General Data Protection Regulation (GDPR) Compliance Guidelines. GDPR.eu. Available from: https://gdpr.eu/

34. European Commission. Question and Answers on the interplay between the Clinical Trials Regulation and the General Data Protection Regulation. https://ec.europa.eu/health/sites/default/files/files/documents/qa_clinicaltrials_gdpr_en.pdf

35. European Commission. Informed consent guidance. https://ec.europa.eu/research/participants/data/ref/fp7/89807/informed-consent_en.pdf

36. European Commission. Eudralex Vol. 10. Risk proportionate approaches in clinical trials. Available from: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/2017_04_25_risk_proportionate_approaches_in_ct.pdf

37. European Medicines Agency. Reflection paper on risk-based quality management in clinical trials [Internet]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-risk-based-quality-management-clinical-trials_en.pdf

38. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) - Module III – Pharmacovigilance inspections. Available from: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

39. European Medicines Agency. Guideline on Missing Data in Confirmatory Clinical Trials [Internet]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-missing-data-confirmatory-clinical-trials_en.pdf

40. 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 Available from: https://www.acpjournals.org/doi/10.7326/M18-3079

41. European Medicines Agency policy on publication of clinical data for medicinal products for human use [Internet]. European Medicines Agency. Available from: https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/clinical-data-publication

42. ENCePP. The ENCePP Code of Conduct [Internet]. Available from: http://www.encepp.eu/code_of_conduct/documents/ENCePPCodeofConduct.pdf

43. Kurz X, Bauchau V, Mahy P, Glismann S, van der Aa LM, Simondon F, et al. The ADVANCE Code of Conduct for collaborative vaccine studies. Vaccine. 2017 04;35(15):1844–55. Available from: https://pubmed.ncbi.nlm.nih.gov/28285984/

44. European Data Protection Board. European Data Protection Board. Available from: https://edpb.europa.eu/about-edpb/board/members_en

45. European Medicines Agency. Clinical efficacy and safety guidelines [Internet]. Available from: https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-guidelines/clinical-efficacy-safety-guidelines

46. Clinical investigation of recombinant and human plasma-derived factor VIII products [Internet]. European Medicines Agency. Available from: https://www.ema.europa.eu/en/clinical-investigation-recombinant-human-plasma-derived-factor-viii-products

47. Clinical investigation of recombinant and human plasma-derived factor IX products [Internet]. European Medicines Agency. Available from: https://www.ema.europa.eu/en/clinical-investigation-recombinant-human-plasma-derived-factor-ix-products

48. European Medicines Agency. Report on Cystic Fibrosis Registries - Workshop - 14 June 2017 [Internet]. Available from: https://www.ema.europa.eu/en/documents/report/report-cystic-fibrosis-registries_en.pdf

49. European Medicines Agency - Report on Multiple Sclerosis Registries - Workshop - 7 July 2017 [Internet]. Available from: https://www.ema.europa.eu/en/documents/report/report-multiple-sclerosis-registries_en.pdf

50. European Medicines Agency. Report on CAR T-cell therapy Registries Workshop - 9 February 2018 [Internet]. Available from: https://www.ema.europa.eu/en/documents/report/report-car-t-cell-therapy-registries-workshop_en.pdf

51. European Medicines Agency - Report on Haemophilia Registries Workshop - 8 June 2018 [Internet]. Available from: https://www.ema.europa.eu/en/documents/report/report-haemophilia-registries-workshop_en.pdf

52. European Medicines Agency. Report of the workshop on the use of registries in the monitoring of cancer therapies based on tumours' genetic and molecular features - 29 November 2019 [Internet]. [cited 2020 May 9]. https://www.ema.europa.eu/en/documents/report/report-workshop-use-registries-monitoring-cancer-therapies-based-tumours-genetic-molecular-features_en.pdf

53. ENCePP Resources Database [Internet]. ENCePP. Available from: http://www.encepp.eu/encepp/resourcesDatabase.jsp

54. EUR-Lex. The Commission Implementing Regulation (EU) No 520/2012 [Internet]. Available from: https://eur-lex.europa.eu/eli/reg_impl/2012/520/oj

55. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) Module I – Pharmacovigilance systems and their quality systems [Internet]. Available from: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

56. Bray F, Parkin D. Evaluation of data quality in the cancer registry: Principles and methods. Eur J Cancer 2009; 45(5):747-64 Available from: https://pubmed.ncbi.nlm.nih.gov/19117750/

57. Gini et al. The ENCePP Code of Conduct: A best practise for scientific independence and transparency in noninterventional postauthorisation studies. Pharmacoepidemiol Drug Safety [Internet]. 2019 Mar 5 Available from: https://www.ncbi.nlm.nih.gov/pubmed/30838708

Appendices

Appendix 1. *Checklist for evaluating the suitability of registries for registry-based studies*

(List adapted from the REQuEST tool published by EUnetHTA) (6)

1. Administrative information

1.1. Governance for collaborations

- Publicly available documentation (with website) of key registry characteristics
- Single contact point for information
- Publicly available policy for collaborations with external organisations
- Governance structure for decision-making on requests for collaboration
- Supportive scientific and technical function
- Supportive function for ethical and legal aspects
- Template for research contracts between the registry and external organisations

1.2. Informed consent and data protection

- Status of implementation of GDPR
- Nature of consent requested in the Informed consent form, incl. permission to use and share data for research purpose (or need for re-consent), to re-contact patients and to use data for quality management, audits and inspections.
- Compliance with applicable data protection rules as set out in chapter 4

1.3. Funding

- Funding sources and impact on short, long-term sustainability and possible conflicts of interests for a specific registry-based study

2. Methods

2.1. Objectives

- Purpose of the data collection system, which may influence the main characteristics of the registry population and the data collected

2.2. Data providers

- Description of data providers, such as patients, carers or health care professionals (with different specialties), their geographical area and any selection process (inclusion and exclusion criteria) that may be applied for their acceptance as data providers

2.3. Patient population covered

- Type of patient registry (disease, condition, time period covered, procedure), which defines the criteria for patient eligibility
- Setting and catchment area
- Patients' inclusion and exclusion criteria
- Methods applied to minimise selection bias and loss to follow-up
- Numbers of patients available in the registry (total number and number of eligible patients if applicable), numbers of new patients entering the registry per year, numbers of patients lost per year (with reasons for exit)
- Mean/median duration of follow-up per patient, person-time of exposure in defined categories, if applicable

2.4. Data elements

- Core data set collected from patients by all centres;
- Optional data set;
- Definition, dictionary and format of data elements
- Standards and terminologies applied
- Capabilities and plans for amendments of list of data elements

2.5. Infrastructure

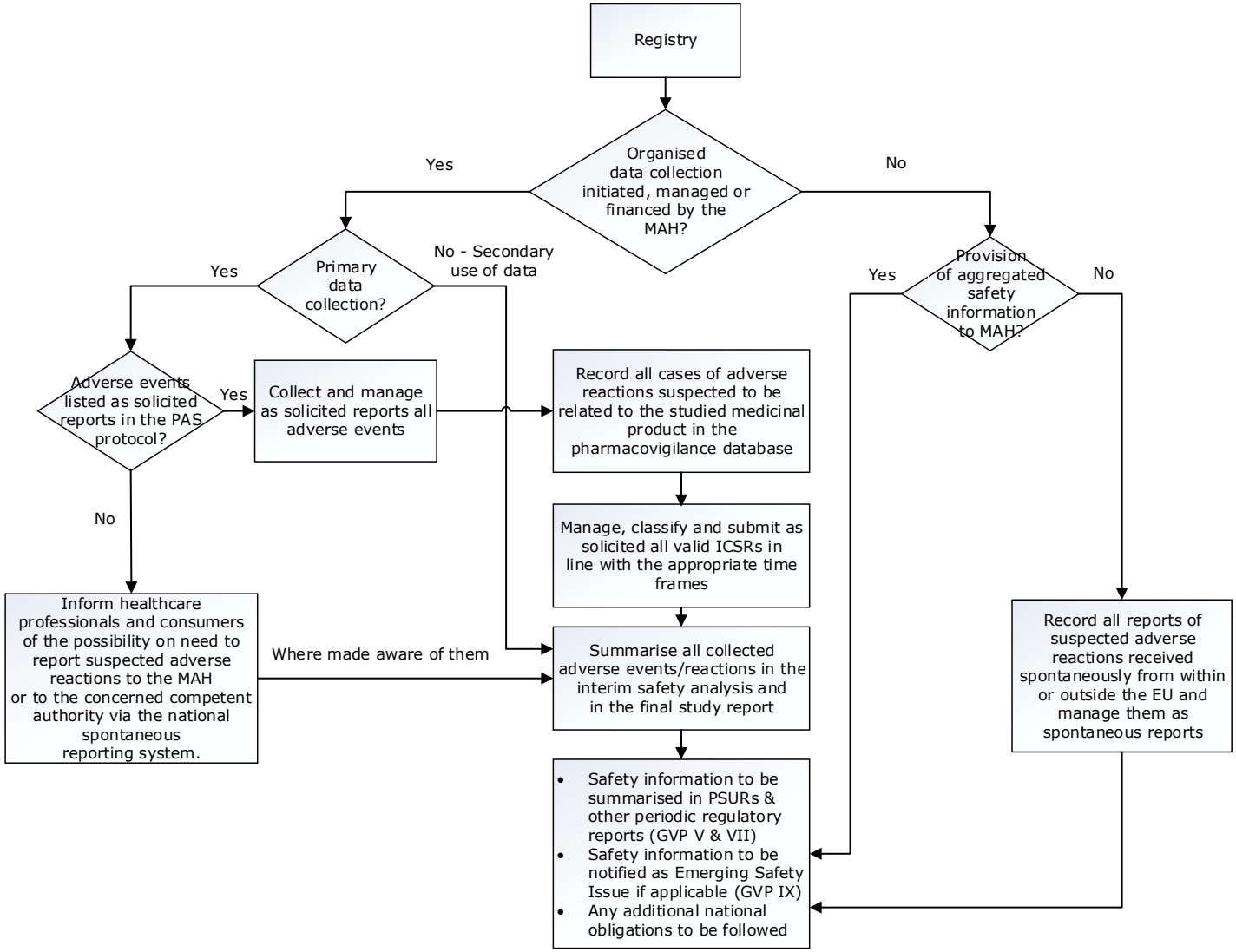
- Systems for data collection, recording and reporting, including timelines
- Capability (and experience) for expedited reporting and evaluation (at physician or registry level) of serious adverse reactions in primary data collection
- Capability (and experience) for periodic reporting of clinical outcomes and adverse events reported by physicians, at individual-patient level and aggregated data level
- Capability (and experience) for data cleaning, extraction, transformation and analysis
- Capability (and experience) for data transfer to external organisations
- Capability (and experience) for record linkage
- Capabilities for amendment of safety reporting processes

2.6. Quality requirements

- Processes in place for quality planning, control, assurance and improvement
- Data verification (method and frequency of verification)
- Missing data (statistics, trends, variables affected, management)
- Auditing practice

Appendix 2. *Safety reporting*

Overview of MAH responsibilities for the collection and reporting of adverse events and suspected adverse reactions when a registry-based study fulfils the definition of a non-interventional post-authorisation study according to the clinical trial legislation (GVP Modules VI, VIII).



Appendix 3. *Examples of recommended international terminologies for data elements*

Data elements	Terminologies	Weblinks
Diseases, diagnostics, symptoms, indication for use of medicine	ICD-9, ICD-10, ICD-11	http://www.who.int/classifications/icd/en/
	ICD-o-3 (cancers)	http://codes.iarc.fr/
	SNOMED-CT	https://www.snomed.org/snomed-ct
	CDISC Standards	https://www.cdisc.org/
	MedDRA	https://www.meddra.org
Rare disorders (disease, malformation syndrome, clinical syndrome, morphological or biological anomaly or particular clinical situation in the course of a disorder), Groups of rare disorders, Sub-types of rare disorders	Orphanet (ORPHAcodes) (entries are cross-referenced with ICD-10, ICD-11, OMIM, UMLS, MeSH, MedDRA) Orphadata	http://www.orphadata.org/cgi-bin/inc/product1.inc.php http://www.orphadata.org/cgi-bin/index.php#ordomodal
	HPO (Human Phenotype Ontology)	https://hpo.jax.org/app/
Medicinal products	Article 57 database (EEA)	https://www.ema.europa.eu/en/human-regulatory/post-authorisation/data-medicines-iso-idmp-standards/public-data-article-57-database
	ISO IDMP standards and related terminologies	https://www.ema.europa.eu/en/human-regulatory/overview/data-medicines-iso-idmp-standards-overview
	ATC classification	https://www.whocc.no/atc_ddd_index/
	RxNorm (US)	https://www.nlm.nih.gov/research/umls/rxnorm/index.html
AESI, other adverse events, suspected adverse reactions	MedDRA[1]	https://www.meddra.org/
	CTCAE (cancer therapies; includes MedDRA)	https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm
	IMDRF	http://www.imdrf.org/documents/documents.asp
	Division of AIDS (DAIDS) AE severity grading tables	https://rsc.niaid.nih.gov/clinical-research-sites/daids-adverse-event-grading-tables
Routes of administration, pharmaceutical dosage	EDQM Standard Terms Database	https://www.edqm.eu/en/standard-terms-database

[1] In accordance with the Commission Implementing Regulation (EC) No 520/2012, Member States, marketing authorisation holders and the Agency shall apply MedDRA as the internationally agreed standard terminology for the classification, retrieval, presentation, benefit-risk evaluation and assessment, electronic exchange and communication of pharmacovigilance and medicinal product information

Data elements	Terminologies	Weblinks
forms, packaging, units of administration		
Test results units	Unified Code for Units of Measure (UCUM)	http://unitsofmeasure.org/ucum.html
Genetic diagnosis	International classification of mutations (HGVS)	http://www.hgvs.org/
	HUGO Gene Nomenclature Committee (HGNC)	https://www.genenames.org/
Classification of functioning/disability	ECOG	Eastern Cooperative Oncology Group (ECOG) Performance Status
	ICF	http://www.who.int/classifications/icf/whodasii/en/
Terminologies and formats for individual case safety reports (ICSR) specification	Code Sets and Object	http://estri.ich.org/e2br3/index.htm
	Identifiers based on the ICH E2B R(3) ICSR	Clinical Trial Facilitation Group (CTFG). Reference Safety Information 2017.pdf
	Implementation Guide	Clinical trials, Regulation (EU) No 536/2014, Questions and Answers
Harmonised data formats	Observational Medical Outcomes Partnership (OMOP) common data model	https://ohdsi.org/who-we-are/

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

15 December 2017
EMA/526382/2017
Inspections, Human Medicines, Pharmacovigilance and Committees Division

ENCePP Resource Database

Inventory of Patient Registries

Table of Contents

Introduction **2**

A. How to enter your registry in the ENCePP Resources Database **2**

B. How to search for a registry in the ENCePP Resources Database **4**

C. Other relevant links **6**

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5520
Send a question via our website www.ema.europa.eu/contact

An agency of the European Union

© European Medicines Agency, 2018. Reproduction is authorised provided the source is acknowledged.

Introduction

The aim of this document is to provide key points on:

- **How to <u>enter</u> patient / disease registries**.
- **How to <u>search</u> patient / disease registries.**

The main objective of the ENCePP inventory of registries is to be a network tool for disease and patient registries.

A. How to <u>enter</u> your registry in the ENCePP Resources Database

1. Go to the ENCePP website Resources database: http://www.encepp.eu/encepp/resourcesDatabase.jsp
2. Select "Add Data Source"

3. After selecting "Add Data Source", you will have to choose a database classification (Centre/Network/Data Source). Select "Data source".
4. Register data source question: "Is data source maintained by a centre registered in the ENCePP inventory of research centres?" Select as appropriate (yes/no/don't know)
5. Questionnaire

Fields marked with (*) are mandatory. They must be completed in order to move forward to the next question.

1. Contact details

1. Name of data source*: complete name (long name)
2. Name under which the data source should be displayed in the inventory (max 50 characters)*: short name – *add disease*

Please note that the patient registry disease must appear next to the short name. This is to ensure that at least the disease appears in one of the two above mentioned fields (that will help when searching specific disease patient registries). Please see the following example:

1. Contact details	
Name of data source	European Register for Multiple Sclerosis
Name under which the data source should be displayed in the inventory (e.g short name if one exists)	EUReMS - Multiple Sclerosis

3. Name of organisation holding the data.
    1. Department/research group
    2. Organisation/affiliation*
    3. Website/Homepage
    4. Administrative contact *
    5. Scientific contact *

2. How would the data source be most appropriately described? (Please choose all that apply):

Patient registries should have disease/case registries **and** exposure registry selected.

Sections 3.1 to 3.4 are optional

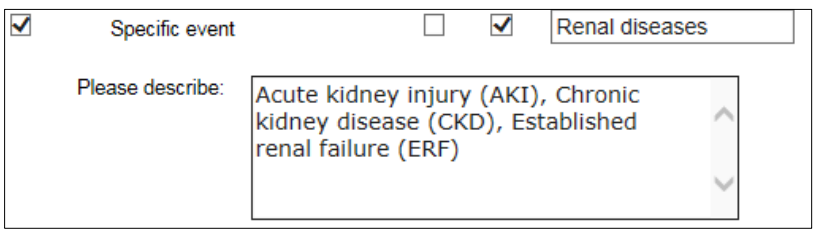
(Please complete all that apply)

4. Which kinds of events are covered?

This field is very important when searching for a specific disease registry stakeholder.

'Specific event' should be selected.

The disease/s and/or the field disease must be included as this assists in searches. The row on the specific event should be for the field disease, and the free box text for all the diseases included in the registry, if more than one disease is included.



Sections 5 to 12 are optional.

(Please complete all that apply)

13. Is there a written policy governing data access?*

Yes / No. If yes, please specify

14. Is there a committee to evaluate requests for data access?*

Yes / No

15. Is a charge made for data access?*

Yes / No. If yes, please specify.

16. Please list at least 1 of the 5 most relevant publications using the data for the last five calendar years. If there is a publication explicitly reporting on assessment of data quality, please include first

Please provide a link to relevant recent publications

B. How to search for a registry in the ENCePP Resources Database

1. Go to search database section: http://www.encepp.eu/encepp/search.htm
2. Select "Data Source"

   It is recommended to search the following 3 options:

   - *2.1 Search a registry by name*
   - *2.2. Search a registry by disease*
   - *2.3. Obtain the entire data sources list*

   The above options are described below:

2.1. Search a registry by name

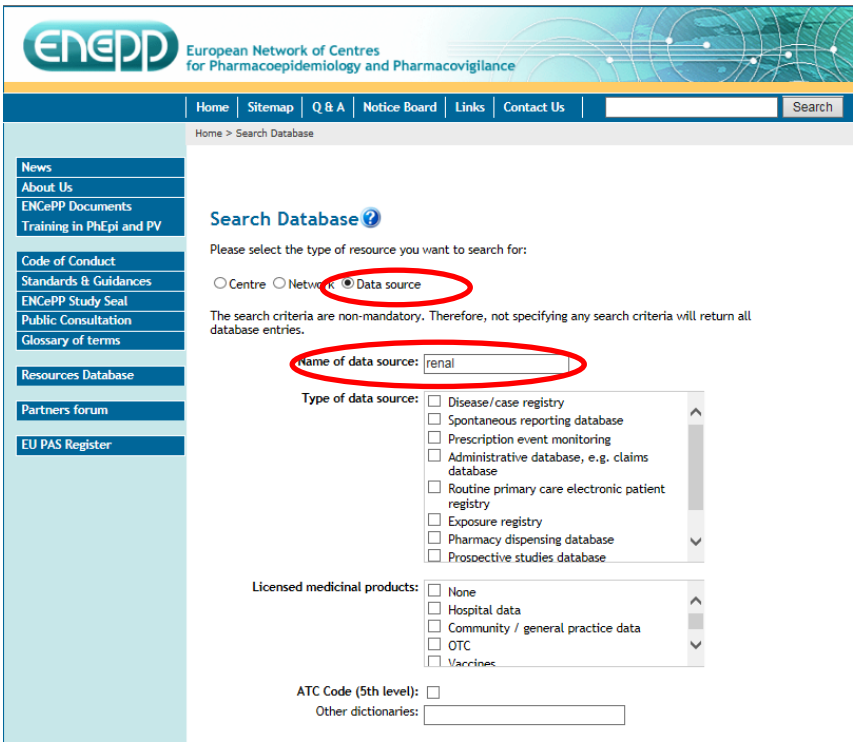
Type the complete name (long name) or short name (acronym) on the "name of data source" field.

2.2. Search a registry by disease

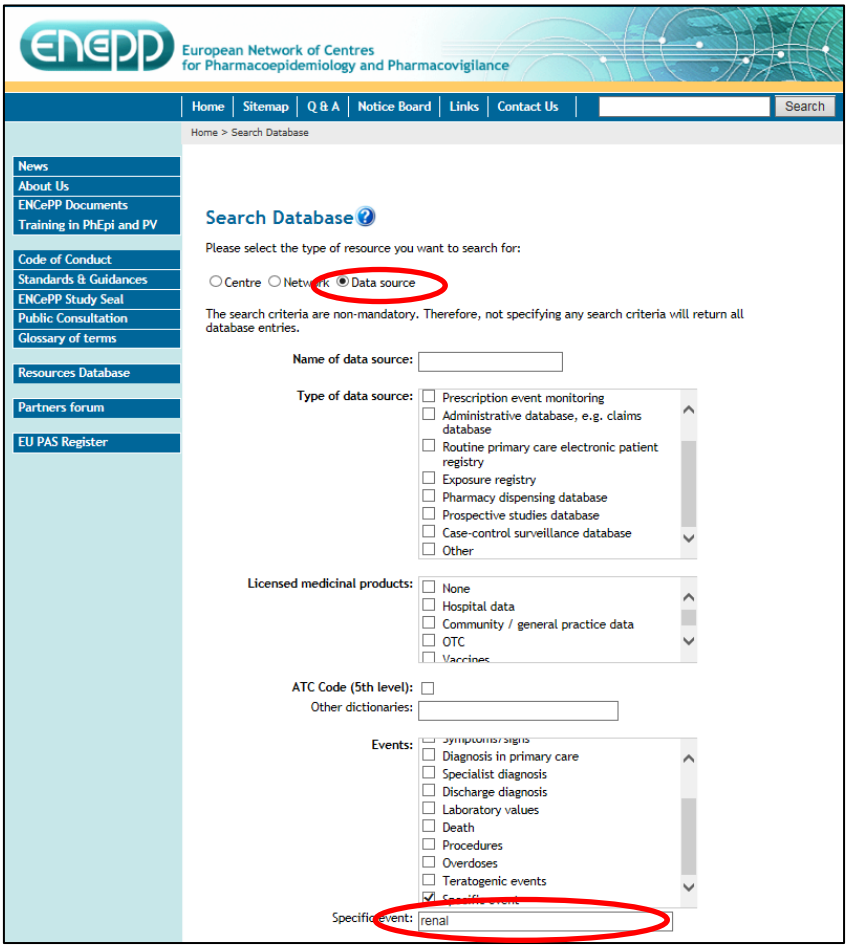
There are 2 different steps to find all the registries by disease:

1. Type the disease in the search field "name of data source

2. You will retrieve registries with the disease on their "name of data source" or "short name" field (acronym) field.



3. Then another new search should be done. Select “specific event” from the list of events. Type the disease in the “specific event” field:



2.3. Obtain the entire data sources list

1. Select “data source”
2. Select “search”
3. The complete list of data sources will be displayed.

C. Other relevant links

- *Patient registries initiative from the European Medicines Agency (EMA)*

  http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000658.jsp

- *European Commission*

  https://ec.europa.eu/commission/index_en

- *RD Connect Network*

  http://catalogue.rd-connect.eu/