5 November 2018

EMA/763513/2018

**Discussion paper:**

**Use of patient disease registries for regulatory purposes – methodological and operational considerations**

The Cross-Committee Task Force on Patient Registries

**Foreword**

This discussion paper has been prepared by the Cross-Committee Task Force on Registries established by the EMA Patient Registries Initiative. The main objective of this initiative is to facilitate use of patient registries to support regulatory decision-making.

The paper discusses methodological and operational aspects of the use of patient disease registries and registry studies for regulatory purposes. It has been published to seek comments and suggestions from all the interested parties. All the responses will be considered for the finalisation of the document with the EMA Committees in Q4 2019.

Please send your comments and suggestions **before 30 June 2019** by sending the [Form for submission of comments](https://www.ema.europa.eu/documents/other/form-submission-comments_en.doc) or an annotated version of the document (mentioning on the first page your name, affiliation and contact details) to: EMAregistries@ema.europa.eu.

Thank you.

Table of Contents

[Abbreviations 5](#_Toc528934916)

[Executive summary 6](#_Toc528934917)

[1. Introduction 10](#_Toc528934918)

[2. Objective 11](#_Toc528934919)

[3. Core concepts 12](#_Toc528934920)

[3.1. Definitions 12](#_Toc528934921)

[3.2. Disease registries and product registries 12](#_Toc528934922)

[3.3. Registry and registry study 13](#_Toc528934923)

[4. Use of patient disease registries in medicines regulation 16](#_Toc528934924)

[5. Good registry practice 18](#_Toc528934925)

[5.1. General consideration 18](#_Toc528934926)

[5.2. Patient population 18](#_Toc528934927)

[5.3. Time elements 19](#_Toc528934928)

[5.4. Core data elements 20](#_Toc528934929)

[5.5. Terminologies 21](#_Toc528934930)

[5.6. Quality management 23](#_Toc528934931)

[5.6.1. Importance of quality management 23](#_Toc528934932)

[5.6.2. Requirements of data quality 24](#_Toc528934933)

[5.6.3. Measures to improve data quality 25](#_Toc528934934)

[5.6.3.1. Data quality activities at management level 25](#_Toc528934935)

[5.6.3.2. Data quality activities at operational level 25](#_Toc528934936)

[5.6.4. Indicators of data quality 26](#_Toc528934937)

[5.7. Safety analysis 28](#_Toc528934938)

[5.7.1. Reporting of safety information 28](#_Toc528934939)

[5.7.2. Monitoring of adverse events of special interests 28](#_Toc528934940)

[5.7.3. Aggregate analysis of adverse events 28](#_Toc528934941)

[5.8. Governance 29](#_Toc528934942)

[5.8.1. Governance principles 29](#_Toc528934943)

[5.8.1.1. Role of registry coordinators 29](#_Toc528934944)

[5.8.1.2. Role of pharmaceutical companies 30](#_Toc528934945)

[5.8.1.3. Role of regulatory authorities 30](#_Toc528934946)

[5.8.2. Data ownership and intellectual property 31](#_Toc528934947)

[5.8.3. Data sharing 31](#_Toc528934948)

[5.8.3.1. General considerations 31](#_Toc528934949)

[5.8.3.2. Data sharing with MAAs/MAHs 32](#_Toc528934950)

[5.8.3.3. Data sharing with regulators 33](#_Toc528934951)

[5.8.4. Informed consent 33](#_Toc528934952)

[5.8.5. Data security 34](#_Toc528934953)

[6. Registry studies 35](#_Toc528934954)

[6.1. Regulatory context 35](#_Toc528934955)

[6.2. Timelines 36](#_Toc528934956)

[6.3. Study protocol 36](#_Toc528934957)

[6.4. Study population 37](#_Toc528934958)

[6.5. Data collection 37](#_Toc528934959)

[6.6. Data quality 38](#_Toc528934960)

[6.7. Data analysis 39](#_Toc528934961)

[6.8. Safety reporting 40](#_Toc528934962)

[6.9. Reporting of study results 42](#_Toc528934963)

[7. Conclusions 43](#_Toc528934964)

[8. References 44](#_Toc528934965)

Abbreviations

ADVANCE: Accelerated Development of VAccine benefit-risk Collaboration in Europe

AESI: Adverse Event of Special Interest

AHRQ: US Agency for Heathcare Research and Quality

CAR-T cell: Chimeric Antigen Receptor T cell

CHMP: Committee for Medicinal Products for Human Use

CIBMTR: Center for International Blood and Marrow Transplant Research

EBMT: European Bone Marrow Transplantation

ECFSPR: European Cystic Fibrosis Society Patient Registry

EEA: European Economic Area

EMA: European Medicines Agency

ENCePP: European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

EU: European Union

GDPR: General Data Protection Regulation

GVP: Good pharmacovigilance practice

HCP: health care professional

HTA: Health Technology Assessment

ICH: International Council for Harmonisation

ICSR: Individual Case Safety Report

MAA: Marketing Authorisation Applicant

MAH: Marketing Authorisation Holder

MedDRA: Medical Dictionary for Regulatory Activities

NCA: national competent authority

PAES: Post-Authorisation Efficacy Study

PARENT JA: PAtient REgistries iNiTiative Joint Action

PAS: Post-Authorisation Study

PASS: Post-Authorisation Safety Study

PRAC: Pharmacovigilance Risk Assessment Committee

PRO: Patient-Related Outcome

PSUR: Periodic Safety Update Report

RCT: Randomised Clinical Trial

RMP: Risk Management Plan

SAP: Statistical Analysis Plan

Executive summary

The objective of this paper is to discuss methodological and operational aspects of patient disease registries. It focuses on important principles from a regulatory perspective and makes proposals on what might be considered good registry practice to support collection of high quality registry data. This paper also discusses important methodological aspects of studies performed with registry data to increase regulators’ confidence in their results.

A *patient registry* is defined as an organised system that uses observational methods to collect uniform data on a patient population defined by a particular disease, exposure or condition (e.g. age, pregnancy, specific patient characteristics), and which is followed over time. Patient *disease* ***registries* may be** established by public organisations such as academia or medical research associations of health care professionals or patients. They may have different overall objectives, such as to describe the natural history of a disorder, to monitor the efficacy or safety of treatments, to describe the impact of a disease on patients’ health and quality of life or to identify patients suitable for new treatments. The data collected may also be useful to support the regulatory evaluation of benefits and risks of medicines. Although *product registries* may have advantages in specific circumstances, e.g. to gather data in a subset of treated patients (such as geriatric patients) underrepresented in clinical trials, regulators generally prefer disease registries as they gather insights on clinical outcomes in patients receiving different treatments and may support a wider range of study designs, e.g. controlled designs without an external data source. The main focus of this document is therefore on patient disease registries even if many considerations also apply to product registries or registries of patients defined by a specific condition.

The data collected in a registry may be used for the purpose of a specific study, i.e. a detailed investigation of a research question or hypothesis. The differences between a registry and a registry-based study (a *registry study*) are highlighted as they must be well understood to avoid confusion in concepts and methods. This discussion paper therefore provides different sets of considerations for patient disease registries (“good registry practice”) and for registry studies.

For **patient disease registries**, the following considerations are emphasised:

* *Patient population*: great care should be exercised to ensure exhaustive enrolment of patients and avoid selection bias, i.e. the situation where enrolment is influenced by patient characteristics that may affect the validity of the analyses. Registries are prone to selection bias because numerous factors may influence the enrolment of patients in a registry and may be difficult to identify in advance and prevent. Four steps are proposed to be considered in the definition and enrolment of a study population.
* *Time elements*: **accurate knowledge and recording of dates of important events are essential components of all registries. A list of core time elements that should normally be collected in all registries is proposed. Definitions should be standardised across different registries.**
* *Core data elements*: a list of core data elements to be collected in all patients is proposed. They should be harmonised or mapped across registries for a same disease to support regulatory evaluations and facilitate implementation of a common data quality system, data exchange, common data analysis and interpretation of results from different registries.
* *Terminologies*: **common terminologies should be used across registries for diseases, diagnostic tests, symptoms, medicinal products, active substances, adverse events and other relevant data. Where local or national terminologies are used, these should be mapped to international terminologies. A list of examples of recommended international terminologies for different data elements is provided.**
* *Quality management*: quality planning, quality assurance, quality control and quality improvement are activities of quality management that should be continuously in place and considered as part of the management of registries. Data quality should address their consistency, completeness, accuracy and timeliness. Measures to improve data quality are described. They may be performed at management level and at operational level. Indicators of data quality are useful tools to measure and improve data quality.
* *Safety analysis*: disease registries conducted by organisations such as academia or medical research associations should follow the national requirements for the management of safety data. Reporting of suspected adverse reactions through the national or regional pharmacovigilance system should be encouraged. Any active data collection system put in place in a disease registry and initiated, managed or funded by a MAH to collect and record suspected adverse reactions to one of its medicinal products must follow the regulatory framework for PASS. Disease registries are generally not suitable for a rapid statistical analysis of new safety signals but they may be useful for the monitoring and characterisation of known or suspected adverse reactions. A list of adverse events of special interest (AESI) can be defined and integrated in the routine data collection system.
* *Governance*: Most registries have a governance model relying on principles and constraints based on their mandate, operating procedures, legal environment or funding sources. Effective collaboration between all parties is needed to ensure early identification of registries relevant for a regulatory procedure and evaluation of the adequacy and quality of the data collected. A number of activities that may be performed by registries, MAHs and regulators are proposed to strengthen the use of registry data. Amongst them is the agreement on principles of data sharing between registries, MAAs/MAHs and regulators. Principles of data ownership, informed consent and data security should be applied in accordance with the General Data Protection Regulation (GDPR).

For **registry studies**, the following aspects are particularly relevant:

* *Regulatory context*: post-authorisation safety (PASS) and efficacy (PAES) studies may be required by regulators to MAHs. For studies imposed as a legal obligation, the MAH holds the responsibility for their supervision, including the obligations to monitor the data generated, to consider its implications for the benefit-risk balance of the medicinal product and to communicate to authorities any new information which might influence this balance. The MAH must also ensure that the fulfilment of these obligations can be audited, inspected and verified, and register the study in a public database. These legal constraints should be taken into account by registry coordinators in their discussion with MAHs.
* *Timelines*: given the time needed to initiate a registry study, preparations and discussions about use of a patient registry post-authorisation should start before –during scientific advice- or at an early stage of the approval application process. It is recommended that such early discussions involve the MAH(s), regulators and registry coordinators. The study proposal should be sufficiently detailed to be used by registries and individual centres to assess whether they can participate in the study in terms of data availability and data quality requirements.
* *Study protocol*: an early decision to be made is the choice of the data collection method: secondary data collection, where the data for the study are already available and extracted from a dataset, or primary data collection, where the events of interest for the study are collected directly from patients as they come to the attention of the investigator. This choice has implications for safety reporting and should be clearly specified in the study protocol.

The study protocol should follow the recommendations of the Good pharmacovigilance practice (GVP) Module VIII and the technical guidance on the format and content of the protocol for non-interventional PASS. For studies addressing several products, all concerned MAHs should participate in a joint registry study based on a single protocol. If a registry study is to be conducted across multiple sites, a common protocol needs to be developed based on core common data elements and common design even if some aspects of the study may vary according to the characteristics of each registry.

The protocol should provide an estimation of the study size and the feasibility of attaining this sample size within the registry should be assessed using conservative assumptions (or a previous feasibility analysis), both in terms of number of patients and duration of follow-up. Milestones and timelines for completing the main study phases should be provided.

* *Study population*: where study data are extracted from data already collected routinely in the registry, methodological challenges to define the study population are comparable to those met in secondary data collection. When the safety or effectiveness of a new treatment need to be monitored in a disease registry, the study population will most often include new users who are either patients newly diagnosed with the disease and who received a first prescription (incident patients) or patients already included in the registry to whom the new treatment is prescribed (prevalent patients). The choice of including incident or prevalent patients in the study population has implications for the data analysis and the interpretation of the results. It is recommended to collect the data needed to distinguish incident and prevalent patients and identify possible differences in their characteristics. In case of prospective recruitment, it is critical that procedures are in place to ensure sequential inclusion of all eligible patients treated in the individual centres. Patients having been involved in clinical trials, often representing a subgroup with certain disease characteristics (e.g. genetic variants), could also be enrolled in a disease registry later on.
* *Data collection*: it is the investigators’ responsibility to collect for the study only the sets of data that are strictly needed to provide valid results. It is also their responsibility to collect all the data needed for this purpose including available data on potential confounders. Data to be collected include those needed for sensitivity analyses as outlined in the study protocol and statistical analysis plan (SAP). The legal status of a study with additional data collection should also be considered. Additional data collection may turn it into an interventional study and its relation to current clinical practice needs to be detailed in the study protocol.
* *Data quality*: the need for additional measures for data quality control in a registry study depends on the measures already applied routinely in the registry. In order to ensure acceptable data quality for individual registry studies conducted for regulatory purposes, source data verification and periodic auditing on a reasonable amount of data may need to be conducted on a risk analysis-based approach and following a strategy dependent on the scope of the study. Data source verification for a minimum of 10% of randomly selected patients registered in individual study centres would be considered adequate. Data quality measures also include quality checks at data entry and monitoring of patient follow-up.
* *Data analysis*: there is no single statistical method applicable to every study and the most appropriate statistical procedure would need to be selected on a case-by-case basis to specifically address the scientific question of interest. Analysis of registry data should take into account that they are selected observational cohorts and hence knowledge of techniques developed to analyse such data is essential. Commonly encountered methodological problems and how they could be approached are highlighted. The handling of missing data should be described in the SAP and justification should be provided for the assumption about their distribution.
* *Safety reporting*: a distinction should be made between studies with a design based on primary data collection directly from health care professionals (HCPs) or consumers and studies with a design based on secondary use of data; the recommendations from Module VI of the GVP should be followed. In case of primary data collection, the concerned MAH(s) should have a data collection or electronic system to collect, analyse and report information on adverse events notified by a HCP or a patient. Investigators of registry studies should also be informed of the mechanisms allowing them to report at any time to the national spontaneous reporting system any adverse event or suspected adverse reaction occurring during the course of the study.
* *Reporting of study results*: for imposed PASS and PAES, legal obligations apply to the final study report to be submitted by the MAH to regulatory authorities. MAHs should therefore be able to comment on the study results and their interpretation as well as on the format of the report. Following the submission of the final study report to the regulatory authority, additional information or clarifications may be requested to the MAH by the regulatory authority and the lead investigator should have the duty and scientific responsibility to address the scientific aspects of the request. The lead investigator should always have the right to independently prepare publications of the study results irrespective of the source of funding.

The EMA is willing to support interactions between MAAs/MAHs and registries and provide tools to facilitate recognition of disease registries as data sources to conduct studies for regulatory purposes. Early dialogue with regulators is encouraged. The Scientific Advice procedure on study protocols and the Qualification procedures of registries are available to pharmaceutical companies and registry coordinators to provide advice and opinions on the validity of registries and study protocols.

1. Introduction

The European Medicines Agency (EMA) is responsible for the scientific evaluation, supervision and safety monitoring of medicines in the European Union (EU). Together with the EU regulatory network, the EMA's scientific committees provide an independent and comprehensive scientific evaluation of data submitted by companies in marketing authorisation applications, and their recommendations provide the basis for the authorisation of medicines in Europe through the centralised procedure. Once medicines have been authorised, the EMA continuously monitors and supervises their safety and effectiveness to ensure that their benefits continues to outweigh their risks (1).

Patient registries are frequently used as a source of data for post-authorisation monitoring of medicinal products. An analysis of 335 centrally approved products from 2005-2013 found that for 30 products (9.0%), a *registry* has been requested by the EMA Committee for Medicinal Products for Human Use (CHMP) as a condition of the marketing authorisation (2). Another review of 116 new drugs approved by the CHMP from 2007 to 2010 found that at least one registry was included in the risk management plan (RMP) of 43 (37%) products and, out of these, it was imposed as an obligation for 9 (20.9%) products (3). These reviews have found that registries requested by the EMA are often created by companies as new *product registries* where inclusion is determined by exposure to a specific medicinal product. However, many patient *disease registries,* where inclusion is determined by occurrence of a specific disease, exist in Europe. They include a wealth of clinical information that has not been extensively used by regulatory authorities for the benefit-risk evaluation of medicines.

The EMA established in 2015 the Patient Registry Initiative and the Cross-Committee Task Force on registries (4) to explore ways of expanding the use of disease registries by supporting a more systematic and standardised approach to their contribution to the benefit-risk evaluation of medicines. Discussions with registry coordinators, patient associations and pharmaceutical companies during the pilot phase of the Patient Registry Initiative **helped identify barriers to the use of disease** registries for regulatory purposes. They included factors such as heterogeneity in the data collected by different registries on a same disease, inconsistent quality assurance in registry processes and verification of data, restricted access, obstacles to data sharing, lack of transparency and lack of sustainability.

1. Objective

The objective of this paper is to discuss methodological and operational aspects of patient disease registries. It focuses on important principles from a regulatory perspective and makes proposals on what might be considered good registry practice to support the collection of high quality registry data. This paper also discusses important methodological aspects of studies performed with registry data to increase regulators’ confidence in their results.

The main focus of this document is on patient disease registries, but many of the considerations presented also apply to patient product registries or to registries of patients defined by specific conditions such as age, pregnant status or other patient characteristics.

This discussion paper reflects recommendations arising from the EMA Patient Registries Workshop (5), from the four product-specific workshops on registries for cystic fibrosis (6), multiple sclerosis (7), CAR-T cell products (8) and haemophilia products (9), from the Qualification opinion on the European Cystic Fibrosis Society Patient Registry (ECFSPR) (10), from the Draft qualification opinion on the Cellular therapy module of the European Society for Blood & Marrow Transplantation (EBMT) Registry (11), and from existing guidance published in the PARENT Joint Action Methodological Guidance (12) and the US Agency for Healthcare Research and Quality (AHRQ)’s handbook (13). The latter documents provide practical information on the design, operation and analysis of patient registries. This information will not be repeated here but some aspects important for regulators are emphasised. The document also integrates some recommendations from the ENCePP Guide on Methodological Standards in Pharmacoepidemiology (14) and the ENCePP Code of Conduct (15).

1. Core concepts
	1. Definitions

This section provides definitions for specific terms used in the text.

* *Patient registry:* **organised system that uses observational methods to collect uniform data on a patient population defined by a particular disease, exposure or condition (e.g. age, pregnancy, specific patient characteristics), and which is followed over time.**

**It is the occurrence of an event that determines the entry of a patient into a registry. This event may be the diagnosis of a particular disease, the start of a treatment with a particular drug or any other defining condition, for example the beginning of a pregnancy. Some registries also include** family members with the disease or who may have an elevated genetic risk of developing the disease.

**The term *register* generally designates the database derived from the data collection system (i.e. the registry). As the terms registry and register are often used interchangeably, the term registry is used throughout this text.**

* ***Incident patients:* patients whose entry in the registry is determined by the occurrence of a new event (a new diagnosis, a new initiation of a medication or a new condition).**
* ***Prevalent patients*: patients who have already experienced the event in the past and who enter the registry based on their first visit to a clinician participating in the registry or fulfilling any other condition for eligibility.**

**For prevalent patients, it is important to ensure that the date of entry in the registry is properly recorded; the data of entry should be clearly distinguished from other relevant dates e.g. the date of the diagnosis of a disease or of a first prescription.**

* ***Registry participant*: health care professional or other person recruiting patients, updating medical records, ensuring data entry, managing the data or performing other tasks at the local level.**
* ***Registry coordinator*: person or entity having a role in the overall coordination of a registry or of a platform of several registries.**
* ***Population registry:*** comprehensive registration of population data at the regional or national level based on a unique identifier allowing linkage between datasets, and which is called *registry* in European Nordic countries, such as the Norwegian Patient Registry (16), the Danish National Patient Registry (17) or the Swedish National Patient Register (18). Population registries are not addressed in this document.
	1. ****Disease registries and product registries****

**There are many reasons why collection of patient data is initiated. The utility of patient registries for different stakeholder groups is described in section 3.2 of the report of the** Patient Registries Workshop (5)**.**

**Patient disease registries are** often created by public organisations such as academia or medical research associations of health care professionals or patients. They may have different objectives, such as to describe the natural history of a disorder, to monitor the efficacy or safety of treatments, for example in subpopulations (e.g. geriatric patients) for which pre-authorisation data are limited, to describe the impact of a disease on patients’ health and quality of life or to identify patients suitable for new treatments. Irrespective of the original aim pursued by the registry, the data collected may also be useful to support the regulatory evaluation of benefits and risks of medicines. For this reason, regulators generally prefer *disease registries,* i.e. patient registrieswhere entry is determined by the diagnosis of a disease,to *product registries,* i.e. patient registries where entry is determined by the prescription of a specific drug, as disease registries gather insights on clinical outcomes of conditions in patients receiving different treatments, rather than on the outcomes of a specific treatment, and they may support a wider range of study designs, e.g. controlled designs without an external data source. They are also generally better integrated into health care systems and are therefore more likely to be sustainable and provide long-term follow-up data on patients.

* 1. Registry and registry study

The data collected in a registry may be used for the purpose of a specific *study*, i.e. a detailed investigation of a research question or hypothesis. The difference between a *registry* and a registry-based study (a *registry study)* must be well understood to avoid confusion in concepts and methods. It is acknowledged that, as previously highlighted (2) (19), regulators have sometimes requested marketing authorisation holders (MAHs) to establish a *registry*, although the objective was to perform a post-authorisation safety study (PASS) to monitor the safety of a product. Some existing guidance seems also to use the terms “registry” and “study” interchangeably.

The main differences between a registry and a registry study are summarised in Table 1 and include the following aspects:

*1)* *Nature*: a patient registry is a data collection system. A registry study is an investigation set up to answer a research question that uses data collected in the registry, and which may be initiated, managed or financed by a pharmaceutical company, a regulatory authority or another organisation.

*2) Timelines*: a disease registry is a long-term endeavour and in principle it has no end. Timelines are driven by schedules for data collection (schedule of medical consultations or other encounters between patients and registry participants) and for routine data analyses. On the other hand, timelines for a registry study are driven by the time needed to collect or extract the data relevant for the specific study objective and to perform the analyses.

*3) Patient enrolment*: patient enrolment in the registry should be exhaustive within the boundaries defined by the purpose of the registry (for example, all patients diagnosed with a disease in a region), while the study population should be defined in line with the research objective and may be a subset of the registry population.

*4) Data collection*: depending on the purpose of the registry, different types of data can be collected, such as data on demographic characteristics, treatments, diseases, patient-related outcomes or comorbidities. In a study, data collection or extraction is restricted to the data necessary to investigate the research question, including data on potential confounders and effect modifiers. A specific study may also require additional data collection from other sources if these data are not routinely collected in the registry.

*5)* *Analysis plan*: data analysis in a registry is generally performed at intervals based on patient accrual or time schedule and it follows a routine analytical plan with additional ad-hoc analyses performed by the registry coordinator or registry participants. In a study, key elements of the planned statistical analysis should be described in the protocol, supplemented as required with a separate Statistical Analysis Plan (SAP). As with the protocol, deviations to the SAP following the start of the study should be accompanied by a formal amendment process. Registries and registry studies may face different methodological and analytical issues.

*6) Collection and reporting of suspected adverse reactions:* disease registries conducted by organisations such as academia or medical research associations should follow the national requirements as regards the management of safety data. Any active data collection system put in place in a disease registry and initiated, managed or funded by a MAH to collect and record suspected adverse reactions to one of its medicinal products should follow the regulatory framework for PASS (20) (see chapter 6.8). In registries without active safety data collection, health care professionals (HCPs) and patients should be reminded of the possibility to report suspected adverse reactions to the MAH of the suspected medicinal product or to the concerned competent authority via the national spontaneous reporting system. Any noxious or unintended response to a medicinal product notified to the MAH should be managed by the MAH as a spontaneous report of a suspected adverse reaction (21) (see chapter 5.7).

For registry studies, requirements to MAH for suspected adverse reactions differ between studies with a design based on primary or secondary data collection. Any system for collecting, reporting and analysing adverse events put in place should be described in the study protocol (20) (see chapter 6.8).

***7) Data quality control*: measures for data quality control should be routinely in place in a registry and these measures therefore also apply to the data and procedures used in a study. Additional measures may be needed for a specific study, for example to verify the completeness of the information on some data elements.**

***8) Regulatory status*: by nature, a registry is observational, i.e. non-interventional; on the other hand, a registry study may be non-interventional (if data collection is restricted to secondary use of already existing data- see chapter 6.5.) or interventional, for example if additional visits or diagnostic tests are required to validate a diagnostic test in patients identified through the registry, or if additional treatment is given.**

**Table 1. Important differences between a registry and a registry study.**

|  |  |  |
| --- | --- | --- |
|  | **Registry** | **Registry study** |
| **Nature** | **Data collection system** | **Investigation of a research question or hypothesis** |
| **Timelines** | **Long-term, open-ended** | **Defined by the study objective and described in the study protocol** |
| **Patient enrolment** | **Exhaustive within the boundaries of the purpose of the registry (e.g. all patients diagnosed with a disease in a hospital, region or country)** | **Defined by research objective and described in the study protocol- it may be a subset of the registry population.** |
| **Data collection** | **Wide range of data may be collected depending on the purpose of the registry** | **Restricted to what is needed by the research question including data on potential confounders and effect modifiers – additional data collection may be required.** |
| **Analysis plan** | **Routine periodical data analysis; additional ad-hoc analyses** | **Statistical analysis plan separate from the study protocol**  |
| **Collection and reporting of suspected adverse reactions**  | **National requirements as regards the management of safety data apply. Any active data collection with involvement of a MAH must follow the regulatory framework for PASS.** | **National requirements may apply. Regulatory requirements to MAHs differ between studies with primary or secondary data collection.**  |
| **Data quality control** | **Applied routinely to all data and processes** | **Additional quality assurance may be needed.** |
| **Regulatory status** | **Non-interventional**  | **Non-interventional or interventional** |

1. ****Use of patient disease registries in medicines regulation****

Pre-clinical and clinical trials are at the core of the scientific evaluation of medicines before their authorisation, but this evaluation may be supported or supplemented with evidence derived from observational data. These include descriptive data on standards of care for the concerned disease, incidence and determinants of disease outcomes in clinical practice, characteristics of the target population for a new medicine, or validity of any surrogate endpoint used to measure efficacy. In some circumstances (e.g. very rare diseases) where randomised controlled trials (RCTs) are unethical or not feasible, observational evidence may also be used to assess efficacy and the benefit-risk profile of medicinal products.

At the time of authorisation, and post-authorisation during the whole life cycle of the product, regulatory authorities may also request MAHs to collect additional data on the utilisation, benefits and risks of their medicines pursuant to a legal obligation (since June 2012) or in the context of the RMP, either as a PASS or as a post-authorisation efficacy study (PAES). Both may be non-interventional, i.e. they may be based on observational data (20) (22).

Patient disease registries represent an important source of information derived from clinical practice. They are particularly useful when dealing with a rare disease or with a special population affected by a disease, such as older people. The registry can be used as a source of patients based on either primary data collection (where the events of interest for the study are collected directly from patients as they come to the attention of the investigator) or secondary data collection (analogously to the use of electronic healthcare records). For this purpose, registry data can be enriched with additional information from linkage to existing databases such as national cancer registries, prescription databases or mortality records.

For common diseases, exposures or conditions, alternative data sources providing access to large numbers of patients, such as electronic healthcare databases of medical records or claims data, should be rather considered (23).

Patient registries may provide data on patient characteristics, diseases and treatments as well as on their determinants and their outcomes (clinical outcomes or patient-reported outcomes). With adequate quality, they are therefore useful for the following studies:

* drug utilisation studies (DUS), including studies measuring the prevalence or incidence of drug prescription in a population and describing patterns of utilisation (dosage, schedule, duration, switching between treatments, etc.);
* post-authorisation efficacy studies (PAES), to study efficacy/effectiveness in patient sub-groups defined by variables such as age, co-morbidities, use of concomitant medication or genetic factors, and to provide historical control data that could be used for comparative purposes;
* post-authorisation drug safety studies (PASS), to collect safety data on adverse events using standardised data collection tools and amplify a safety signal, particularly for rare outcomes, to assess the incidence of important identified and potential risks, to compare the risk of some adverse events between relevant exposure groups or to assess the effectiveness of risk minimisation measures. Registries may be particularly valuable when examining the safety of medicinal product used for an orphan disease.

Incorporating data from clinical practice into the drug development process is also of growing interest for health technology assessment (HTA) bodies and payers since reimbursement decisions can benefit from better estimation and prediction of effectiveness of treatments at the time of product launch. An example of where registries could provide clinical practice data is the building of predictive models that incorporate data from both RCTs and registries to help in generalising results of RCTs to specific real-world settings or sub-populations.

1. ****Good registry practice****
	1. ****General consideration****

**Some guidelines describe different designs for a registry, including the traditional case series, cohort, case-control or case-cohort designs. This reflects the misrepresentation of a registry as a study rather than a data collection system. This reflection paper establishes that a registry, as a system that collects data on patients followed-up over time, inherently leads to the creation of a cohort of patients that may be secondarily used to investigate a research question through different study designs. Therefore, registry designs will not be discussed.**

* 1. ****Patient population****

Definition and enrolment of the registry patient population requires close attention. Great care should be exercised to ensure a comprehensive enrolment of patients and avoid selection bias, i.e. the situation where enrolment is influenced by patient characteristics that may affect the validity of the analyses. Registries are prone to selection bias because numerous factors may influence the enrolment of patients in a registry (including clinical, demographic and socio-economic factors) and may be difficult to identify in advance and prevent.

The following steps need to be considered in the definition and enrolment of a registry population.

1. The first step is to develop a clear conceptual definition of the target population i.e. the patients who would benefit from the information recorded from the registry and from the analysis of registry data. This definition should consider the main purpose of the registry, for examples clinical monitoring of disease, follow-up of treatment effects or another purpose. This is an important step that needs in-depth consideration and consultation of specialists and patients concerned.
2. The second step is to translate the conceptual definition into an operational definition, for example all patients diagnosed with a certain disease by a hospital specialist. It may include exclusion criteria whose rationale should be justified and documented.
3. The third step is to put in place a process whereby the patients enrolled in the registry would be representative of this operational definition. The best method to achieve this is to ensure an exhaustive enrolment of all eligible patients fulfilling the operational definition.
4. The fourth step is to ensure complete follow-up of enrolled patients and completeness of information collected on each of them to allow accounting for confounding factors and investigation of effect modifiers.

A good enrolment may generally be achieved for rare diseases but this is more difficult for more common diseases that may require more time and resources to collect data and elicit informed consent from a large number of patients. As traditional sampling methods are difficult to implement in routine clinical practice, anticipation of a non-exhaustive enrolment may require solutions such as:

• adoption of narrower conceptual and operational definitions to select patients who would best benefit from registry follow-up based on disease severity or other relevant criteria;

• collection of minimum information at baseline on all patients meeting the operational definition in order to compare characteristics of the actual registry population with those of the target population in the region or country; this information may possibly include: age and sex, indicator of socio-economic status (such as educational level), severity of disease or current treatment for the disease;

* benchmarking of the actual registry population in comparison with another data source covering the same population (e.g. electronic health care records) allowing to assess the representativeness of the registry, identify potential confounders, guide the recruitment strategy for a specific study or adapt the statistical analyses.
	1. ****Time elements****

**Accurate knowledge and recording of dates of important events are essential components of all registries as they allow computation of time periods critical to the valid analysis of the data, such as: time between entry into the registry and first exposure to medicines, time under different treatments, time of onset of adverse events of special interests (AESIs), time to recovery from an adverse event, time to achieve effectiveness, time from clinical improvement to relapse, survival time, duration of follow-up, etc. 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. In addition to the dates relevant to the clinical follow-up of patients, each registry may also need to record administrative dates related to data collection and reporting to a central registry.**

**Although the choice of time elements depends on each disease and each situation, Table 2 proposes a list of core time elements that should normally be collected in all registries. Date format and definitions should be standardised or mapped across different registries to facilitate data analysis with the same programming codes and pooling of results obtained with identical statistical analyses.**

**Table 2. Core list of dates to be collected and recorded in a patient registry**

|  |  |
| --- | --- |
| **Date** | **Comments** |
| **Patient dates****- Date of birth** **- Date of death****- Date of pregnancy****- Registry entry date****- Registry exit date****- Informed consent date** | **Exact date****Exact date****Date of last menstrual period or other estimated date of pregnancy start****Date at which observation time starts, for ex. date of first contact, or date of diagnosis****Date at which observation time has ceased due to death, move outside the catchment area, or other reason**  |
| **Disease dates****- Date of first symptoms** **- Date of first diagnosis****- Date of definitive diagnosis****- Date(s) of cure or significant improvement(s)****- Date of relapse** **- Date of occurrence of significant events associated with the disease****- Date of resolution of significant events associated with the disease** | **Relevant for diseases where diagnosis may be difficult or unprecise****May be difficult to determine for diseases without clear diagnostic criteria****Depending on the disease, for ex. date of definite diagnosis using a validated method such as MRI, histology, cyto-genetic method, etc.****Cure or improvement(s) documented by objective criteria, or milestones in treatment patterns****If applicable, relapse(s) documented by objective criteria****Significant events associated with the disease to be defined and documented.** |
| **Investigation date****- Date of test1, test2, test3…** | **Dates of key investigation tests applicable in the course of disease management.** |
| **Treatment dates****- Start date****- Stop date** | **Dates to be collected for each relevant disease-related treatment, symptomatic therapies and relevant co-therapy; date of prescription or dispensing; dates of dose changes** |
| **AESI dates****- Date of occurrence of any AESI****- Date of resolution of any AESI** | **Source of AESI occurrence date to be defined in advance and documented: occurrence of first symptoms, first consultation, diagnosis, or any other event****Source of AESI resolution date to be defined in advance and documented.** |
| **Other events dates****- Date of occurrence of any other significant adverse event** **- Date of resolution of any other significant adverse event** | **Other significant adverse event to be reported as defined in registry protocol, for example event requiring treatment or life-threatening event** |
| **Observation dates****- Dates at which follow-up status is recorded** | **Date of encounter, date of hospitalisation, etc.**  |

* 1. ****Core data elements****

Core data elements are data to be collected on all patients. Ideally, core data elements and formats should be harmonised across all registries of patients with a same disease. This is an important basis for regulatory evaluation as harmonisation to international standards facilitates implementation of a common data quality system (automated data entry control, checks for data consistency, routine statistical screening), data exchange, common data analysis on a large number of patients and interpretation of results from different registries. Lack of harmonisation may require a mapping of data elements representing the same concept but used with different definitions and terminologies. As this mapping process may be time-consuming and resource intensive, common core data elements and formats should be preferably implemented at the design or amendment stages of registries.

**Core data elements should be defined with clinicians and experts concerned by the disease as well as representatives of patients and end-users of registry information. Where it is anticipated that the data will be used for registry studies with a regulatory purpose, regulators could also be consulted in order to support adoption of an adequate set of data elements relevant for medicines evaluation. For all data elements, associated definitions and data dictionaries should be included in the registry documentation and published or made available on request.**

Examples of lists of core data elements agreed for disease registries have been published in the reports of workshops on registries for cystic fibrosis (6), multiple sclerosis (7), CAR T cell products (8) and haemophilia (9) registries. The European Platform on Rare Diseases Registration (EU RD Platform) has developed a "Set of Common Data Elements for RD Registration" (24). It is aimed to the European Reference Network's existing registries and registries under development, to all other RD registries at national, regional, and local level in the Member States, to researchers and to patient organisations. This set of core data elements for rare diseases does not include information on therapies.

**The adoption of a common set of core date elements across registries does not exclude collection of additional data that may be useful for patient monitoring or other purposes. “Crucial” data elements are those which should be collected in all registries and on which greater amounts of resources should be allocated to ensure completeness, standardisation, data quality and verification of the information. “Should have” data elements are those considered of interest and useful for some stakeholders, but not essential to all. In the perspective of regulatory registry studies, crucial data elements should include the following information:**

* **Patient data: date of birth, gender, height, weight, indicator of socio-economic status (e.g. highest educational level completed), smoking status, centre;**
* **Disease: list of core data elements to be determined for each disease with core data on diagnosis (date, test, result), grade/severity/burden of disease, important milestones in disease progression and core disease outcomes (e.g. relapse, disabilities, functional status, quality of life measure, cause of death);**
* **Co-morbidities: list of co-morbidities differentiating past and current ones;**
* **Disease-related treatments: substance, brand name, start and end dates, reason for discontinuation, important stages in response to treatment, dose, route, schedule;**
* **Other therapies: indication, start and end dates, dose, route, schedule;**
* **Adverse events of special interest and serious suspected adverse reactions: adverse event MedDRA terms, start and end dates, treatment suspected to be associated, seriousness, dechallenge, outcome of adverse event;**
* **Pregnancy: date of start of pregnancy, pregnancy outcome (spontaneous abortion, live birth, etc.).**
	1. ****Terminologies****

**As part of the harmonisation between registries covering a same disease, international terminologies should be used for diseases, diagnostic tests, symptoms, medicinal products, substances and adverse events. Table 3 provides examples of recommended international terminologies for different data elements. Where national or local terminologies are used, they should be mapped to international terminologies.**

**Table 3. Examples of recommended international terminologies for disease registries**

| **Data elements** | **Standard(s)** | **Weblinks** |
| --- | --- | --- |
| **Diseases, diagnostics, symptoms, indication for use of medicine** | **ICD-9, ICD-10, ICD-11****ICD-o-3 (cancers)** **MedDRA1**  | <http://www.who.int/classifications/icd/en/><http://codes.iarc.fr/><https://www.meddra.org/> |
| **Rare disorders (disease, malformation syndrome, clinical syndrome, morphological or biological anomaly or particular clinical situation in the course of a disorder).** | **Orphadata**(entries are cross-referenced with ICD-10, OMIM, UMLS, MeSH, MedDRA) | <http://www.orphadata.org/cgi-bin/inc/product1.inc.php> |
| **Medicinal products** | **Article 57 database (EEA)****ISO IDMP standards and related terminologies (forthcoming)** | **EMA website2****EMA website3** |
| **AESI, other adverse events, suspected adverse reactions** | **MedDRA** | <https://www.meddra.org/> |
| 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> |
| Tests | **MedDRA, ICD-10, ICD-11,**  | <https://www.meddra.org/> |
| Genetic diagnosis | International classification of mutations (HGVS) | <http://www.hgvs.org/><https://www.genenames.org/> |
| Classification of functioning/disability | International Classification of Functioning and Disability (ICF) | <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> |

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

**Medical terminologies most frequently used in clinical practice are ICD-9 and ICD-10. The ICD-11 version has been published in June 2018. The Medical Dictionary for Regulatory Activities (MedDRA) is a clinically validated and regularly updated international medical terminology dictionary used by regulatory authorities to facilitate data entry, retrieval, evaluation, presentation and sharing of regulatory information internationally for medical products used by humans. It is also the adverse event classification dictionary used internationally for regulatory purposes.**

**Another clinical terminology is the SNOMED Clinical Terms (CT) terminology. It** provides codes, terms, synonyms and definitions used in clinical documentation and reporting and is frequently used in electronic health records (25). It has been mapped to ICD-9 and ICD-10 (26).

**The EMA publishes data from the Article 57 database on all medicines authorised in the European Economic Area (EEA).** An Excel table contains the following data fields: product name (product short name: brand name or the concatenation of the generic name and the company name); active substance; route of administration; country of authorisation; name of the marketing authorisation holder; country of location of the pharmacovigilance system master file; marketing authorisation holder's contact details for pharmacovigilance enquiries. The Article 57 data will be integrated as part of the ISO IDMP standards implementation when it becomes available. Where national registries use national drug catalogues, a mapping to the Article 57 database would facilitate multi-national studies.

Another international classification for drugs is the Anatomical Therapeutic Chemical (ATC) classification and Defined Daily Dose (DDD) system, which has been developed as a tool for exchanging and comparing data on drug use at international, national or local levels (https://www.whocc.no/). This classification does not provide information on brand names. For medicinal products marketed in the US, RxNorm also provides normalised brand and substance names for clinical drugs available on the US market and links them to many of the other drug vocabularies commonly used (27).

Besides the examples of terminologies given in Table 3, standards defined by the International Council for Harmonisation (ICH) exist in different domains, for instance to record age categories of results of tests without numerical values.

The EMA may provide assistance on identifying relevant international terminologies.

* 1. Quality management
		1. ****Importance of quality management****

**Uncertainties about the quality of the data collected in registries and the level of quality management applied have often been a barrier to the use of patient disease registries by regulators and MAHs as these uncertainties undermine the confidence in the validity and reliability of the evidence generated. Concerns about data quality are particularly important in the context of post-authorisation registry studies imposed to MAHs by regulators as a condition of the marketing authorisation, where the legal responsibility to conduct the study and provide valid and reliable results lies with the MAHs. This legal context has often stimulated MAHs to create their own product registry providing them full control of the data collection.**

Quality management consists in four main activities (14):

* Quality planning: set of activities with the purposes to define quality system policies, objectives and requirements, and to explain how these will be applied and achieved;
* Quality assurance: quality standards for the registry data, procedures and computerised systems to be followed in order to meet the quality requirements;
* Quality control: activities undertaken to ensure that the defined standards are followed at every step;
* Quality improvement: activities undertaken to enhance an organisation's ability to meet quality requirements.

**Quality management should be part of the coordination of all registries in order to achieve a level sufficient for the intended purpose of the registry. In addition, registry coordinators should regularly assess and report on quality assurance activities. These activities may be costly and should be supported by dedicated funding mechanisms.**

* + 1. ****Requirements of data quality****

**Data quality includes four main components:**

* Consistency: the formats and definitions of the data entered in the registry are consistent over time and across different registries, especially across those of patients with a same disease; this component relates to the availability and use of a minimum set of common core data elements, common definitions, common data entry procedures and of mechanisms to implement changes in nomenclature systems if needed.
* Completeness: complete information on all eligible patients is recorded, with verification of missing data to keep them at a minimum.
* Accuracy: the data available in the registry is a correct representation of patient data, e.g. data available in medical charts or records.
* Timeliness: there is a timely recording and reporting of data based on their intended use and in compliance with an agreed procedure.

These four components should be considered as distinct elements, even if some of them can be addressed with the same data quality assurance measures. A good balance will need to be found between introducing additional control measures, avoiding redundancy in data collection and keeping a manageable time for data entry, as a cumbersome data entry process may increase the amount of missing data.

Data quality is a characteristic of the registry as a whole and quality control and assurance measures should be continuously in place. These routine measures will be beneficial to registry studies, even if studies can introduce additional temporary measures justified by the study objectives, for example a validation of the data entries against medical charts by two independent reviewers, verification of clinical diagnoses through data linkage to hospital data or verification of the information on drug exposure in the study population through data linkage to a claims database.

The extent of the work to be done to ensure data quality is highly dependent on the status of the registry. For new registries, the choice of the same data elements, terminologies and data entry and control software in all contributing centres will greatly facilitate management of data quality. However, many existing registries use different platforms and processes, and their harmonisation or mapping to international standards should be considered and implemented as far as possible. Real-time data collection, recording and uploading in a central registry platform is not feasible in many registries. While it should remain a long-term objective, efforts should be directed towards harmonisation of schedules for uploading data to the central platform to allow analysis on pooled data from all centres.

* + 1. Measures to improve data quality

**Measures that can be implemented to improve data quality are presented in two categories. Measures at management level are measures that should be initiated or implemented at the level of the registry coordinator (if there is a single registry with several contributing centres) or at the central level of a registry platform (for example in case several national registries collaborate to record data in a similar way for a common purpose). Measures at operational level are measures that should be initiated or implemented at a local level. i.e. in each centre contributing to the registry.**

* + - 1. ****Data quality activities at management level****
* **Registry coordinators should provide to local registries and centres harmonised definitions and data elements; a support function should be made continuously available at central level; use of a common data collection and reporting software should be considered.**
* **Standard operating procedures, work instructions, manuals and users’ guide should be developed, distributed, maintained and updated as necessary by the central office, and their compliance to their use should be regularly checked.**
* **Training sessions on processes should be organised for local registry coordinators, data managers and data custodians as applicable.**
* **If legally and technically possible, a linkage system with other national databases to double-check data or extract additional information should be established.**
* **Routine descriptive statistical analyses should be performed at each data upload into the central database to detect missing data, inconsistent data, outliers and losses to follow-up, and to request additional information to the local centres if needed.**
* **Internal or external audit with on-site review of processes and data audit may be considered and could be especially useful when planning for a registry study.**
* **Registry data can be benchmarked to external data source such as national electronic health records to compare the distribution of categories of important variables such as age, gender or prevalence of disease-related drug exposure.**
* **Indicators of data quality are defined and measured periodically (see section 5.6.4).**
* **Registries may take the opportunity of a regulatory Qualification of registries by the Committee for Medicinal Products for Human use (CHMP), as illustrated by the Qualification Opinions issued for the European Cystic Fibrosis Society Patient Registry (ECFSPR)** (10) **and the Cell therapy module of the European Bone Marrow Transplantation (EBMT) Registry** (11)**.**
	+ - 1. ****Data quality activities at operational level****
* **Standard operating procedures and work instructions should be introduced and followed.**
* **Accurate data entry is a key component of data quality; adequate training of data managers and other persons involved in data entry is critical.**
* **Automated data quality checks and visual prompts should be in place at data entry to prevent introduction of erroneous or inconsistent data and trigger source data verification.**
* **Standard data quality control reports should be produced at local level to check for missing, unusual or incorrect data; source data verification through manual checks of medical charts or records should be performed secondarily.**
* **The procedure for data submission to the central registry should include in a first stage a check of the completeness of the data and provide the sender a missing data report used to increase completeness.**
* **Indicators of data quality are measured periodically.**
	+ 1. ****Indicators of data quality****

**Indicators of data quality may be useful tools to measure and improve data quality but they should be associated with remedial measures if acceptable levels of quality are not observed. The feasibility of measuring a large number of indicators may also be an issue. Table 4 provides the example of the quality indicators agreed in the EMA CAR T-cell Therapy Registries Workshop in the context of the monitoring of the CAR-T cell products by the European Bone Marrow Transplantation Registry (EBMT) and the US-based Centre for International Blood and Marrow Transplant Research (CIBMTR), as well as the feasibility of their introduction in routine** (8)**. Other examples of quality indicators have been discussed for cystic fibrosis registries** (6) **and for multiple sclerosis registries** (7)**.**

**Table 4. Proposed indicators of data quality and feasibility of implementation in the EBMT and CIBMTR registries** (8)**.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data Quality Component | Definition | Proposed indicators of quality | Quality Solutions to facilitate data quality | EBMT & CIBMTR feasibility |
| **Consistency** | Uniformity of the data overtime (e.g. lab data routinely entered) | Number of fields changed over time | Manual checks at centres level, audits | Both: Feasible |
| % of fields missing over time | Standard terminology, coding | Both: Feasible |
| % of forms reported per scheduled follow-up  | Standard operating procedures, user guides | Both: Feasible |
| Campaigns, dashboards for clinicians | Both: Feasible |
| **Accuracy** | Accuracy of data entry: no errors, no contradictions or impossibilities in data, absence of duplicates | Change in value of data filed by x% creates alerts | Drop down menus, alerts, text prompts, flags | EBMT: FeasibleCIBMTR: Simplify data collection to avoid redundancy |
| Variability across fields    | Validate against source data (e.g., 10%), cross form validation |  EBMT: Costly and only currently done for funded studies CIBMTR: Suggests ‘crucial’ elements be audited and ‘acceptable’ error rate defined (3% in CIBMTR)  |
| Staff training, software checks. | Both: Feasible |
| Help screens/desks, training, newsletter | Both: Feasible  |
| Funding for data managers | EBMT: Requires new fundingCIBMTR: Necessary to motivate data collection |
| **Complete-ness** | How much data is missing? | Agreed % of fields completed in audit procedures (e.g. >90%) | Audits | EBMT: Costly and only done for funded studies currentlyCIBMTR: May be reported directly from the registry |
| Lost to follow up %   | Mandatory fields | EBMT: FeasibleCIBMTR: Feasible for ‘crucial’ and ‘should have’ elements |
| Engagement with patients and/or health care providers (HCPs) | Both: high engagement with HCPs, less with patientsCIBMTR: Implementing systems to collect patient reported outcomes |
| Absence of core variables | Minimum agreed core common data elements reported | Agreed list of data elements and definitions | EBMT: FeasibleCIBMTR: Feasible for ‘crucial’ and ‘should have’ elements |
| All treated patients reported, not selected patients only | Cross check patient numbers with numbers of products used at treating centres during a defined period  | Both: Feasible if there is access to orders/product supply information |

* 1. ****Safety analysis****
		1. ****Reporting of safety information****

**Disease registries conducted by organisations such as academia or medical research associations should follow the national requirements for the management of safety data. Reporting of suspected adverse reactions to the national or regional pharmacovigilance system could be facilitated by integrating an electronic reporting module in the registry software or by giving access to an electronic reporting module put in place by the NCA. The fact that the information was generated from a disease registry could be included in the case description.**

**Any active data collection system put in place in a disease registry and initiated, managed or funded by a MAH with the objective to collect and record suspected adverse reactions to one of its medicinal products must follow the regulatory framework for PASS** (20)(21) **(see chapter 6.8). In this situation, collected adverse events should be managed by the MAH as *solicited reports* andthe mechanisms implemented to record, document and evaluate case information in order to allow the submission of valid ICSRs to EudraVigilance should be described in a protocol.** Relevant safety information should be summarised in the product Periodic Safety Update Report (PSUR) by the MAH and in the registry reports. **Contractual agreements should be in place to clearly define the role and responsibilities of each party.**

In registries without active safety data collection, HCPs and patients should be reminded of the possibility to report adverse reactions (for which they suspect a causal role of a medicine) to the MAH of the suspected medicinal product or to the concerned competent authority via the national spontaneous reporting system. Any suspected adverse reaction notified to the MAH should be managed by the MAH as a *spontaneous report* (21).

* + 1. ****Monitoring of adverse events of special interests****

**In disease registries, data on medical events are generally uploaded periodically into a central database for statistical analyses. Disease registries are therefore generally not suitable for a rapid statistical analysis of new safety signals. Disease registries may however be useful for the monitoring and characterisation of known or suspected adverse reactions previously identified from clinical trials, the pharmacology of the drug, the spontaneous reporting system or other sources. In such case, a list of adverse events of special interest (AESI) can be defined and integrated in the routine data collection system with sufficient information to allow subsequently the conduct of further investigations on risk of occurrence, association with exposure(s), delay of onset, and severity and clinical outcome of the adverse reaction. For medicinal products with a RMP, AESIs may represent the important identified risks, important potential risks or missing information. Standard definitions and the MedDRA terminology should be used for the coding of AESIs.**

* + 1. Aggregate analysis of adverse events

**Routine statistical analyses of collected adverse events could be based on standard statistical programming that can be run periodically by registries or upon requests from a regulator or a MAH, and agreed with them. Analyses can be descriptive but can also integrate comparisons between different treatment groups (based on therapies, dose, duration, etc.) or categories of patient characteristics. Routine statistical analyses should be defined in advance and described in a SAP. Comparative analyses could require defining comparator exposure groups and controlling for confounding factors, especially confounding by indication. Consideration should also be given to various treatment periods to which patients may be exposed, especially in case of switching between therapies, in order to avoid time-related bias (hence the need for an accurate recording of times of entry into the registry and start and end times of different therapies).**

* 1. ****Governance****
		1. ****Governance principles****

Most registries and potential users of registry data have a governance model relying on principles and constraints based on their mandate, operating procedures, legal environment or funding sources. Effective collaboration between all parties is therefore needed to ensure early identification of relevant registries and evaluation of the adequacy and quality of the data collected if the data are to be used for regulatory purposes. A number of activities that may be performed by registry coordinators, MAA/MAH and regulators are proposed below to strengthen the governance for the use of registry data.

* + - 1. Role of registry coordinators

Many registries have been established for clinical evaluation and academic research purposes and some do not have processes in place to allow use of the data for other purposes, e.g. to support monitoring of drug utilisation or safety of medicinal products, with limited flexibility for providing individual patient level or aggregated data to external organisations. Divergent priorities among stakeholders must also be recognized: many patient registries will participate in research only if they can preserve their scientific independence, maintain control of the management of registry data and provide only aggregated data without having to re-consent patients for specific registry studies. This principle may diverge from priorities of other stakeholders, especially where a PASS has been imposed to a MAH as a legal obligation by a regulatory authority.

The consideration of patient registries for regulatory purposes respecting stakeholders’ priorities could be facilitated by the development and publication of clear governance principles and support functions by registry coordinators. Such principles and functions could include:

* a published documentation of key registry characteristics, such as patient inclusion and exclusion criteria, extent and detail of data collected and available for secondary use, procedures and instructions for data collection, upload and analysis (with timelines), and quality management;
* a single contact point for requesting information on available data and data access conditions, including for multi-registry platforms;
* a published policy for collaboration with external stakeholders (pharmaceutical companies or other organisations), including information such as the acceptable scope of collaboration and the range of activities that can be routinely performed by the registry in such collaboration, the extent of safety monitoring, the type and detail of data that may feasibly be collected, principles of data ownership, data sharing and data analysis, possibilities for pooling data from multiple registries and for linkage with other data sources, the approval process with estimation of typical timelines for making decisions, as well as principles for receiving funding from private and public sources; this principle would be supported by the registration and description of the registry in a public inventory of registries;
* a process for managing requests for collaboration together with a governance structure for decision-making (e.g. steering committee, ethics committee, advisory board);
* a support function for scientific aspects of the requests such as additional data collection, study protocols, interoperability between registries, relevance and feasibility of pooling patient-level or aggregated data, statistical analysis plans, consideration of methodological differences between centres or registries, estimation of sample size, etc.;
* a support function for ethical aspects of the requests such as data ownership, data protection, scientific independence, compliance with the informed consent system in place, acceptability of the proposed publication process, etc.;
* a template for a research contract between the registry and external stakeholders.
	+ - 1. Role of pharmaceutical companies

Pharmaceutical companies could also support early consideration of the feasibility and appropriateness of using a disease registry for the post-authorisation monitoring of their products. Their activities could include:

* to understand before or at an early stage of the authorisation application process the regulatory data requirements that are likely to arise during the application process especially in planning for post marketing surveillance given e.g. limited exposure to the drug in rare diseases or the prolonged duration of follow-up that is required for some products;
* to initiate discussions with disease registry coordinators and regulators before –using the scientific advice procedure- or at an early stage of a marketing authorisation application on the relevance and adequacy of one or several existing patient disease registries for the long-term monitoring of their specific product;
* to have an in-depth understanding of the extent and detail of data available in patient disease registries when planning registry-based post-authorisation studies;
* to develop a preliminary study protocol for post-authorisation studies of any new product and explore with the registry coordinator and the regulator how the registry could fulfil the data needs, for example through the Scientific Advice procedure; joint study protocols developed by several MAHs of products of the same therapeutic class are encouraged;
* as applicable and needed, to liaise with registry coordinators to discuss means (including additional funding) to increase data quality control, to support modifications of an existing registry or to meet other requirements to comply with their regulatory obligations.
	+ - 1. Role of regulatory authorities

To some extent pharmaceutical companies and regulatory authorities have complementary roles in relation to disease registries. Their common objective is to make best use of disease registries to support medicines evaluation and to discuss at an early stage of the marketing authorisation application whether and which registries can provide adequate information.

Specific roles may include:

* to identify relevant patient disease registries in the pre-submission process or early application process and gather information on which data are routinely collected or can be additionally collected for specific products;
* to support interactions between MAAs/MAHs, registry coordinators and regulators to agree on data elements, collection methodologies, timelines and protocols;
* to support quality management by describing expectations as regards data quality and providing formal scientific advice or qualification opinion if requested;
* to provide methodological guidance on core data elements, terminologies, data quality, data collection and analysis for specific disease registries, upon request from registry coordinators and pharmaceutical companies.

Discussions between regulatory authorities and MAAs/MAHs can take place under the umbrella of requests for scientific advice. At the EMA level, existing regulatory procedures also include the Innovation Task Force, the EMA Qualification procedure or the parallel EMA/HTA Scientific Advice in support to high quality data.

* + 1. Data ownership and intellectual property

Patients remain in the control of the use of their data. They may or may not consent for their use for clinical or research purpose and they may withdraw their previous consent.

It is the responsibility of local centres and registry coordinators to use and share the data in accordance with the EU General Data Protection Regulation (GDPR) and the informed consent provided by the patients. In situations where data are shared with MAHs based on a contractual agreement, the research contract should clearly describe the level of access to data, the intellectual property rights arising from the use of the data and the dissemination of the results. In case of specific studies performed by registry investigators and funded fully or partially by MAHs/MAAs, it is recommended to refer to the ENCePP Code of Conduct (15) (and ADVANCE Code of Conduct for vaccines (28)) in the research contract to ensure scientific independence and transparency, while allowing sharing of unpublished results with regulators and the MAHs/MAAs concerned.

Some registry coordinators have expressed concerns that unpublished data communicated to regulators or to MAHs in the context of regulatory procedures could be made publicly available through Regulation (EC) No 1049/2001 regarding public access to documents, thereby undermining their intellectual property rights. However, as highlighted in the EMA policy on access to documents (29), the EMA may restrict the access to documents and, prior to the release of documents, will always liaise with the entity having produced the document (in this case a registry coordinator or a MAH) to discuss potential steps necessary to protect commercially confidential information and personal data if applicable.

* + 1. Data sharing
			1. General considerations

Data from disease registries are important for regulators as they may help understand the disease, monitor the safety and effectiveness of medicinal products, and take appropriate decisions to protect patients’ health. Aggregated data (supported by statistical analysis if needed) are generally sufficient. Regulators will rarely request patient-level data or analytical datasets, and requests for pseudonymised patient data would be based on important public health reasons. Sharing of registry data should be based on a voluntary agreement between different parties.

The data sharing process may be time consuming, especially if this involves sharing data not previously collected or aggregated. It is therefore recommended to MAAs/MAHs to initiate discussions with regulatory authorities and registry coordinators on a possible use of registry data at an early stage during the authorisation process, for example at the time of pre-submission meetings. Early interactions between companies, regulators and registry coordinators would facilitate mutual in-depth understanding of the extent and data available in patient registries and of the regulatory data requirements that are likely to arise during the application process, especially in planning for post-authorisation surveillance.

The format of the data shared with MAHs and regulators will vary according to the objectives of data sharing and the agreement between different parties. The data should be presented using international terminologies, especially the MedDRA terminology for adverse events or suspected adverse reactions.

* + - 1. Data sharing with MAAs/MAHs

There are a number of situations where data sharing with MAAs/MAHs will support the development, authorisation and monitoring of medicinal products:

* natural history cohort: registries may provide important data for the planning of the clinical development and for deriving assumptions on the sample size needed, expected effect size and endpoints, such as data on disease epidemiology in terms of prevalence, incidence and outcomes of standards of care, on the validity of biomarkers, and on characteristics of the target population according to co-morbidities, medication use, complications and safety concerns;
* external control cohort: as part of the clinical development, disease registries may provide an external comparator cohort with data on unexposed patients with similar demographic and clinical characteristics (i.e., disease severity, age group, median/mean follow-up time, comorbidities) to the clinical programme under study;
* post-authorisation: MAHs may have to provide PSURs with cumulative tabulations of adverse reactions they have been made aware of during the period covered by the PSUR. These adverse reactions may be derived from various sources, including disease registries from which periodic reports can be generated. These reports may include drug utilisation data, incidence data of suspected adverse reaction and clinical description of cases. The standard submission schedule of PSURs for a marketed product is 6 monthly following initial placing on the market in the EU for 2 years, then once a year for the following 2 years and thereafter at 3-yearly intervals. This schedule may be amended depending on circumstances; for example, regulators may require that PSURs are submitted annually for a period longer than 2-years post-marketing. The periodicity of reports generated by registries could be adapted to the standard PSUR submission schedule;
* in the context of the investigation of a new safety signal, a MAH may be asked to collect cases of the suspected adverse reaction from additional sources; registries may therefore be requested to provide any available information on cases related to the signal or cumulative data that may be extracted from a previous registry report or from an ad-hoc statistical report;
* in some circumstances, a PASS or a PAES may be imposed as a legal obligation where the results of such studies are considered critical for the assessment of the benefit-risk balance of a medicinal product; an example of such circumstance is where it is deemed necessary to verify the implementation of risk minimisation measures introduced to prevent occurrence of a severe adverse reaction; disease registries are very important sources of information in such situations as they may provide data on the everyday clinical practice. For such studies, the study protocol needs to be agreed between the MAH and the registry coordinators before being endorsed by the Pharmacovigilance Risk Assessment Committee (PRAC); the PRAC may also request amendments which may need to be discussed by the MAH and the registry coordinators.

The range of situations where MAHs may contact registries to get data is therefore potentially very large and each situation may require different agreements for data sharing, for example provision of statistical reports generated by the registry, of electronic tables of study results to be included in a study reports, of pseudonymised analytical datasets including only the data elements to be analysed, or of study reports to be submitted by companies to regulators. Provisions of pseudonymised patient-level data or analytical datasets to companies are generally not accepted by registries for policy reasons or concerns about patient privacy. In such cases, data analysis might be performed either by the registry team or by a trusted third-party. Irrespective of the type of information provided by the registry, it is recommended that agreements between registries and companies include a provision on the use of the ENCePP Code of Conduct on scientific independence and transparency (15) (or the ADVANCE Code of Conduct for collaborative vaccine studies (28)). Additional national provisions may be in place.

* + - 1. Data sharing with regulators

Regulators also need data on disease epidemiology, standards of care, and safety and effectiveness of medicinal products. In some circumstances, regulatory authorities may directly contact registry coordinators to get access to data. For safety, the context of such request may be:

* signal detection, to clarify data on adverse reactions submitted to a NCA;
* signal validation and assessment, to confirm or update the number of cases of a suspected adverse reaction, to get additional information on cases, or to provide further evidence for or against a causal association or a new aspect of a known association;
* analysis of specific aspects of a suspected causal association, such as the relationship between a suspected causal association with drug dose, concomitant medications or some patient characteristics;
* feasibility analysis for a planned registry study;
* investigation of a safety concern, either through direct analysis of data or through funding agreement with a third-party;
* assessment of the effectiveness of risk minimisation measures, for example through an analysis of the indications for which a product has been prescribed.

In such circumstances, discussions should address patients’ privacy, the informed consent in place and the possible need to re-consent patients.

* + 1. ****Informed consent****

Patients need to be aware of why data is collected, what data is collected, how it will be used, and by whom and with whom it will be shared, and at what level of details. In addition, some patient registries have been expanded to include additional data such as genetic profiling and other biochemical analyses. This data is sensitive and it is important that patients have a good understanding of the data that could be provided to external organisations. Principles of informed consent should be applied in accordance with the GDPR (30).

Patients must grant their permission to process and use their data in full knowledge of the possible benefits and risks of their participation and of the measures applied to mitigate risks. Patients must also be aware that they can restrict the scope of their consent and withdraw the consent at any time. For consent to be valid, it must be *voluntary* and informed, and the person consenting must have the capacity to make the decision. It is the responsibility of the treating centre to ensure that patients have consented to the recording and use of their data. For children, the consent is signed by the parent or legal guardian and children need to provide their own consent once they reach the adult age of consent. There is therefore a risk that their data could not be made available if no procedure is in place to obtain their informed consent as adults. The specific situation of older people with impaired cognition (dementias) and supported by carers needs also to be considered.

As the current framework for informed consent is not harmonised across countries, the informed consent process is dependent on local requirements. An electronic system of informed consents would facilitate harmonisation of consent forms in all registries, and keeping track of consents, restrictions and withdrawals. This system would also support registry coordinators to get confirmation about patients consenting or not to share their data and about compliance to legal obligations.

If a new informed consent form is developed or if the current consent needs to be amended, it should be ensured that the informed consent is 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 across country borders and with other stakeholders including regulators and MAHs.

* + 1. ****Data security****

Security measures should be implemented to maintain the privacy of patients enrolled in a registry, described and documented in standard operating procedures. Security measures should be applied for data collection, storage, archiving, transmission and access. Each registry has the responsibility to ensure the security of the data collected in line to the provisions of the GDPR (30).

1. Registry studies
	1. Regulatory context

A number of obligations described in the legislation (Directive 2001/83/EC and Commission Implementing Regulation (EC) No 520/2012) apply to the MAHs when they conduct post-authorisation studies involving their own products, irrespective of the data source used for the study, and these obligations therefore also apply to registry studies (31) (32).

A non-interventional post-authorisation safety study (PASS) may be initiated, managed or financed by a MAH voluntarily or pursuant to obligations imposed by regulatory authorities (20). Article 107m of Directive 2001/83/EC describes the obligations to be followed for all studies, and Articles 107n to 107q provide additional obligations for imposed studies. According to this legislation, the MAH holds the responsibility for the supervision of the PASS, including the obligations to monitor the data generated, to consider its implications for the risk-benefit balance of the medicinal product and to communicate to authorities any new information which might influence this balance. If the conduct of registry study is partially or fully sub-contracted by a MAH to a third-party (for example a contract research organisation, an academic institution or the registry coordinator), the MAH should ensure that these activities are carried-out and should provide adequate funding for their implementation. The research contract established by a MAH and a third-party for such studies should clearly describe roles and responsibilities and it should follow the recommendations of GVP Module VIII (20) and of the ENCePP Code of Conduct (15) (as well as the ADVANCE Code of Conduct for vaccines (28)).

In addition to its pharmacovigilance obligations in relation to the study, the MAH should ensure that the fulfilment of these obligations can be audited, inspected and verified. For PASS imposed as an obligation, the MAH shall ensure that 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. Obligations also apply to other PASS required in the risk management plan or conducted voluntarily by the MAH in terms of the recording, retrievability and traceability of pharmacovigilance information. Registry coordinators should be aware of these legal obligations to MAHs and these obligations should be reflected in the research contract.

In order to support transparency on non-interventional PASS conducted voluntarily or pursuant to an obligation, and to facilitate exchange of pharmacovigilance information between the EMA, member states and marketing authorisation holders, Module VIII of the GVP recommends that MAH should make study information available in the EU electronic Register of post-authorisation studies (EU PAS Register) maintained by the EMA. This registration is a legal obligation for imposed PASS for which the registration number must be mentioned in the final study report. For imposed studies, the EMA has the obligation to make public the protocol and an abstract of the study results (33). GVP Module VIII therefore recommends that MAHs anticipate this obligation by taking the initiative to enter themselves this information in the EU PAS Register.

A PAES may also be imposed where there is a well-reasoned scientific uncertainty to be addressed post-authorisation to enhance the understanding of therapeutic efficacy and the benefit-risk of a medicine with implications for better use in clinical practice. The PAES guidance (22) recommends that agreement on a suitable methodology giving reliable and interpretable answers to the research question should be sought as early as possible between the regulator and MAH, taking account of the post-authorisation setting and the required timeframe. Randomisation can be incorporated in a disease registry setting, if called for by the study design best to address the research question (e.g. efficacy).

To support transparency about registries used in regulatory studies, it is also recommended that the registry coordinator enters information on the registry in the ENCePP inventory of data sources (34).

* 1. Timelines

Initiation of a registry study requires consideration of many different aspects that can take overall more than one year: identification of suitable registry(-ies), feasibility analysis, writing of study proposal, submission of request to registry steering or scientific committee, decision-making within registry governance structure, development of study protocol and statistical plan, legal considerations, including submission to ethics committees, agreement on data ownership and publication rights, contract agreement per country and site, database creation, infrastructure implementation, centre identification, budget calculation for different study phases, etc. For this reason, preparations and discussions about use of a disease registry post-authorisation should start before or at an early stage of the approval application process making use of existing regulatory procedures such the pre-submission meeting or scientific advice. It is recommended that such early discussions involve the MAH(s), regulators and registry coordinators.

It is important that the study proposal submitted by a MAA/MAH to a registry holder should be sufficient detailed as it will be used by registries and individual centres to assess whether they can participate in the registry study in terms of data availability and required quality standards. The study proposal should therefore provide detailed requirements in terms of data needed and procedures for data collection, analysis and dissemination.

* 1. Study protocol

Patient disease registries provide cohorts of patients who can be analysed using different types of study design and the choice of the study design should be the most appropriate for the research question. Registry studies should follow the best methodological standards applicable to pharmacoepidemiological research and the protocol should describe measures foreseen to account for bias and confounding and ensure the internal validity of the study. It is not an objective of this document to replicate methodological guidance available elsewhere such as in the ENCePP Guide on methodological standards in pharmacoepidemiology (14). The ENCePP Checklist for Study Protocols (35) lists important questions to be considered when designing a study and writing a protocol.

An early decision to be made when designing a registry study is the choice of the data collection method: secondary data collection, where the data for the study are already available and extracted from a dataset, and primary data collection, where the events of interest for the study are collected directly from patients as they come to the attention of the investigator. This choice has implications for safety reporting (see chapter 6.8) and should be clearly specified in the study protocol.

The study protocol should follow the recommendations of the GVP Module VIII and the technical guidance on the format and content of the protocol for non-interventional PASS (36). 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 study based on a single protocol. If a registry study is to be conducted across multiple sites, a common protocol needs to be developed based on core common data elements and common design even if some aspects of the study may vary according to the characteristics of each registry. In some registries, additional data may need to be collected and these may need to include a request for pseudonymised individual patient data if appropriate and to require additional informed consent. The protocol should also provide details on mechanisms put in place to identify and collect missing data as well as to minimise the number of patients lost to follow up.

The protocol should provide an estimation of the study size and the feasibility of attaining this sample size within the registry should also be assessed using conservative assumptions (or a previous feasibility analysis), both in terms of number of patients (taking into account the inclusion and exclusion criteria) and in terms of duration of follow-up based on assumptions for losses to follow-up. Where there are doubts about the possibility to achieve the sample size, the possibility to extend the study population by recruiting one or several other registry(-ies) should be discussed.

For imposed (category 1 and category 2) regulatory PASS, the final protocol should be endorsed by the PRAC before the study starts. For non-imposed (category 3) studies included in a RMP, the protocol may also need to be agreed by regulators upon request from a PRAC or CHMP Rapporteur. Although subsequent amendments to the protocol are possible, these should be avoided as much as possible as new requests, especially for additional data elements, may be time-consuming to implement and they may require additional funding.

* 1. Study population

The registry patient population represents the source population from which the study population will be selected for a registry study based on inclusion and exclusion criteria. Depending on the objectives of the study, the study population may be identified through extraction of data already collected routinely in the registry. In this situation, methodological challenges to be considered are comparable to those met in using electronic health records, such as choice of index date, definition of exposure periods, definition of outcomes, choice of comparator groups, etc.

When the safety or effectiveness of a new treatment need to be monitored in a disease registry, the study population will most often include new users who are either patients with a new diagnosis of the disease and a first prescription (incident patients) or patients already included in the registry and switched from an earlier treatment to the new one (prevalent patients). The choice of including incident or prevalent patients in the study population has important implications for the data analysis and the interpretation of the results (see section 6.7). It is therefore recommended to include all patients receiving the new treatment in the study population and collect the data needed to distinguish incident and prevalent patients and identify possible differences in their characteristics. A description of the characteristics of new users in comparison to patients included in clinical trials may also provide useful information to interpret the data collected in the study.

In case of prospective recruitment, it is critical that procedures are in place to ensure sequential inclusion of all eligible patients treated in the individual centres. Patients having been involved in clinical trials, often representing a subgroup with certain disease characteristics (e.g. genetic variants), could also be enrolled in a disease registry study later on. Particular attention should be given in studies where procedures and data collection forms are added to the routine practice as time pressure and administrative workload may discourage investigators to recruit some of the eligible patients. Procedures should therefore be in place at the analysis phase to allow identification of eligible patients not recruited in the study and compare important socio-demographic (e.g. older age) and clinical characteristics (e.g. disease severity) between eligible patients recruited and not recruited.

* 1. Data collection

Registry studies may not need the totality of the information collected in the registry to answer the research question. Even if the informed consent allows use of the patients’ data for research purpose, it is the investigators’ responsibility for ethical reasons to collect **only** the set of data that is strictly needed to provide valid results. However, it is also their responsibility to collect **all** the data needed for this purpose as it could be considered unethical not to make best use of patients’ data if, for example, data on important confounders were not extracted and bias is introduced, compromising the validity of the reported/published results.

Data to be collected or extracted include data required for sensitivity analyses. Based on the assumptions used in the study design, investigators should plan in advance (and include in the protocol) sensitivity analyses that will test if any association that has been found holds true with different assumptions about the value of study parameters. Sensitivity analyses may also be planned to test the impact of effect modifiers such as dose or duration of treatment or patients characteristics. Sensitivity analyses must be planned at an early stage as retrospective collection of additional data for post-hoc analyses may be particularly difficult in registries.

For certain studies, modifications to the existing data collection procedure may be implemented for additional data collection, e.g. to address a particular research question or to address specific issues required in PASS/PAES which were not part of the core registry dataset. For registries with participation of several individual centres, it may be preferable to add a data collection form or module rather than to modify the existing one. In such cases, relevant modifications to the consent form may be needed for prospective data collection or patients may need to re-consent for retrospective data analysis.

The legal status of the study should also be considered as additional data collection may possibly lead to the study being considered interventional. According to the legislation (37), non-interventional studies may include those involving primary data collection, provided that 1) the medicinal product is prescribed in the usual manner in accordance with the terms of the marketing authorisation, 2) the assignment of the patient to a particular therapeutic strategy is not decided in advance by a trial protocol but falls within current practice, and the prescription of the medicine is clearly separated from the decision to include the patient in the study, and 3) no additional diagnostic or monitoring procedures are applied to the patients and epidemiological methods are used for the analysis of collected data. There is a possibility that the third condition would not be met in case of additional data collection, even if interviews, questionnaires, blood samples and patient follow-up may be part of normal clinical practice.

Additional data collection will need to be detailed in the study protocol submitted to the relevant ethics committee(s). The fact that a registry study is considered interventional has no influence on the non-interventional nature of the registry itself.

Data collection on suspected adverse drug reactions are discussed in chapter 6.8.

* 1. Data quality

The need for additional measure for data quality control in a registry study depends on the measures applied routinely in the registry. In order to ensure acceptable data quality for individual registry studies conducted for regulatory purposes, source (registry) data verification and periodic auditing on a reasonable amount of data may need to be conducted on a risk analysis-based approach and following a strategy dependent on the scope of the study. As a general rule, source data verification for a minimum of 10% of randomly selected patients registered in individual study centres would be considered adequate. The level of data verification will have to be agreed upfront between the registry coordinators and the MAHs in the context of the study performed. Appropriate measures in case relevant findings are observed should be specified. Quality of short and long term data should be assured. Agreements on relevant logistical aspects should be made between the registry and MAHs in advance of study start.

* 1. Data analysis

Data analysis for a registry study performed as a regulatory requirement should preferably be performed by the registry coordinator or by a third-party (e.g. academic centre or contract research organisation) rather than by MAAs/MAHs. If data analysis is conducted by the registry coordinator or a third party, results of product-specific data analysis should be shared with regulators and the concerned MAAs/MAHs. The process for this communication should be specified in the study protocol.

There is no single statistical method applicable to every study and the most appropriate statistical procedure would need to be selected/tailored on a case-by-case basis to specifically address the scientific question of interest. Of particular importance is the type of outcome to be analysed, i.e. descriptive analysis of drug utilisation patterns, comparison of safety between different treatments or effectiveness of a new therapeutic strategy. Pre-specification of the analytical approach before data are generated on any investigational or comparator treatment is critical, so that choices on analytical approach are not determined by the data generated (22).

Analysis of studies based on registry data should take into account that they are selected observational cohorts and hence knowledge of techniques developed to describe such data is essential. In depth discussion can be found in many texts on epidemiological methods but some common problems may arise:

* Measurement of the incidence of events of interest should clearly distinguish between event rates (number of events in a specified group in a specified period) and incidence rates (number of individuals presenting at least one event in a specified group in a specified period). The distinction is important where an individual may present several events. Event and incidence rates should be stratified by sex, age categories, registry (in multiple registry studies), treatment categories and other relevant variables. Comparisons between categories should take observation periods into account and use parameters such hazard ratios (HR) or incidence rate ratios (IRR) adjusted for potential confounders.
* In the absence of randomised decisions concerning treatment allocation it must be recognized that patient groups given different treatments are likely to differ systematically. Hence differences in clinical outcomes may not be attributable to treatment. One feature that is particularly likely to differ is indication, i.e. the reasons why a patient was prescribed a particular treatment. Such decision may be influenced by a large number of different factors that need to be taken into account in the analyses as these factors may also be associated with the risk of occurrence of the outcome of interest. The propensity score approach can be used to address the problem, but the appropriate method and the assessment of whether the propensity score approach is valid will depend on the research question, patient population, outcomes for the specific study and data available. Consideration should also be given to the facts that the potential set of variables required for the propensity score approach may vary among studies and will need to be available for all exposure groups, and that the appropriate use of propensity scores in the analysis (i.e. matching, inverse probability of treatment weighting (IPTW), stratification and adjusting) depends on the study characteristics. When an instrumental variable (IV) can be identified it may also be considered to address unidentified confounding in comparative studies, for example if the confounders are unknown or are likely to be measured with error.
* Registries offer the opportunity to compare patients receiving a new drug with patients who are untreated or who have received different therapies over a long period of time. Use of prevalent drug users in a registry study (i.e. patients already treated for some time before study follow-up begins) can cause two types of bias: firstly, prevalent drug users are “survivors” of the early period of treatment, which can introduce substantial bias if risk varies with time, and secondly, covariates for drug use at study entry (e.g. disease severity) may be affected by previous utilisation of the drug itself, which may introduce confounding. A new-user design eliminates these biases by restricting the analysis to incident drug users, i.e. persons enter the study cohort only at the start of the first course of the treatment of interest during the study period. Its weaknesses, however, include reduced precision of estimates due to lower sample size, a likely reduction in the number of patients with long-term exposure, and the possibility of accentuated selection bias when a new treatment is compared with new-users of one that is potentially outdated. Moreover, comparison of new-users of competing treatments may not be the question of interest.
* A common error in analysis may be made in registry studies when the follow-up period starts (sometimes a long time) before the start of the treatment under study, especially where the patients switch from one drug to another and a comparison is made with untreated patients. Immortal time bias can arise when the period between the start of follow-up and the date of first exposure to the drug of interest, which is unavoidably event-free in case of a drug-specific effect, is either misclassified or simply excluded and not accounted for in the analysis. Both situations may lead to bias. The solution is to use a time-dependent definition of exposure that properly classifies the immortal person-time as unexposed until the start of drug use and exposed thereafter. In nested case-control studies, a same person may therefore contribute to both exposed and non-exposed person-time.
* Two other situations merit consideration:
	+ use of a comparative non-exposure group from outside the registry, for example another registry or from electronic health care records in a country/region where the drug has not yet been marketed; in this case, considerations should be given to:
		- correctly define a comparable index date of entry into the study in both groups,
		- correctly account for exposure periods to different drugs, and
		- collect data on determinants of exposure to different drugs and, if relevant, of inclusion into two different data collection systems.
	+ use of historical controls who were enrolled in the registry at the time where the new treatment did not exist yet; in addition to the above-mentioned aspects, one should take into account the historical time period, the best therapeutic approaches existing in that period and secular trends in the occurrence of important events in the study.

The handling of missing data should be described in the statistical analysis plan and justification should be provided for the assumptions about their distribution, reasons for and timings of missing data and whether missingness is related to treatment assignment or outcome. Sensitivity analyses may be proposed to show that conclusions drawn from the data are not sensitive to the particular strategy used to handle missing values.

* 1. Safety reporting

The legal requirements to MAHs for the reporting of safety findings from non-interventional PASS make a distinction between studies with a design based on primary data collection directly from HCPs or consumers (i.e. where the events of interest for the study are collected directly from patients as they come to the attention of the investigator), and studies with a design based on secondary use of data (i.e. where the events of interest have already occurred and have been collected for another purpose, for example when a registry study only use data that have already been collected (21)).

In the case of primary data collection, the following provisions apply to MAHs (see also Table VI.1 of GVP Module VI (21)):

* Information on all adverse events should be collected and recorded from HCPs or consumers in the course of the study unless the protocol provides with a due justification for not collecting certain adverse events.
* Cases of adverse reactions, which are suspected to be related to the studied medicinal product by the HCP, the consumer or the MAH should be recorded and submitted as ICSRs to the EudraVigilance database.
* All fatal outcomes should be considered as adverse events which should be collected. In certain circumstances, suspected adverse reaction with fatal outcome may not be subject to submission as ICSRs, for example because they refer to study outcomes (efficacy end points), because the patients included in the study have a disease with high mortality, or because the fatal outcomes have no relation to the objective of the study. For these particular situations, the rationale for not submitting as ICSRs certain adverse reactions with fatal outcome should be clearly described in the protocol together with a list using the appropriate level of the MedDRA classification.
* All adverse events collected during the study should be summarised in the interim safety analysis and in the final study report.
* For adverse events specified in the study protocol which are not systematically collected, healthcare professionals and consumers should be informed in the protocol (or other study documents) of the possibility to report adverse reactions (for which they suspect a causal role of a medicine) to the marketing authorisation holder of the suspected medicinal product, or to the concerned competent authority via the national spontaneous reporting system. Where made aware of them, these reports should also be summarised in the relevant study reports by the MAH.

In the case of secondary data collection, the submission of suspected adverse reactions in the form of ICSRs is not required. All adverse events/reactions collected for the study should be recorded and summarised in the interim safety analysis and in the final study report unless the protocol provides for different reporting with a due justification.

Assessment of adverse events to identify suspected cases of adverse reactions may represent a complex task requiring knowledge of the safety profile, pharmacology and pharmacodynamic properties of the concerned product. This task may be difficult to perform in the context of routine registry activities. In case of primary data collection, the concerned MAH(s) should therefore have in place a specific data collection or electronic system to collect, analyse and report if applicable information on the adverse event notified by a HCP or a patient.

The safety analysis in the final report is descriptive and includes summary statistics. Registry coordinators should also receive the results of the safety analysis even if this analysis was carried-out by the MAH. A summary of suspected adverse reactions should also be included by the MAH in PSURs of the concerned product.

In all situations, investigators of registry studies should be informed of the mechanisms allowing them to report at any time to the national pharmacovigilance system of any adverse event or suspected adverse reaction occurring during the course of the study.

* 1. Reporting of study results

Whilst the ENCePP Code of Conduct (15) and the ADVANCE Code of Conduct for vaccines (28) recommend that the responsibility for preparing the final study report lies within the lead investigator, legal requirements may apply to MAHs for regulatory PASS. For PASS imposed by regulators as a legal obligation, the final study report has to follow the format of the *Guidance for the format and content of the final study report of non-interventional post-authorisation safety studies* (38). This format is also recommended for PASS required in the RMP agreed in the EU or conducted voluntarily in the EU. For a registry study contracted externally by a MAH, the MAH should therefore 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 regulatory requirements. The lead investigator should keep the responsibility for the scientific interpretation of the results.

Following the submission of the final study report, additional information or clarifications may be requested to the MAH by the regulatory authority. The lead investigator should have the duty and scientific responsibility to address the scientific aspects of the request, with the possibility for the MAH to provide comments.

The lead investigator should always have the right to independently prepare publications of the study results irrespective of the source of funding. The MAH which funded the registry study should be entitled to view the final results and interpretations thereof prior to submission for publication and to comment in advance of submission within a reasonable time limit, e.g. one month, and without unjustifiably delaying the publication. The MAH may also require that the presentation of the results be changed to delete confidential information.

1. Conclusions

Patient registries play an important role in medicines regulation as valuable data sources on diseases and their treatments. The EMA Patient registry initiative has been initiated to support the use of high quality data from disease registries. During the pilot phase and in several specific disease–related workshops, it became clear that there was a need for clarifications and recommendations about the expectations from regulators on the type of data to be provided by registries, the acceptable level of quality and the requirements as regards safety reporting. This information is particularly important to understand whether registries can be used by pharmaceutical companies to support regulatory evaluations.

A distinction is made in this paper between methodological recommendations that apply to registries and those that apply to specific studies using registry as a source of data. There are several reasons why this distinction is justified:

* a registry is a routine, long-term, data collection system and a registry study is a time-limited investigation of a specific research question; different objectives require different methods;
* the study team is generally different from the registry group; while a registry coordinator may also take the role of the principal investigator for a study, a study may require additional expertise such as statistical expertise;
* a registry may contribute to regulatory evaluations outside the context of a study; it can provide routine data to regulatory authorities and MAHs;
* registries and registry studies are faced with different methodological issues; important challenges to be addressed by registries are quality management and organisational aspects linked to long-term retention and follow-up of patients, completeness of information and sustainability; in addition to challenges regarding to data quality and completeness of information, registry studies may be faced with very complex epidemiological and statistical questions, and challenges of registry data analyses that may introduce bias and compromise the validity of the results need to be carefully considered when designing the protocol and SAP of registry studies.

The experience has shown that registries are particularly useful to monitor treatments of rare diseases. A single registry with limited size is therefore unlikely to provide enough data for precise analyses and collaborations between registries of patients with a same disease is encouraged. For this reason, emphasis has been put on adoption of a core list of common data collected by all registries, of common terminologies and of common protocols to access the data. Routine collection of a common set of data elements will facilitate initiation of studies and use of the data in the best interest of patients and the patient community. In this respect, clarity is needed regarding data ownership, including patients’ wishes regarding the use of their data.

Sustainability of disease registries has been a common issue discussed in the course of the EMA Patient registry initiative. Studies conducted with registry data may provide an additional source of funding from the public or the private sector. Governance principles are therefore proposed to facilitate interactions between all parties concerned while preserving the registry participants’ scientific independence. For this aspect, quality management is an important activity to provide confidence in the quality of the data that can be generated.

The EMA is willing to support interactions and provide tools to facilitate recognition of disease registries as data sources to conduct studies for regulatory purposes. The Scientific Advice procedure on study protocols and the Qualification procedures of registries are available to pharmaceutical companies and registry coordinators to provide advice and opinions on the validity of registries and study protocols.

1. References

1. European Medicines Agency (EMA). About Us. EMA/338312/2016 Rev. 1 [Internet]. 2017. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Other/2016/08/WC500211862.pdf

2. Bouvy JC, Blake K, Slattery J, De Bruin ML, Arlett P, Kurz X. Registries in European post-marketing surveillance: a retrospective analysis of centrally approved products, 2005-2013. Pharmacoepidemiol Drug Saf. 2017 Dec;26(12):1442–50.

3. Jonker CJ, van den Berg HM, Kwa MSG, Hoes AW, Mol PGM. Registries supporting new drug applications. Pharmacoepidemiol Drug Saf. 2017 Dec;26(12):1451–7.

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

5. European Medicines Agency (EMA). Patient Registries Workshop, 28 October 2016 - Observations and recommendations arising from the workshop. EMA/69716/2017 [Internet]. 2017. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Report/2017/02/WC500221618.pdf

6. European Medicines Agency (EMA). Report on Cystic Fibrosis Registries - Workshop 14 June 2017. Patient Registries Initiative. EMA/510601/2017 [Internet]. 2017. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Report/2017/10/WC500236631.pdf

7. European Medicines Agency (EMA). Report on Multiple Sclerosis Registries - Workshop 7 July 2017. Patient Registries Initiative. EMA/548474/2017 [Internet]. 2017. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Report/2017/10/WC500236644.pdf

8. European Medicines Agency (EMA). Report on CAR T-cell therapy Registries - Workshop 9 February 2018. Patient Registries Initiative. EMA/204454/2018 [Internet]. 2018. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Report/2018/05/WC500249247.pdf

9. European Medicines Agency (EMA). Report on Haemophilia Registies - Workshop 8 June 2018. Patient Registries Initiative. EMA/487643/2018 [Internet]. Available from: https://www.ema.europa.eu/documents/report/report-haemophilia-registries-workshop\_en.pdf

10. European Medicines Agency (EMA). Qualification opinion - The European Cystic Fibrosis Society Patient Registry (ECFSPR). EMA/CHMP/SAWP/802259/2017 [Internet]. Available from: https://www.ema.europa.eu/documents/regulatory-procedural-guideline/qualification-opinion-european-cystic-fibrosis-society-patient-registry-ecfspr\_en.pdf

11. European Medicines Agency (EMA). Draft qualification opinion on Cellular therapy module of the European Society for Blood & Marrow Transplantation (EBMT) Registry. EMA/CHMP/SAWP/423488/2018 [Internet]. Available from: https://www.ema.europa.eu/documents/regulatory-procedural-guideline/draft-qualification-opinion-cellular-therapy-module-european-society-blood-marrow-transplantation\_en.pdf

12. Zaletel M, Kralj M, Magajne M, Doupi P. Methodological Guidelines and Recommendations for Efficient and Rationale Governance of Patient RegistriesMetka Zalatel. Eur J Public Health [Internet]. 2015 Oct 1 [cited 2018 Aug 6];25(suppl\_3). Available from: https://academic.oup.com/eurpub/article/25/suppl\_3/ckv169.006/2578054

13. Gliklich R, Dreyer N, Leavy M, eds. Registries for Evaluating Patient Outcomes: A User’s Guide. Third edition. Two volumes. (Prepared by the Outcome DEcIDE Center [Outcome Sciences, Inc., a Quintiles company] under Contract No. 290 2005 00351 TO7.) Rockville, MD: Agency for Healthcare Research and Quality. [Internet]. 2014. Available from: https://effectivehealthcare.ahrq.gov/topics/registries-guide-3rd-edition/research

14. ENCePP, European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. ENCePP Guide on Methodological Standards in Pharmacoepidemiology [Internet]. Available from: http://www.encepp.eu/standards\_and\_guidances/methodologicalGuide.shtml

15. European Medicines Agency (EMA). The ENCePP Code of Conduct For Scientific Independence and Transparency in the Conduct of Pharmacoepidemiological and Pharmacovigilance Studies. EMA/929209/2011 [Internet]. 2018. Available from: http://www.encepp.eu/code\_of\_conduct/documents/ENCePPCodeofConduct.pdf

16. Norwegian Patient Registry [Internet]. Available from: https://helsedirektoratet.no/english/norwegian-patient-registry

17. Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol. 2015;7:449–90.

18. Swedish National Patient Registry [Internet]. Available from: http://www.socialstyrelsen.se/register/halsodataregister/patientregistret/inenglish

19. Pacurariu A, Plueschke K, Olmo CA, Kurz X. Imposed registries within the European postmarketing surveillance system: Extended analysis and lessons learned for regulators. Pharmacoepidemiol Drug Saf. 2018 Jul;27(7):823–6.

20. European Medicines Agency (EMA). Guideline on good pharmacovigilance practices (GVP) Module VIII – Post-authorisation safety studies (Rev 3). EMA/813938/2011 Rev 3\* [Internet]. 2017. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Scientific\_guideline/2012/06/WC500129137.pdf

21. European Medicines Agency (EMA). Guideline on good pharmacovigilance practices (GVP) Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2) EMA/873138/2011 Rev 2\* [Internet]. 2017. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Regulatory\_and\_procedural\_guideline/2017/08/WC500232767.pdf

22. European Medicines Agency (EMA). Scientific guidance on post-authorisation efficacy studies. EMA/PDCO/CAT/CMDh/PRAC/CHMP/261500/2015 [Internet]. 2016. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Scientific\_guideline/2016/12/WC500219040.pdf

23. Pacurariu A., Plueschke K, McGettigan P, Morales D R, Slattery J, Vogl D, et al. Electronic healthcare databases in Europe: descriptive analysis of characteristics and potential for use in medicines regulation. BMJ Open 2018;8:e023090. doi:10.1136/bmjopen-2018-023090

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

25. SNOMED International. SNOMED CT. The Global Language of Healthcare. [Internet]. Available from: https://www.snomed.org/snomed-ct

26. NIH. U.S. National Library of Medicine. SNOMED CT to ICD-10-CM Map [Internet]. Available from: https://www.nlm.nih.gov/research/umls/mapping\_projects/snomedct\_to\_icd10cm.html

27. NIH. U.S. National Library of Medicine. Unified Medical Language System® (UMLS®) [Internet]. Available from: https://www.nlm.nih.gov/research/umls/rxnorm/

28. Kurz X, Bauchau V, Mahy P, Glismann S, van der Aa LM, Simondon F. The ADVANCE Code of Conduct for collaborative vaccine studies. Vaccine. 2017 Apr 4;35(15):1844–55.

29. European Medicines Agency (EMA). European Medicines Agency policy on access to documents. POLICY/0043. EMA/729522/2016 [Internet]. 2017. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Other/2017/02/WC500221814.pdf

30. EU General Data Protection Regulation (GDPR) [Internet]. Available from: https://www.eugdpr.org/

31. European Commission. DIRECTIVE 2001/83/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 6 November 2001 on the Community code relating to medicinal products for human use [Internet]. Available from: https://ec.europa.eu/health//sites/health/files/files/eudralex/vol-1/dir\_2001\_83\_consol\_2012/dir\_2001\_83\_cons\_2012\_en.pdf

32. European Commission. COMMISSION IMPLEMENTING REGULATION (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 of the European Parliament and of the Council and Directive 2001/83/EC of the European Parliament and of the Council (Text with EEA relevance) [Internet]. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:159:0005:0025:EN:PDF

33. European Commission. REGULATION (EC) No 726/2004 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 31 March 2004 laying down Community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency (Text with EEA relevance) (OJ L 136, 30.4.2004, p. 1) [Internet]. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2004R0726:20120702:EN:PDF

34. ENCePP, European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. ENCePP Resources Database [Internet]. [cited 2018 Aug 28]. Available from: http://www.encepp.eu/encepp/resourcesDatabase.jsp

35. ENCePP, European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. ENCePP Checklist for Study Protocols. Revision 4 [Internet]. Available from: http://www.encepp.eu/standards\_and\_guidances/checkListProtocols.shtml

36. European Medicines Agency (EMA). Guidance for the format and content of the protocol of non-interventional post-authorisation safety studies EMA/623947/2012 [Internet]. 2012. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Other/2012/10/WC500133174.pdf

37. European Commission. The Rules Governing Medicinal Products in the European Union. Volume 10 - Guidance documents applying to clinical trials. Questions & Answers. Version 11.0 (May 2013). Draft – Submitted for discussion to the ad-hoc group for the development of implementing guidelines for the “Clinical Trials Directive” 2001/20/EC.

38. European Medicines Agency (EMA). Guidance for the format and content of the final study report of non-interventional post-authorisation safety studies. EMA/48663/2013 [Internet]. 2013. Available from: http://www.ema.europa.eu/docs/en\_GB/document\_library/Regulatory\_and\_procedural\_guideline/2013/01/WC500137939.pdf