EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

24 September 2020
EMA/502388/2020

Guideline on registry-based studies

Draft

Draft approved by the Cross-Committee Task Force on Registries	May 2020
Draft sent to the EU Regulatory Network for consultation	9 July 2020
Start of public consultation	24 September 2020
End of consultation (deadline for comments)	31 December 2020

Comments should be included in the form published with this draft guideline, and should be sent to EMAregistries@ema.europa.eu by 31 December 2020.

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

Table of contents

1. Introduction

A *registry-based study* is an investigation of a research question using the infrastructure of (a) new or (an) existing registry(-ies) for patient recruitment and data collection. A registry-based study may be a clinical trial, to which the provisions of Directive 2001/20/EC or of Regulation (EU) No 536/2014 (when it becomes applicable) apply, or a non-interventional study if it fulfills the corresponding requirements specified in Directive 2001/20/EC (see Annex of Questions & Answers document, Version 11.0, May 2013) or Regulation (EU) No 536/2014 [1] (1). A registry-based study may apply primary data collection and/or secondary use of data collected in a patient registry for another purpose than the given study (see definitions in Appendix 1). A *patient registry* is defined in this Guideline as an organised system that collects data and information on a group of people defined by a particular disease or condition, and that serves a pre-determined scientific, clinical and/or public health (policy) purpose. The use of the term 'patient' in combination with 'registry' (i.e. patient registry) is used to highlight the focus of the dataset on health information. The terms 'people' and 'patients' used in this definition and Guideline are synonyms, independently of the health status of the individual.

The EMA Patient Registry Initiative and the Cross-Committee Task Force on Registries (2) have explored ways to improve the use of patient registries for registry-based studies in order to support the benefit-risk evaluation of medicinal products. Recommendations on aspects to be addressed for such studies were issued in five workshops on specific registries (3) and in the CHMP Qualification Opinions for two registry platforms 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 Guideline, 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 (8), and the European Platform on Rare Diseases Registration (9).

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 studies. To support these recommendations, aspects of patient registries that regulators consider important for their use in registry-based studies are included in the Annex. Relevant legal basis and regulatory requirements that apply to these studies are listed in Chapter 4.

This Guideline focusses on studies based on *disease registries* or *condition registries* to study the utilisation, safety and effectiveness of medicines prescribed to or consumed by patients included in the registry. Such registries are characterised by the presence or occurrence of a particular disease or disease-related patient characteristic, such as a set of signs or symptoms, or a specific condition, such as a pregnancy (pregnancy registry), a birth defect or a molecular or genomic feature. They may have different purposes, such as to collect data on natural history of the disease, 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 may provide an important source of information on diseases, patients, standards of care, utilisation of drugs, devices and procedures and outcomes of treatments. They may, in particular, represent an important source of data on rare diseases or populations such as those treated with advanced therapy medicinal products (ATMP) (10), including gene therapy (11).

[1] In this Guideline, the terms "non-interventional study" is used to indicate both a non-interventional study (Regulation (EU) No 536/2014) and a non-interventional trial (Directive 2001/20/EC).

The term *product registry* is sometimes used to indicate a system of data collection targeting patients exposed to a specific medicinal product, single substance or therapeutic class and who are followed over time with the aim to evaluate the use, safety, effectiveness or another outcome of this exposure. This type of data collection system corresponds to a clinical trial or a non-interventional study and does not include specific aspects related to the use of patient registries. For these reasons, the term product registry is not used in this Guideline.

Details on procedural aspects related to the interactions with regulators 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 as appropriate.

Although this Guideline is primarily targeted to MAAs/MAHs, it is also relevant to patients and to persons involved in the funding, creation and management of registries, those participating in the collection and analysis of registry data, and those planning to use the registry information and infrastructure to perform registry-based studies with a possible regulatory purpose.

3. Methods and processes

3.1. Use of registry-based studies for evidence generation

The use of a registry-based study for a regulatory purpose depends on many factors related to its relevance to answer a specific research question, the characteristics of the concerned registry, the quality of the data collected and the design and analytical plan of the proposed study (12). Prior consultation with national competent authorities, where applicable, and with EMA via the procedure for Scientific advice and protocol assistance is therefore recommended when a registry-based study is proposed to be used (13). Examples where registry-based studies may be useful for evidence generation are presented below.

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

  Pre-clinical studies and clinical trials are at the core of the scientific evaluation of the efficacy and safety of medicines prior to granting a marketing authorisation. In some circumstances, this evaluation may be supported by observational evidence derived from patient registries. Examples of such evidence include information on standards of care for the disease, incidence and determinants of disease outcomes in clinical practice, characteristics of the target population, or validity of a surrogate endpoint used in the evaluation. In some Member States, diagnostic monitoring of patients, e.g. imaging methods such as CT-scans and laboratory testing, should be strictly limited to normal clinical practice if the registry-based study is not registered as a clinical trial.

  Studies based on patient registries may also contextualise the results of uncontrolled trials, provide comparator groups of patients for a single arm trial on a case-by-case basis where a randomised controlled trial (RCT) is deemed not feasible or unethical, and support registry-based randomised controlled trials (RRCTs) for patient recruitment (for example to identify patients meeting inclusion/exclusion criteria) and data collection (14) (15). It is recommended to obtain Scientific Advice from EMA and, where applicable, of the concerned national competent authorities on the acceptability of the chosen approach to evidence generation in case deviations from a traditional RCT design are considered.

- To provide data sources or infrastructure for post-authorisation evidence generation

Patient registry-based studies can be data sources for RCTs and non-interventional studies, post-authorisation efficacy studies (PAES) (16) or post-auhorisation safety studies (PASS) (17) that may be performed after marketing authorisation. The interventions performed to monitor efficacy or safety compared to the SmPC and normal clinical practice determines if the post authorisation study is a clinical trial or a non-interventional study, and randomisation of subjects results in the registry-based study being considered a clinical trial. In the context of products that have been previously investigated in RCTs, registry-based studies may help, for example, to estimate and predict the effectiveness of adapted drug dosing schemes applied in clinical practice and understand effectiveness and safety of medicinal products in a broader clinical disease-related context and a more heterogenous patient population. 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, or to assess patterns of drug utilisation that add to knowledge on the benefit risk profile of the medicinal product. Registry-based studies may require linkage between different data sources through a unique patient identifier, if feasible.

A large proportion of ATMPs are developed for very rare diseases. This has an impact on the type of clinical trials (e.g. single arm trials with external control groups) and the size of the safety and efficacy database at the time of approval. The follow-up of safety and efficacy of ATMPs after approval is therefore mandatory (10), and PAES and PASS are often imposed for post-authorisation evidence generation. These are frequently and preferentially performed on the basis of existing disease registries.

- To evaluate the effects of medications received during pregnancy

Pregnancy registries include pregnant women followed up to collect information on outcomes of pregnancy and in the offspring for a given medicinal product. Besides the challenges of recruitment and retention of pregnant women, specific challenges of such studies relate to the completeness of information on pregnancy outcomes and the ascertainment of the exposure window/trimester, which may require linkage with data captured in birth defects registries, teratology information services or electronic health care records where mother-child linkage is possible (18).

3.2. 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.

	Registry-based study	**Patient registry**
1. Definition	Investigation of a research question or hypothesis using data from an existing patient registry or from a registry newly set-up for the study.	Data collection system on a group of people defined by a particular disease or condition, established for a specific purpose and used to conduct a registry-based study.
2. Timelines	Timelines driven by the collection/extraction and analysis of the data relevant for the specific study objective(s).	Generally planned to be long-term; timelines driven by schedules for routine data collection and any anticipated data analyses which prompted the registry.
3. Patient enrolment	Defined by research objective(s) - may be a subset of a registry population; in	Aimed at complete enrolment within the boundaries of the purpose of the

	case of a clinical trial, allocation to treatment (e.g. with randomisation) is to be documented; representativeness and generalisability of the study results to be analysed and documented.	registry; representativeness and 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; additional data collection may also be required; if such additional data includes subject monitoring outside SmPC and normal clinical practice, the legislation for clinical trials apply; study may involve primary data collection or secondary use of data.	Wide range of data may be collected depending on the purpose of the registry; there should be an agreed core set of data elements to be collected with harmonised definitions, common coding system and common data entry procedures.
5. Analysis plan	Detailed statistical considerations most commonly defined in separate document in addition to study protocol and to registry protocol; hypothesis driven statistical analysis plan.	Statistical analysis plan with analyses that are often descriptive and performed routinely at intervals based on patient accrual or defined time schedules described in the registry protocol.
6. Data quality control	Additional quality assurance to be performed for the study data; quality control to be prospectively defined and assessed with a risk-based approach; for RRCTs, data quality control involves central adjudication of events and treatment complications.	Applied routinely to data and processes with a focus on core set of data elements; data systems to ensure data integrity (i.e. system validation).

3.3. Planning a registry-based study

Planning a registry-based study requires to identify one or several suitable registry(-ies), to obtain the agreement to collaborate from each registry as well as from each individual centre if no central registry coordination exists, to identify a third-party to be possibly involved in the study and to set up a database, a data extraction process and quality control activities.

It is therefore recommended to discuss early with regulators, through Scientific Advice, both nationally and at EMA, the feasibility of the use of the registries to meet regulatory needs and the legal requirements for clinical trials. The EMA PRIority MEdicines (PRIME) procedure (19), if applicable, and pre-submission meetings can also be used in the pre-authorisation phase. In case of ATMPs, a strategy for post-authorisation activities should be developed in the pre-authorisation phase and discussed in scientific advice and PRIME procedures if applicable.

Early discussions should take place with involvement of the concerned Rapporteurs or Lead Member States (and concerned EMA Committees) as well as the MAA/MAH, registry holders and health technology assessment (HTA) bodies if relevant. It is the responsibility of the MAA/MAH to include in the discussion the holders of the registry(-ies) intended to be used.

MAAs/MAHs proposing a registry-based study should provide adequate information regarding the availability of data, the quality management applied and the feasibility of introducing any additional data collection and quality control measures. In case of primary data collection, adequate measures may be needed to detect and promptly report adverse events of interest. A feasibility analysis should be considered by the MAA/MAH or research organisation initiating the study prior to the writing of the study protocol to guide its development and facilitate the discussion with regulators, HTA bodies and

any other party. The feasibility analysis should be performed in collaboration with registry holders and include the following information:

- General description of the registry(-ies) or coordinated registry network; the Checklist for evaluating the suitability of registries for registry-based studies (see Appendix 2) can be used to prepare this description.
- Analysis of the availability of the data elements needed for the study (including relevant confounding and effect-modifying variables) and of the capacity to collect any additional data elements or introduce additional data collection methods if necessary.
- Analysis of the quality and completeness of the available data elements needed for the study, information on missing data and possible data imputations, and results of any verification or validation (e.g. through an audit) performed; if several registries are planned to be used, analysis of the differences that may exist between them and of the possible impact of these differences.
- Description of processes in place for the identification, analysis and reporting of adverse events of special interest (AESIs), suspected adverse reactions (ADRs) or suspected unexpected serious adverse reactions (SUSARs), and capacity to introduce additional processes for their collection if needed.
- 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, duration of follow-up, missing data and losses to follow-up; based on this information, analysis of the feasibility of the study and of the time needed to complete patient recruitment for the study.
- Analysis of any potential information bias, selection bias due to the inclusion/exclusion criteria of centres and patients, potential time bias between and within registry(-ies), and potential losses to follow-up.
- Analysis of any potential confounding bias that may arise in the proposed registry-based study if some data elements are not available or cannot be collected or measured.
- Analytical issues that may arise based on the data characteristics and the study design.
- Any data privacy issues and governance-related issues such as data sharing and funding source (see chapter A.5 of the Annex).
- 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 and should be published in the EU PAS Register (20) in order to inform on the feasibility of other studies in the same registry and avoid duplication of work.

For regulatory studies addressing a class of products where all concerned 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. For clinical trials, this could be performed through joint trial sponsorship as provided for in 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 scientific question of interest, how the study will be conducted and how the validity (both internal and external) of the results will be ensured. The legislation on clinical trials should be followed

to determine if a registry-based study should be labelled as a clinical trial or a non-interventional study.

Designing a registry-based study also implies to consider how the requirements of the data protection legislation will be fulfilled in terms of adequate procurement of patient informed consent, depending on the type of study (clinical trial vs. non-interventional study) and the patient information consent that was signed when the patient initially registered (21). The study protocol should specify how the data protection regulation will be followed, e.g. if the data is not already provided in an anonymised way excluding the identification of the patient (see Chapter 4).

The study protocol should follow the existing regulatory requirements for the study topic and for its type of design, such as the ICH E6 (22), ICH E8 (23) and ICH E9 (24) guidelines, technical guidance on the format and content of the protocol for non-interventional PASS (25) or the Scientific Guidance on PAES (16). The study protocol should apply the best methodological standards, such as the ENCePP Guide on Methodological standards in pharmacoepidemiology (26).

The framework of the ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials (27) should be considered for studies aiming to measure treatment effects. The ENCePP Checklist for Study Protocols (28) identifies important questions to be addressed when designing a non-interventional study and writing the study protocol.

A registry-based study may include primary data collection (i.e. collection of information on the events of interest for the purpose of the study directly from the patients, caregivers, healthcare professionals or other persons involved in the patient care) and/or secondary use of data (i.e. use of data collected in the registry for a purpose other than the given study) (see Appendix 1). The method of data collection should be clearly specified in the study protocol as it has implications with regards to prior data knowledge, potential sources of bias and safety reporting requirements. Where the registry-based study entails secondary use of data, the study protocol should specify the events of interest that are/are not collected in the registry and discuss the risk for bias in such secondary data use. The protocol should also specify agreements made with the registry holder on the additional variables that can be added to the registry prior to study start, with timelines for their introduction and data availability. To avoid misclassification of outcome and information bias, dedicated and complete search strategies and appropriate definition of the outcome of interest should be performed.

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 (taking into account the 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)] (27). 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.

If a registry-based study is to be conducted across multiple registries, a common study protocol should be developed based on core data elements 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 combined across all registries. Nevertheless, the protocol should also describe differences between registries, critically discuss the potential impact of such differences and propose sensitivity analyses addressing these. Additional legal requirements apply if the registry-based study is a clinical trial.

3.5. Study population

The choice of patients for the study population should be driven by the study objectives (for example, need for an internal control group for comparison of different treatments) and has important implications for the interpretation of the results. When studying a drug of interest, the study population may include various groups of patients: newly diagnosed patients entering the registry with a first prescription of the drug of interest, patients already diagnosed with the disease and switched from another treatment or patients having already received the drug of interest (e.g. in a clinical trial). When all treated patients are included in the study population, it may be useful to collect the data needed to describe the overall population and identify possible differences between subsets in order to assess the homogeneity and representativeness of the overall population.

In case of primary data collection, it is critical that procedures are in place to promote the participation of all individual centres enrolling the population of interest and the inclusion and follow-up of all eligible patients treated in these centres. In order to document possible selection bias and to evaluate generalisability of the study results, eligible patients not recruited in the study or withdrawing from the study could consent in writing to provide a small set of baseline data. This will allow comparing important socio-demographic and clinical characteristics between recruited patients, withdrawn patients and non-recruited eligible patients. The protocol should describe the procedure for consented but not enrolled patients for data collection and handling of personal data.

3.6. Data collection

Registry-based studies may not need the totality of the information collected in the registry to answer the research question. Only the set of data that is needed to ensure the validity and usefulness of the results should be collected or extracted, 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 or extracted. Mechanisms should also be put in place to identify and retrieve initially missing data, if possible.

Data collection should be planned as early as possible, including sensitivity analyses, and should be detailed in the study protocol from early stages on, as collection of additional data for post-hoc analyses may not only be difficult but also prone to additional sources of bias.

Some registry-based studies may require modifications to the existing registry data collection system to address a particular research question, e.g. by adding a specific data collection form or module for additional data collection. The impact of this modification on the legal status of the study should be taken into account as it may require additional informed consent and may impact on the status of the study as clinical trial or non-interventional study, depending on whether the additional data collection is considered part of normal clinical practice. If the data collection system is amended, a validation of the new system should be implemented.

3.7. Data quality management

The nature and extent of the 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 primary or secondary use of registry data.

Risk-based methodologies and measures should be planned. 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 measures if needed. The validity of any data cleaning, extraction and transformation processes

performed centrally should be verified and monitored, especially if it involves mapping of data to a common terminology. The collected information per time interval for the main outcome parameters should be compared to the amount of expected information. Other possible measures include random source data verification, on-site review of processes and computerised systems used for data collection and management, and internal or external audit of the registry-based study. The European Commission's Risk proportionate approaches in clinical trials (29), the EMA Reflection paper on risk-based quality management in clinical trials (30) and the GVP Module III on Pharmacovigilance inspections (31) should be consulted on these aspects.

The thresholds of data quality measures, the level of data verification and the measures to be taken in case relevant findings are observed should be agreed upfront with the registry holders. This information should be included in the study protocol.

3.8. Data analysis

The analytical approach applied for the outcomes of interest should be pre-specified in the registry-based study protocol and statistical analysis plan 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 statistical analysis plan. All changes should be presented, explained and discussed in the study report.

The ICH E9 (R1) addendum (27) 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.

The data analysis should include an evaluation of the representativeness of the study population in relation to the pre-defined target population, as it influences the external validity of the registry-based study (see also chapter A.2 of the Annex). In particular, a comparison between eligible registry patients that were recruited, withdrawn and not recruited, or 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 available electronic health care databases or other population-based data sources.

The handling of missing data should be carefully described in the study protocol and, if applicable, in the statistical analysis plan, and a thorough justification should be provided for the assumptions about their distribution, causes and timing. The ICH E9 (R1) addendum (27), the EMA Guideline on Missing Data in Confirmatory Clinical Trials (32) and the ENCePP Guide on Methodological Standards in Pharmacoepidemiology (26) provide useful guidance on how to handle missing data. It will be necessary to investigate the robustness of the results through appropriate sensitivity analyses that make different, clinically plausible, assumptions.

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

- Measurement of the incidence of outcomes of interest should clearly distinguish between the number of events and the number of individuals presenting at least one event. Comparisons between groups should take person-time of observation into account.
- In the absence of randomised treatment allocation, it must be recognised that the characteristics of patient groups given different treatments are likely to differ. Treatment decisions may be influenced by different factors that may also be associated with the risk of occurrence of the outcome of interest, such as disease severity, or with the monitoring practice of patients (ascertainment bias due to different monitoring requirements of treatments). Even though

methods to adjust for confounding factors may account for underlying differences in clinical outcomes, it must be acknowledged that such confounding adjustment may not be comprehensive and appropriate sensitivity analyses should be considered. In addition, ascertainment of causes for changes of treatments may require complete collection of such information over the course of the study and adjustment only for baseline covariates may not fully address this if the observation expands over several years.

- Registries offer the opportunity to compare patients receiving a treatment of interest with patients who are untreated or who have received different therapies over a long period of time. Inclusion of prevalent drug users (i.e. patients already treated for some time before study follow-up begins) can introduce two types of bias. Firstly, prevalent drug users are "survivors" of the early period of treatment, which can introduce substantial (selection) bias if risk varies with time (for example, if treatments carry a risk of hypersensitivity reactions or affect cardiovascular risk). Secondly, covariates relevant for drug use at study entry (e.g. disease severity) may be affected by previous drug utilisation or patients may differ regarding health-related behaviours (healthy user effect). A new-user design reduces these biases by restricting the analysis to incident drug users, i.e. patients enter the study cohort only at the start of the first course of the treatment of interest during the study period. Consequences of a new-user design may include reduced precision of estimates due to lower sample size and likely reduction in the number of patients with long-term exposure.
- When the follow-up period starts before initiation of the treatment under study, immortal time bias can arise due to misclassification of the non-exposed study period, as the period between start of follow-up and date of first exposure to the drug of interest is event-free by definition when investigating a drug-specific effect. A time-dependent definition of exposure is needed to correctly classify the immortal person-time and causes for changes of exposure need to be taken into account.
- Time-related bias and information bias may also occur in a comparison to a historic control group, i.e. to data collected at earlier time points. The landscape 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.
- In case no other adequate data source is available, some analyses may use a comparative non-exposed control group from outside the registry, for example from another registry or electronic health care records in a country/region where the drug 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. Moreover, one should also strive to correctly define a comparable index date of entry into the study in both groups to correctly account for exposure periods to different drugs and account for determinants of exposure to these different drugs. 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.

3.9. Data reporting

The methods used in the study should be published in sufficient detail 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 Chapter 8 of the ENCePP Guide on Methodological standards in Pharmacoepidemiology (26).

National and European Union obligations and requirements for the registration of studies and the publication of study results (clinical trials and non-interventional studies) should be followed. Post-authorisation registry-based studies should be registered in the EU PAS Register (20) and the study protocol, the statistical analysis plan if applicable and the final study report should be included. For post-registration registry-based clinical trials, the results should be presented in line with clinical trial legislation requirements. The EMA policy on publication of clinical data for medicinal products for human use should also be followed (33). The final reports must contain all study results, whether favourable or unfavourable.

For non-interventional studies, the principles of scientific independence and transparency for reporting of study results described in the ENCePP Code of Conduct (34) and the ADVANCE Code of Conduct for vaccines (35) 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 or investigator. However, where legal requirements apply to an MAH who has contracted a study externally, the 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 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 regulatory authority may request additional information or clarifications from the MAH or may initiate a regulatory inspection. Therefore, the research contract should foresee a duty for the registry holder to address the scientific aspects of the request, with the possibility for the MAH to provide comments, as well as a duty to allow a possible regulatory inspection of the registry-based study.

4. Legal basis 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, with reference to relevant legislation and guidelines.

Activities	Requirements	Legal basis
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 and regulators where applicable.	For a clinical trial: Directive 2001/20/EC, Regulation (EU) No 536/2014, guidance in Volume 10 of The Rules Governing Medicinal Products in the European Union, the Guideline for good clinical practice (GCP, ICH E6), the General considerations for clinical trials (ICH E8), the Statistical Principles for Clinical Trials (ICH E9), the Scientific Guidance on Post-Authorisation Efficacy Studies and the Guidance in Post-Authorisation Efficacy Studies: Questions and Answers on PAES (36), the GVP Module VIII on PASS; For an non-interventional study – prior verification that the study is considered as non-interventional by checking the table in Annex I of the Questions & Answers, version 11.0 (MAY 2013) published by the European Commission,

		Directive 2001/83/EC and Regulation (EC) No 726/2004, Implementing Regulation No 512/2012; GVP Module VIII on PASS, the Scientific Guidance on Post-Authorisation Efficacy Studies and the Guidance in Post-Authorisation Efficacy Studies: Questions and Answers on PAES; Others to consider: GVP Module III - Pharmacovigilance inspections, GDPR, ICH E9, ENCePP Checklist for Study Protocols, ENCePP Code of Conduct and the ADVANCE Code of Conduct for vaccines, REQueST Tool, national requirements in the country of conduct. Relevant national legislation.
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 Agency for scientific advice on the most suitable methods and study designs to generate robust evidence for the development or maintenance of a medicine. Scientific advice in parallel with consultations from another regulatory authority or a health technology assessment (HTA) body is facilitated through EMA procedures.	EMA Scientific Advice and Protocol Assistance: Regulatory and Procedural Guidance, Questions and Answers. National scientific advice procedures.
Safety monitoring and reporting of adverse events and suspected adverse reactions	For registry-based studies initiated, managed or financed by a MAA/MAH, appropriate activities include: • **Individual case safety reports (ICSR) – GVP VI**: See Appendix 3 providing an overview of requirements for ICSRs arising from use of registries in the EU outside the context of a clinical trial; • **Study reports – GVP Modules VI and VIII:** All adverse events/adverse reactions collected in studies should be recorded and summarised in the interim and final study report unless the study protocol provides for different reporting with a due justification; • **Emerging Safety Issues (ESIs) – GVP IX:** should be notified as soon as possible and no later than 3 working days in writing to the competent authority(-ies) of Member State(s) where the medicinal product is authorised and to the EMA. Information affecting the risk-benefit balance of the medicinal product may include an analysis of suspected	For a clinical trial: Directive 2001/20/EC, Regulation (EU) No 536/2014, guidance in Volume 10 of The Rules Governing Medicinal Products in the European Union; GVP Module VIII; For a non-interventional study: Directive 2001/83/EC and Regulation (EC) No 726/2004, Scientific Guidance on Post-Authorisation Efficacy Studies and the guidance in Post-Authorisation Efficacy Studies: Questions and Answers on PAES; GVP Modules VI, VII, VIII and IX.

	adverse reactions and aggregated data; • **Periodic safety update report (PSURs) – GVP VII** Safety information to be summarised in PSURs and other periodic and regulatory reports.	
Transparency, registration of PASS and PAES	• **PASS or PAES that are clinical trials** must be registered in the EU Clinical Trial Register with their protocol and summary of results and the related provisions need to be followed.	Website of the EU Clinical Trial Register;
	• **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.	Commission Implementation Regulation (EU) 520/2012 Annex III; GVP Module VIII; EU PAS Register website;
	• **Non-interventional PAES**: (initiated, managed or financed by an MAH) should be registered in the EU PAS Register, independently from whether they are imposed or not.	EMA Scientific Guidance on Post-Authorisation Efficacy Studies; Post-Authorisation Efficacy Studies: Questions and Answers EU PAS Register website
	• All other post-authorisation PASS/PAES that are not initiated, managed or financed by an MAH are encouraged to be registered in the EU PAS Register.	National requirements.
Record keeping	• **For all PASS and PAES**: Pharmacovigilance data and documents relating to individual authorised medicinal products shall be retained as long as the product is authorised and for at least 10 years after the marketing authorisation has ceased to exist. The documents shall be retained for a longer period where EU or national law so requires. This applies even when the MAA/MAH is not involved in the registry-based study.	Commission Implementation Regulation (EU) 520/2012 Articles 12 and 36; GVP VIII on PASS Scientific Guidance on Post-Authorisation Efficacy Studies; Guidance on Post-Authorisation Efficacy Studies: Questions and Answers ;
	• **For imposed PASS and PAES**: the MAH shall ensure that all pharmacovigilance information as well as the analytical dataset and statistical programmes used for generating the data included in the final study report are kept in electronic format and are available for auditing and inspection.	Commission Implementation Regulation (EU) 520/2012 Articles 12, 36 ;
	• **For PASS and PAES that are clinical trials**: the record keeping requirements in Volume 10 of The	Volume 10 of The Rules Governing Medicinal Products in the European Union.

	Rules Governing Medicinal Products in the European Union apply.	
Personal data protection	• The MAA/MAH and any other organisation involved in the collection, management, use and storage of data from registries must follow the national legislation on personal data protection and the EU General Data Protection Regulation (GDPR). • Informed consent should be broad enough to cover all potential uses of registry data in line with the applicable legislation, including the option for data sharing/ pooling between registries and with other stakeholders including competent authorities and MAAs/MAHs • It is recommended to contact the data protection authorities (DPAs) (37) of the Member States who are competent for monitoring and enforcing the application of the GDPR and other national data protection legislation that may be applicable in their territories.	General Data Protection Regulation (EU) 2016/679; Recital 26; Article 5(1)(e); Article 89(1), national requirements as applicable.
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 MAH and the other organisation should be detailed, up-to-date and clearly describe the responsibilities of each party. Where the 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.	Commission Implementing Regulation (EU) No 520/2012 Article 6(1), 6(2) and 11(2); GVP Module I.
Information of 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.	Commission Implementing Regulation (EU) No 520/2012 Article 10(2), GVP Module I.

430

Annex : Considerations on patient registries

A.1. Introduction

This Annex reviews aspects of good regulatory practice in the establishment and ongoing management of patient registries considered relevant to their use for registry-based studies and other possible regulatory purposes. There are many other factors influencing the suitability of a particular registry for regulatory purposes, such as the size of the target patient population and the patterns of utilisation of a medicinal product in the population covered by the registry (12). For clinical trials, i.e. including randomisation or additional monitoring compared to normal clinical practice, the clinical trial legal requirements should be met by the registries.

A.2. Registry population

The data generated from a patient registry should be representative of the target population of the product. Ideally, the registry should cover a broad patient population covering all disease aspects and patient characteristics. Selection bias can affect the validity of the data derived from the registry and can occur at the level of site selection (i.e. if sites 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 can be considered prior to the enrolment of a registry population:

1. To clearly define the purpose of the registry and the corresponding target population.
2. To translate the target 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 description of the target population definition. This should include prospective enrolment of all newly eligible patients fulfilling the definition and enrolment of already eligible patients by other methods ensuring representativeness and avoiding selection bias, for example by using any 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 best 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. Sensitivity analyses on the effects of incomplete follow-up might 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, 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' representatives and experts of the disease as well as regulators, HTA bodies 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) (38). Definitions, lag times for data availability and data dictionaries should be included in the registry documentation and published or made available on request 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 coordinated registry network. 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 and the capacity of centres to collect and report data in routine clinical care. Note that collection of such data elements, e.g. involving additional laboratory tests, could lead to the categorisation of a registry-based study as a clinical trial.

The dataset should ideally contain the core data elements listed below. This list should be adapted to each situation, for example as regards data elements that remain fixed 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, gender, lifestyle factors (smoking, alcohol);
- 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), dose, route, schedule;
- Relevant concomitant therapies: substance, brand name, indication, start and end dates, dose, route, schedule;
- Safety recording and reporting: adverse events of special interest (AESI), serious adverse reactions (SARs) and suspected unexpected serious adverse reactions (SUSARs); selection of AESI that will be collected should be based and motivated by the previous clinical safety experience with this study population/condition and/or this medication;
- Pregnancies: pregnancy status, pregnancy outcome;
- Patient-reported outcomes collected in clinical practice;
- Additional core data elements defined in disease-specific regulatory guidelines.

An exact 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 (39) (40)), the reports of EMA workshops on registries for cystic fibrosis (41), multiple sclerosis (42), diseases for which CAR-T cell products are indicated (43), haemophilia (44), and on registries for cancers which therapies are based on the tumours' genetic and molecular features (45). The European Platform on Rare Diseases Registration has developed a "Set of Common Data Elements for Rare Diseases Registration" (9). 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 (46) contains information on disease registries that may be consulted when developing a list of data elements. Appendix I of GVP Module P III (18) 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 coordinated registry network. 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 4 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 (47) and GVP Module I (48) provide a framework for the quality of pharmacovigilance systems 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 they should be regularly assessed and made available to patients, health care professionals and potential users of the registry data to provide assurance 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.

A.4.2. Requirements for data quality

In this context, data quality may include 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 coordinated registry network;
- 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 coordinated registry network; implementation of the same data elements, terminologies, data entry procedures and data control software may not be feasible simultaneously in all centres. If needed, 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 a registry-based study once they have achieved an adequate level of data quality as agreed between the concerned parties according to the objective and data needs of the study.

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 (8). Examples of practical aspects and techniques for addressing data quality in patient registries exist in the medical literature (49).

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 and processes in place within the registry or coordinated registry network. 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 (41), multiple sclerosis registries (42) and CAR T-cell Therapy Registries (43).

A.4.4. Data quality management activities

Quality management can be supported by the activities described below, taking into account that compliance with European Union's and national regulations on data protection and storage should always be ensured. 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 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 performed centrally should be verified and monitored, 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.

- If possible, aggregated registry data should be compared to 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 that may be influenced by their purpose, operating procedures, legal environment or funding sources. 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, are therefore useful. Useful guidance is the ENCePP Code of Conduct (50), which provides principles of scientific independence and transparency for pharmacoepidemiological research, and the ADVANCE Code of Conduct (35), which provides governance principles for collaborative studies. The AHRQ User's Guide on patient registries provides a complementary source of recommendations on the governance of registries (8).

Registry holders should consider the following aspects to ensure 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 published in the ENCePP Resources Databases (46).
- To establish a single contact point within the registry or coordinated registry network 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 process for 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 establish a governance structure for management of requests for collaboration to participate in a coordinated research network or in a registry-based study, including a structure for decision-making on such requests (e.g. independent steering committee, ethics committee, advisory board).
- 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 Chapter 4.6.2 of the ENCePP Guide on Methodological standards in pharmacoepidemiology) (26), resources available in the registry and the contractual agreements proposed.

- To establish a supportive function for ethical and legal aspects of collaborations such as compliance with national legislation or the GDPR regulation and ethical approvals.
- To develop a template for research contracts between the registry and external organisations, in line with those recommended by the ENCePP Code of Conduct (50) or the ADVANCE Code of Conduct (35).

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

There may be situations where registry data can be shared in the format of counts, aggregated data or statistical reports with regulators, MAAs/MAHs, HTA bodies, or other parties for clinical development planning or the evaluation or monitoring of medicinal products. These data may concern:

- disease epidemiology in terms of prevalence, incidence, outcomes, prognostic factors, potential confounding variables for defined outcomes;
- size and characteristics of the target population 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 or co-medications;
- medical device utilisation, with number, types and indications and times for specific implanted products;
- surgical procedures with numbers, types and indications and times for and relevant details for the procedures;
- 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 analysis within the registry or, if allowed by the registry governance, 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 coordinated registry network and the other concerned parties.

References

1. EudraLex - Volume 10 - Clinical trials guidelines. Public Health - European Commission https://ec.europa.eu/health/documents/eudralex/vol-10_en

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. 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

9. E-Rare. JRC EU RD Platform releases the Set of Common Data Elements for RD Registration | ERA-Net E-Rare [Internet]. Available from: http://www.erare.eu/news/jrc-eu-rd-platform-releases-set-common-data-elements-rd-registration

10. 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

11. 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

12. 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.

13. 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

14. 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

15. 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 [cited 2020 May 9]; Available from: https://www.jclinepi.com/article/S0895-4356(16)30350-X/fulltext#secsectitle0010

16. 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

17. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) - Module VIII - Post-authorisation safety studies (Rev 3) [Internet]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf

18. European Medicines Agency. Draft 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/documents/scientific-guideline/draft-guideline-good-pharmacovigilance-practices-product-population-specific-considerations-iii_en.pdf

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

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

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

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

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

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

25. 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

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

27. 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

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

29. 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

30. 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

31. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) - Module III – Pharmacovigilance inspections. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-module-iii-pharmacovigilance-inspections_en.pdf

32. 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

33. 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

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

35. 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/

36. European Medicines Agency. Post-authorisation efficacy studies: questions and answers. Available from: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/post-authorisation-efficacy-studies-questions-answers

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

38. 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

39. 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

40. 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

41. 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

42. 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

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

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

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

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

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

854 48. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) Module I –
855 Pharmacovigilance systems and their quality systems [Internet]. Available from:
856 https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-
857 practices-module-i-pharmacovigilance-systems-their-quality-systems_en.pdf

858 49. Bray F, Parkin D. Evaluation of data quality in the cancer registry: Principles and methods. Eur J
859 Cancer 2009; 45(5):747-64

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

863

864

Appendices

Appendix 1. Glossary

Coordinated registry network

Coordinated network of registries and other linked data sources set up to to support 1) the implementation of structured information, core minimum data elements and definitions, and 2) the ability to share data across registries and other linked data sources.

Disease registry

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

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.

Interventional and Non-interventional study

A non-interventional study on safety and/or effect of medicinal products should follow normal clinical practice in the Member State where it is authorised. All interventions in the study, i.e. treatment, diagnostic or monitoring procedures, should fall within the standard of care, as interpreted by the competent authority/ethics committee in that Member State. Typically, a clinical trial application is required when a study involves additional diagnostic or monitoring procedures compared to normal clinical practice, i.e. if these measures are required in a protocol and not based on patient care decisions taken by the treating physician. Such study-specific interventions that could lead to a change of category from a non-interventional study to a clinical trial include additional patient visits, sampling of biological samples including blood as well as other study-specific burdensome procedures. For long-term follow-up the same principles apply, even if the medicinal product was administered before the long-term follow-up clinical trial starts (Directive 2001/20/EC). For clarity on the difference between non-interventional studies and interventional trials, see the Annex I table, Questions & Answers, 11.0 (May 2013) published by the European Commission (1).

Missing data

Data that would be meaningful for the analysis of a given estimand but were not collected (ICH E9 (R1)) (27).

Such data should be distinguished from data that do not exist or data that are not considered meaningful because of an intercurrent event (ICH E9 (R1)) (27).

Intercurrent events are events that occur after treatment initiation and either preclude observation of the variable or affect its interpretation (ICH E9 (R1)) (27).

The estimand is the target of estimation addressing the scientific question of interest posed by the registry-based study (adapted from ICH E9 (R1) addendum) (27).

Non-available data

Data that has not been collected for the purpose of the registry.

Primary data collection or secondary use of data in the context of a registry-based study

Primary data collection: collection of information on the events of interest for the purpose of the study directly from the patients, caregivers, healthcare professionals or other persons involved in patient care.

Secondary use of data: use of information from an existing registry that has collected data for a purpose other than the specific study.

Quality requirements

Those characteristics of a system that are likely to produce the desired outcome, or quality objectives.

Quality requirements may be fulfilled by implementing processes of quality planning, quality control and assurance and quality improvement (Implementing Regulation 520/2012 Art 8(3)) (47).

- Quality planning: establishing structures and planning integrated and consistent processes
- Quality control and assurance: monitoring and evaluating how effectively the structures and processes have been established and how effectively the processes are being carried out
- Quality improvements: correcting and improving the structures and processes where necessary.

Register, synonym: Registry database

Database derived from the registry.

Registry, synonym: Patient registry

For the purpose of this Guideline: organised system that collects data and information on a group of people defined by a particular disease or condition, and that serves a pre-determined scientific, clinical and/or public health (policy) purpose (definition derived from the PARENT guidelines) (6).

In this guideline, the term 'people' refers to persons with a disease or a condition, and to persons using a medicinal product for treating or preventing a disease, restoring, correcting or modifying physiological functions or making a medical diagnosis.

The use of the term 'patient' in combination with 'registry' (i.e. patient registry) is used to highlight the focus of the dataset on health information. The terms 'people' and 'patients' are used in this definition and guideline as synonyms, independently of the health status of the individual.

Examples of a particular condition are pregnancy, a birth defect, a molecular or a genomic feature, and other specific patient characteristics.

The term "registry" and the epidemiological term "cohort" have different meanings. A registry may lead to the creation of a cohort of patients followed over time.

Registry-based study

Investigation of a research question using the infrastructure of (a) new or existing registry(-ies) for patient recruitment and data collection.

A registry-based study may use the infrastructure of a single registry or coordinated registry network, and it may link data from different registries at individual patient level.

A registry-based study may be a clinical trial, or a non-interventional study as defined in Directive 2001/20/EC or Regulation (EU) No 536/2014 when it becomes applicable (1). Post-authorisation, a registry-based study may be a post-authorisation safety study (PASS), a post-authorisation efficacy study (PAES) or another type of study with other objectives. The clinical trial categorisation above also applies to PASS and PAES.

A registry-based study may apply primary data collection and/or secondary use of data collected through a registry for a purpose other than that of the specific study.

Appendix 2. *Checklist for evaluating the suitability of registries for registry-based studies*

(List adapted from the REQuEST tool published by EUnetHTA) (7)

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. Data privacy

- Status of implementation of GDPR
- Informed consent form and its validity for registry-based studies (or need for re-consent)

1.3. Funding

- Funding sources and impact on short, long-term sustainability and possible conflicts of interest 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

- Adequate 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

- Adequate description of the type of patient registry (disease, condition, time period covered, procedure), which defines the criteria for patient eligibility
- Relevance of setting and catchment area
- Clarity on 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 data elements

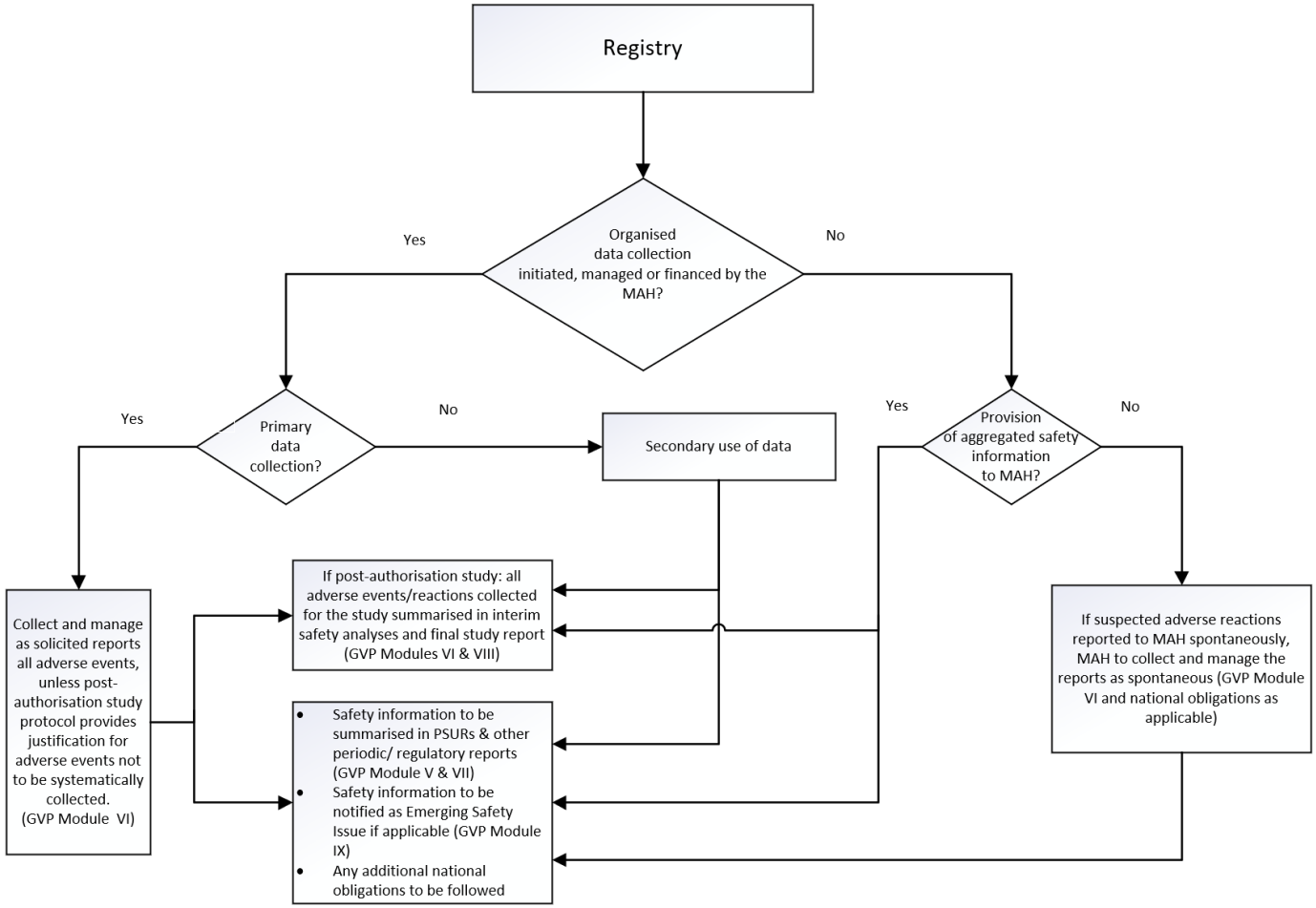
2.5. Infrastructure

- Systems for data collection, recording and reporting, including timelines
- Capability (and experience) for expedited reporting and evaluation (at physician or registry level) of severe suspected 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
- 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 3. *Overview of MAH responsibilities for individual case safety reports (ICSRs) where a registry-based study fulfils the definition of a non-interventional study according to the clinical trial legislation.*



Appendix 4. 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 ICD-o-3 (cancers) SNOMED-CT	http://www.who.int/classifications/icd/en/ http://codes.iarc.fr/ https://www.snomed.org/snomed-ct
Rare disorders (disease, malformation syndrome, clinical syndrome, morphological or biological anomaly or particular clinical situation in the course of a disorder)	Orphadata (entries are cross-referenced with ICD-11, OMIM, UMLS, MeSH, MedDRA) HPO (Human Phenotype Ontology)	http://www.orphadata.org/cgi-bin/inc/product1.inc.php https://hpo.jax.org/app/
Medicinal products	Article 57 database (EEA) ISO IDMP standards and related terminologies ATC classification RxNorm (US)	EMA website[2] EMA website[3] https://www.whocc.no/atc_ddd_index/ https://www.nlm.nih.gov/research/umls/rxnorm/index.html
AESI, other adverse events, suspected adverse reactions	MedDRA CTCAE (cancer therapies; includes MedDRA)	https://www.meddra.org/ https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm

Data elements	Terminologies	Weblinks
Routes of administration, pharmaceutical dosage forms, packaging, units of administration	EDQM Standard Terms Database	https://www.edqm.eu/en/standard-terms-database
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/ https://www.genenames.org/
Classification of functioning/disability	ECOG ICF	https://www.mdcalc.com/eastern-cooperative-oncology-group-ecog-performance-status http://www.who.int/classifications/icf/whodasii/en/
Terminologies and Formats for individual case safety reports (ICSR) specification	Code Sets and Object Identifiers based on the ICH E2BR(3) ICSR Implementation Guide	http://estri.ich.org/e2br3/index.htm https://www.hma.eu/fileadmin/dateien/Human_Medicines/01-About_HMA/Working_Groups/CTFG/2017_11_CTFG_Question_and_Answer_on_Reference_Safety_Information_2017.pdf https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/regulation5362014_qa_en.pdf

[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, risk-benefit evaluation and assessment, electronic exchange and communication of pharmacovigilance and medicinal product information.

[2] http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_001940.jsp&mid=WC0b01ac0580dd91db

[3] http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_001849.jsp&mid=WC0b01ac0580bf85bb