ENGLISH | DECEMBER/2022

Real-World Data (RWD) Real-World Evidence (RWE)

Review Guidance to Support Regulatory Decision

Abbreviations

ANVISA	Agência Nacional de Vigilância Sanitária
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
FDA	Food and Drug Administration
GPP	Good Pharmacoepidemiology Practices
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
RCT	Randomized Clinical Trial
RDC	Resolução da Diretoria Colegiada
RWD	Real-world data
RWE	Real-world evidence

Glossary

- **Clinical Trials -** research study conducted in humans with the aim of discovering or confirming the clinical and/or pharmacological effects and/or any other pharmacodynamic effects of the investigational drug and/or identifying any adverse reaction to the investigational drug and/or studying the absorption, distribution, metabolism and excretion of the investigational drug to verify its safety and/or efficacy. Therefore, its is considered an interventional type of study since the principal investigator assigns an interventional and placebo groups.
- **Observational Studies.** Observational studies are non-interventional clinical study designs that are not considered clinical trials. It is a type in which participants may receive diagnostic, therapeutic, or other types of interventions, but the investigator does not assign participants to specific interventions (as in an interventional study
- **Patient registry (synonym: registry):** Organized system that collects uniform data (clinical and other) to identify specified outcomes for a population defined by a particular disease, condition or exposure. The term 'patient' highlights the focus of the registry on health information. It is broadly defined and may include patients with a certain disease, pregnant or lactating women or individuals presenting with another condition such as a birth defect or a molecular or genomic feature.
- **Pragmatic clinical trials:** Clinical trials designed to evaluate the effectiveness of available medicines in real-life routine practice conditions, whereas explanatory clinical trials aim to assess efficacy of investigational medicines.
- **Primary data:** Data collected directly from patients, caregivers, healthcare professionals or other persons involved in patient care, in the context of a clinical study.
- **Real-World Data (RWD):** Data routinely collected related to the patient's health or clinical condition, captured as part of routine care from a variety of sources, such as electronic health records, medical procedure financing (administrative claims) data, patient-generated data.
- **Real-world Evidence (RWE):** Clinical evidence about the use or potential risks/benefits of a drug derived from the analysis of real-world data.
- **Secondary Use data:** Use of existing data for

a different purpose than the one for which it was originally collected. (Guideline on registry-based studies from EMA)

- **Data curation:** management activities related to data collection, storage, and maintenance among other actions to ensure data is procured and maintained appropriately.
- **Data transformation:** Process to change the format, structure or values of data, in order to prompt them for analysis.
- **Registry-based study:** Investigation of a research question using the data collection infrastructure or patient population of one or several patient registries. (Guideline on registry-based studies from EMA).

Contextualization

Introduction

On the last years, regulatory agencies, such as FDA (Food and Drug Administration) and EMA (European Medicines Agency) started the use of evidence from clinical practice to support regulatory decisions, and terms as Real World Data or RWD and Real World Evidence or RWE became more familiar. Although the concepts and their analysis are used for a long time in pharmacoepidemiology, their use for regulatory decision-making is evolving.

Real-world studies are already included in regulatory submissions to ANVISA (Agência Nacional de Vigilância Sanitária or Brazilian Health Surveillance Agency), although the Agency recognizes that there were no specific local guidelines for its usage. Real-world studies were more commonly used in contexts such as incidence and prevalence of diseases information on target populations, as a comparator of standards of care in clinical studies, and drug performance assessment in clinical practice before a formal and planned evaluation of their effectiveness.

Assessing the efficacy and safety of therapeutic interventions through randomized controlled trials (RCTs) is still considered the gold standard for generating evidence needed for regulatory decision-making. However, some factors intrinsic to the design of RCTs can limit the generation of evidence, such as:

- Strict selection criteria, which reduces the external validity of findings.
- For certain conditions, the randomized study design may not be feasible.
- The duration of an RCT is not always sufficient for an adequate assessment of the long-term treatment effect or to identify rare side effects.
- Randomized studies for populations with specific diseases could not always be possible due to difficulty in recruiting patients.
- RCTs are usually more time consuming than Real World studies.

With RWE, it became possible to understand the effects, risks and benefits of clinical management in a broader context, with different factors and variables, and several actors can benefit from the use of RWE: patients and medical teams, patient associations, sectors responsible for decision-making in the healthcare area, in the public and private spheres, pharmaceutical industries.. In this context, RWD/RWE emerged to bring complementary data, which cannot be obtained through traditional RCTs, but can bring greater robustness to the evidence of safety and effectiveness of health technologies. Thus, RWE can be used to assist regulatory approval of a new drug applications, new indication or to expand the indication of an already marketed drug or to assess post-marketing safety.

Like all evidence used in regulatory decisions, RWE must be valid and of good quality. In the same way as RCTs, studies using RWD also have limitations/ barriers that need to be considered: the quality of sources and data generated; the use of adequate analytical methodologies that can minimize the biases of a non-randomized study.

The regulatory environment for RWD is rapidly evolving, and major international authorities (e.g US FDA, EMA, NMPA) have published final guidelines on the subject. Therefore, this document can suffer changes to incorporate the adoption of additional international guidelines/ documents.

In this context, as an important step of the process it is recommended that, in case of intention of RWE submission in the regulatory setting, a pre-submission discussion of the sponsor with the regulatory agency is crucial for alignment of stakeholders, for any kind of usage, such as providing supporting data in a new indication of an already approved medicine, renewal process of a medicine that requires further clinical evidence, or to answer a scientific question or support a regulatory decision related to the approval of a new medicine.

The General Data Protection Law (Law n. 13,709, of August 14, 2018) represents a milestone in the regulation of data collection, data processing, and data storage for healthcare researchers in Brazil, across institutions and companies, which may be taken into consideration in the planning for RWE submissions to ANVISA.

Real-world data can be derived from a variety of sources, such as:

- Patient records, including electronic medical records;
- Hospitalization data;
- Medical prescriptions;
- Claims data;
- Product and disease registries;
- Non-interventional / observational studies;
- Pragmatic trials;
- Medical devices for home use;
- Digital health solutions, wearables, biosensors and technology accessories, including those with patient-reported health data;
- Social media.

The basis of this document were guidelines from major regulatory agencies (e.g US FDA, EMA, NMPA) albeit guidelines from other authorities may also be referenced and accepted by ANVISA. Thus, this guideline was developed to guide the regulated sector in Brazil for optimal planning and criteria for technical assessment of real-world studies submitted to ANVISA.

Use of real-world data to support regulatory approvals

As discussed before, while RCTs may be the gold standard for demonstrating the efficacy and safety of a medication, more and more regulatory agencies around the world are considering RWE as a complementary basis for supporting regulatory decisions. In addition to the FDA and EMA, other agencies such as Health Canada and Japan's Pharmaceuticals and Medical Devices Agency (PMDA) also have cases of RWE-based regulatory approvals (PETRACCI; GHAI; PANGILINAN; SUAREZ et al., 2021). Following this trend, countries such as China and Taiwan are ahead of other economies around the world and already have published RWE guides, having generated the largest portion of RWE publications among them during

the period from 2015 to 2019 (FOOD AND DRUG ADMINISTRATION, 2019).

In 2019, the FDA circulated a draft guide entitled Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biological Products, with the aim of bringing more flexibility to the type of evidence needed to support effectiveness of products, considering alternative study designs, such as conducting single-arm studies using evidence as external controls (12). More recently, in September 2022, FDA issued the guidance for industry for submitting documents using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products (FOOD AND DRUG ADMINISTRATION, 2022b). In Europe, in 2017, the European agency, EMA, launched the Task Force on Big Data in order to explore the opportunities and challenges of using big data for regulatory decision-making (EUROPEAN MEDICINES AGENCY, 202). The EMA defines big data as a set of extremely large data that can be complex, multidimensional, unstructured, and heterogeneous, that accumulates rapidly, and that can be analyzed computationally to reveal patterns, trends and associations. These data sources include the RWD. In 2018, a European public-private consortium, the Innovative Medicines Initiative, launched the GetReal Initiative, as a way of facilitating the adoption and implementation of RWE for regulatory decisions, and boosting the adoption of tools, methods and good practices in quality in the generation and use of these data (INNOVATIVE MEDICINES INITIATIVE, 2020).

A recent published study on the contribution of Real-World Evidence in European Medicines Agency's Regulatory Decision Making aimed to (i) characterize RWD/RWE presented by applicants to support claims on medicines' efficacy within initial marketing authorization applications (MAAs) and extension of indication applications (EoIs), and (ii) analyze the contribution of RWD/RWE to regulatory decisions on medicines' benefit–risk profile. RWD/RWE was included to support efficacy in 32 MAAs and 14 EoIs submitted 2018–2019. Of these, RWD/RWE was part of the preauthorization package of 16 MAAs and 10 EoIs, and was (i) considered supporting the regulatory decision in 10 applications (five MAAs, five EoIs), (ii) considered not supporting the regulatory decision in 11 (seven MAAs, four EoIs), and (iii) not addressed at all in the evaluation of 5 applications (four MAAs, one EoI). Common limitations of submitted RWD/RWE included missing data, lack of representativeness of populations, small sample size, absence of an adequate or prespecified analysis plan, and risk of several types of bias. The suitability of RWD/RWE in a given application still requires a case-by-case analysis considering its purpose of use, implying reflection on the data source, together with its assets and limitations, study objectives and designs, and the overall data package issued. Early interactions and continuous dialogues with regulators and relevant stakeholders are key to optimize fit-for-purpose RWE generation, enabling its broader use in medicines development (BAKKER; PLUESCHKE; JONKER; KURZ et al., 2022).

Therefore, although the use of RWD has great potential to serve as a complementary source of evidence in the regulatory context, its use depends not only on the evaluation of the methodologies used, but also on the reliability of the data, requiring high quality control in the collection, maintenance, infrastructure and treatment. Likewise, it is important to consider the relevance of the data used, which must be adequate to answer the regulatory question (FRANKLIN; GLYNN; MARTIN; SCHNEEWEISS, 2019; FRANKLIN; PLATT; DREYER; LONDON et al., 2022). Assuming that the data are appropriate, and their source is reliable, it is important to design the most appropriate methodology for

the analysis of the outcomes of interest, applying techniques that can minimize potential biases arising from a non-randomized study design. In all studies to generate RWE, it is necessary to observe the characteristics of the patients considered in the analysis of confounding variables and to identify the variables that could be important in terms of the effect identified in the study that were not considered. In situations where the number of variables that can influence the measured outcome is large and interact in complex ways, advanced statistical or machine learning methods can be used to help limit the effect of confounders on the analysis results (FRANKLIN; PLATT; DREYER; LONDON et al., 2022; MARCHENKO; RUSSEK-COHEN; LEVENSON; ZINK et al., 2018).

Objective of development of this document

Considering the increased trends of RWD and RWE usage worldwide and the participation of ANVISA in several discussions on the usage of RWD/RWE on the regulatory perspective, we developed this document in a joint taskforce of Sindusfarma, Interfarma and ABRACRO to propose an initial framework for the application of RWE for regulatory purposes. In order to achieve a comprehensive approach, we developed the framework in a Q&A format. The intention of this document is to provide ANVISA compilated and helpful information to potentially be considered in the process of a guidance/guideline of RWD/RWE for regulatory usage in Brazil.

Discussion points

Application of RWE for regulatory purpose

Criteria for RWE submissions	Consideration for the questions OR Data Source	Reference guide or guidelines
	Application and documentation submitted	
Cover Letter & Rationale for Selecting RWE Study for Regulatory Purpose Is the application complete and adequately identified?	In the cover letter, it should be specified that the application contains RWD/E. The application should include the final RWE study report, the study protocol and any amendments, which may be found in the Appendix. Also, the cover letter should refer to previous interactions with ANVISA relevant to the RWE study and its design. The cover letter must include a summary of the rationale for using RWD/RWE as part of the submission, study design and sources of data, in a similar format as in the Annex 2 of FDA guideline.	START-RWE. (FOOD AND DRUG ADMINISTRATION, 2022b; WANG; PINHEIRO; HUA; ARLETT et al., 2021)
Documentation of Data Management Process Is Real World Data suitable for regulatory use?	The quality and the purpose of real-world data will dictate whether RWD is suitable for regulatory use. For that, the following points should be considered: 1. Whether the RWD are fit for use, relevant and reliable to adequately address the study question (including studies conducted to test validity of important variables and outcomes found in data source); 2. Whether the study design and analyses are fit for purpose, adequate to address potential bias or confounding to generate scientifically robust evidence to answer or help answer the specific regulatory question. 3. Whether the study conduct meets agency´s regulatory requirements (e.g., for study monitoring and data collection) Regardless of a study's interventional or non-interventional design, the evidence submitted by a sponsor in a marketing application to support the safety and/or effectiveness of a drug must satisfy the applicable legal standards for the application to be approved or licensed. Therefore, a well-documented and transparent process with the use of proper discussion articles, good registry practices should be followed. Also, for registry studies, the regulatory context, timelines, study protocol, study population, data collection, data quality, safety reporting and reporting of study results are relevant.	(EUROPEAN MEDICINES AGENCY, 2018; FOOD AND DRUG ADMINISTRATION, 2018; 2021a)
Suitability of Real-World Data Source to Address Regulatory Question Can the study design provide adequate scientific evidence to answer or help answer regulatory questions?	RWD can be used in a variety of ways and be included in different study designs. As with RCTs, RWE studies and study designs have their intrinsic limitations and strengths. Therefore, it is crucial that those are discussed with the agency upfront and documented during the process. Potential for Study Designs Using RWD to Support Effectiveness: **Randomized Designs Using RWD:** There is a promise in the opportunities created by pragmatic clinical trials, including broader inclusion/exclusion criteria and streamlined data collection. However, it´s important to consider the following factors: • What types of interventions and therapeutic areas might be well-suited to routine clinical care settings? • What is the quality of data that can be captured in these settings? Are they captured consistently over time (over the patient disease journey)? • How many patients can be accessed (particularly when outcomes are rare)? • What are the variations inherent in clinical practice? **Non-randomized, Single Arm Trials with External RWE Control:** External controls (e.g., historical controls) are a possible type of control arm in an adequate and well-controlled study. In the past, the external controls have relied on data from past traditional clinical trials, but increasingly, RWD has been used as the basis for external controls. Using external controls has limitations, including difficulties in reliably selecting a comparable population because of potential changes in medical practice, lack of standardized diagnostic criteria or equivalent outcome measures, and variability in follow-up procedures. Furthermore, hybrid RWE and clinical trial placebo/standard of care control arms and pragmatic trial data may emerge as new approaches and must be described in detail, including comparability, ability to aggregate or pool data, prespecified sensitivity analyses, and sample size considerations. **Observational Studies:** Stand-alone observational or RWE studies can be valuable to contextualize clinical trial results, particularly in the case or rare diseases where only smaller, single arm clinical trials are possible. In the context of retrospective observational studies using RWD, the following critical questions should be considered: • What are the characteristics of the data (e.g., contain data on a relevant endpoint, consistency in documentation, lack of missing data) for improving the validity of the study? • What are the characteristics of the study design and analysis that improve the chance of a valid result?	(EUROPEAN MEDICINES AGENCY, 2021; FOOD AND DRUG ADMINISTRATION, 2018; MEDICINES & HEALTHCARE PRODUCTS REGULATORY AGENCY, 2021; XIA; SCHAEFER; SZENDE; JAHN et al., 2019)

Criteria for RWE submissions	Consideration for the questions OR Data Source	Reference guide or guidelines
	Application and documentation submitted	
	a. Can an active comparator improve the chance of a valid result? b. Given potential unmeasured confounders in non-randomized RWD studies, as well as potential measurement variability in RWD, is there a role for non-inferiority designs? • What sensitivity analyses and statistical tests should be pre-specified for observational studies using RWD to generate RWE for effectiveness? In addition to study design and data considerations, transparency about study design and analysis before execution is critical for ensuring confidence in the results (e.g registration within ClinicialTrials.gov or EU PAS register from ENCePP (EU PAS Register (encepp.eu) is highly suggested) The potential lack of up-front transparency, especially in retrospective observational study design and conduct, coupled with the fact that retrospective analyses in electronic datasets can be conducted multiple times relatively inexpensively with varying study design elements, makes it possible to conduct numerous retrospective studies until the desired result is obtained and then submit only favorable results as if they were the result of a single study with a prespecified protocol. Policies should be considered in order to prevent such practices, including recommendations from experts and other stakeholders. To develop robust evidence, the choice of study design should be described in detail and must represent the most appropriate choice evaluated to adequately address the research. Studies should follow the best methodological standards applicable to pharmacoepidemiologic research and the protocol should describe measures foreseen to account for bias and confounding and ensure the internal validity of the study. After identification of the scientific question(s) to be addressed it is important to critically consider the appropriate study design and source to provide the desired answers and perform a feasibility analysis.	
Type of outcomes to be measured What types of outcomes can be measured/considered?	**Patient Reported Outcome** – A measurement based on a report that comes directly from the patient (i.e., study subject) about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else (e.g., Pain scale measurement, counts of events) **Clinician Reported Outcome –** A measurement based on a report that comes from a trained health-care professional after observation of a patient's health condition. Most measures involve a clinical judgment or interpretation of the observable signs, behaviors, or other manifestations related to a disease or condition (e.g., Psoriasis Area and Severity Index, Hamilton Depression scales). **Performance Rated Outcome –** A measurement based on standardized task(s) actively undertaken by a patient according to a set of instructions (e.g. Measures of memory and gait speed). **Observer Reported Outcome –** A measurement based on a report of observable signs, events or behaviors related to a patient's health condition by someone other than the patient or a health professional (e.g. Acute Otitis Media Severity of Symptoms scale)	(FOOD AND DRUG ADMINISTRATION, 2020)
Type of interventions and therapeutic areas appropriate for collecting Real World Data What type of interventions and therapeutic areas are appropriate for collecting RWE?	The use of RWD may be applicable in all distinct therapeutic areas and interventions, however, some factors influence directly in the collection of RWD. In this context, RWD usage in regulatory context may be evaluated case by case, always considering the clinical relevance of the chosen endpoints and study designs. Therefore, RWD use for regulatory purposes cannot be limited to any kind of intervention/ therapeutic area but evaluated case by case.	
Number of patients to be evaluated How to define the number of patients to be evaluated?	The number of patients evaluated may be aligned with the study´s primary objective and subsequent sample size calculation considering the primary endpoint. In a descriptive study an acceptable precision margin needs to be assessed upfront and in comparative studies, the sample size strategy choice will be similar to the ones used in interventional trials. There should be a discussion about the definition of the study population using inclusion/exclusion criteria.	(EUROPEAN MEDICINES AGENCY, 2021; JOHNSTON; LAKZADEH; DONATO; SZABO, 2019)
Limitation in the selection of the control population (changes and/or variations in medical practices) What are the limitations of the control population?	RWE studies have as an important strength the high external validity compared to RCTs. RWE provides an understanding of treatment patterns, therapy choices, patients' characteristics variabilities and outcomes in different settings (and different clinical practices may be included in the RWD source). The clinical setting may have an impact in patient selection and outcomes, and pharmacoepidemiological and statistical methods should address potential bias to normalize and balance the populations included in the study (e.g., use as covariables or variables in propensity score matching, etc.). A description of the underlying data source(s) selected for the study should be provided with details on the clinical setting and any influences in the treatment patterns observed. A justification of the research/study design should include a discussion of strengths and limitations of the study. There are differences in the practice of medicine around the world and between health care systems that may affect the relevance of the data source to the study question. Patients in different types of commercial or government health care payment programs can differ in a range of characteristics, such as age, socioeconomic status, health conditions, risk factors, and other potential confounders. Various factors in health care systems and insurance programs, such as medication tiering (e.g., first-line, second line), formulary decisions, and patient coverage, can influence the degree to which patients on a given therapy in one health care system might differ in disease severity, or other disease characteristics, from patients on the same therapy in another health care system. It is also important to identify whether the data sources cover all populations relevant to the study if those sources are to be used to examine the study hypothesis.	(EUROPEAN MEDICINES AGENCY, 2021; FOOD AND DRUG ADMINISTRATION, 2021d)

Criteria for RWE submissions	Consideration for the questions OR Data Source	Reference guide or guidelines
	Application and documentation submitted	
	As real-world study populations are diverse compared to clinical trials strategies to address underlying differences in the study population and bias to balance groups and make them comparable is a must. In this context, the use of several pharmacoepidemiological and statistical strategies in study design and analyses, such as propensity score matching, inverse probability weighting and multivariable regressions, may be considered and included in detail in the study documentation (e.g., study protocol and statistical analysis plan). **Recommendations:** • The reason for selecting the particular data sources to address the specific hypotheses. Preliminary and feasibility analyses conducted prior to study to evaluate selected data source against alternative sources may be summarized. • Background information about the health care system, including (if available) any specified method of diagnosis and preferred treatments for the disease of interest, and the degree to which such information is collected and validated in the proposed data sources • A description of prescribing and use practices in the health care system (if available), including for approved indications, formulations, and doses. Usage of proper methodology to adjust/balance the groups	
Registry Studies as RWE support for regulatory submission Can Registry Studies be used to support regulatory submission?	Yes. **A registry** is defined as an organized system that collects clinical and other data in a standardized format for a population defined by a particular disease. Registry data collected initially for one purpose (e.g., to obtain comprehensive clinical information on patients with a particular disease) may or may not be fit-for-use for another purpose (e.g., to examine a drug-outcome association in a subset of these patients). The data must have reliability: accuracy, completeness, provenance, and traceability. Sponsors interested in using a specific registry as a data source to support a regulatory decision should meet with the relevant regulatory agency review division before conducting a study that will include registry data. Sponsors should submit protocols and statistical analysis plans for the agency review and comment before conducting an interventional or a non-interventional study when including data from registries The acceptability of registry-based studies as a source of evidence for regulatory purposes depends on several factors related to the specific regulatory assessment procedure for the concerned medicinal product, the characteristics of the concerned registry and the objectives, design and analytical plan of the proposed study. Early consultation with competent authorities is recommended when a registry-based study is proposed to be used and study protocols should be published.	(EUROPEAN MEDICINES AGENCY, 2018; 2021; FOOD AND DRUG ADMINISTRATION, 2021e)
Situations/Applications where RWE can be used for regulatory submissions (not limited to registry data) When can RWE be submitted to ANVISA for regulatory purposes?	Relevant submissions may include RWE used to support study objectives, such as the following: • To support safety and/or effectiveness for a product not previously approved by the agency • To support labeling changes for an approved product, including: - Add or modify an indication - Change dose, dose regimen, or route of administration - Expand the labeled indication of the product to a new population - Add comparative effectiveness information - Add or modify safety information - Other labeling change • To support or satisfy a post marketing requirement (PMR)/post marketing commitment (PMC)	(EUROPEAN MEDICINES AGENCY, 2018; 2021; FOOD AND DRUG ADMINISTRATION, 2022b)

Data management and quality

Point of analysis	Data Source	Reference guide or guidelines
	Application and documentation submitted	
Data management and data quality: presentation and considerations Has the Real-world data management and quality assurance presented in the documentation?	The strength of RWE submitted in support of a regulatory decision depends on the clinical study methodology and the reliability (data accrual and data quality control (data assurance)) and relevance of the underlying data. Data should be selected based on their suitability to address specific regulatory questions. While the reliability assessments consider whether the codes or combinations of codes adequately represent the underlying medical concepts they are intended to represent, the relevance assessment considers whether the data are fit for purpose and include an assessment of whether the data capture relevant data on exposure, outcomes, and covariates. Although different RWD sources will have different strengths and limitations, the selection of appropriate RWD sources should be based on the regulatory question of interest and should be collected and maintained in a way that provides an appropriate level of reliability. Data availability and quality as well as potential biases will vary according to a serious of factors, including disease prevalence, healthcare coverage, levels of care complexity covered by data source (e.g., primary care, secondary care, etc.). Methods and specific measures should be guided by the feasibility analysis and be selected with a view to minimize risk of invalid study results: • The validity of any data cleaning, extraction and transformation processes should be verified and monitored. • Quality checks of the data used in the study should be performed to alert on erroneous, missing or out-of-range values and logical inconsistencies, and trigger prompt data verification and remedial measures if needed. • In studies with primary data collection, the various factors (e.g., limited human or material resources or inadequate training) influencing quality should be identified and addressed to preserve the integrity of the study. To increase the chances of obtaining reliable data, it is necessary to ensure that: • Data are obtained consistently, using clear definitions and pre-established rules, and ensuring that a data dictionary is available. • Periodic assessments of data consistency are carried out to improve the accuracy of results. • Sufficient information about the disease, products of interest, outcomes and confounders is correctly identified and available in the data source. There are many challenges for standardizing data from the real world, such as: (a) variety of sources and inconsistent formats; (b) differences in data sources that capture terminology and formats for representing similar or identical data elements; (c) wide range of methods and algorithms used to create datasets intended to aggregate data, etc. Therefore, any raw data transformation and data curation methods must be detailed in standardized documents and processes. In addition, evidence must be generated transparently and with integrity from planning to conducting and preparing the study report. The methods used in the study should be described in sufficient detail, while protecting patient privacy, to allow for reproducibility of the findings using the same real-world database. The study-specific RWE should be traceable to the source. Study-specific data sets should be made available to reviewers for data quality and audit related purposes. The data collection method applied should clearly be described in the study protocol as it has implications with regards to the potential sources of bias and confounding, adequate retrieval of missing data and safety reporting requirements. Study-specific data quality management should be prospectively defined and implemented with a risk-based approach and may include verification and monitoring of the validity of any data cleaning, extraction and transformation process; quality checks to alert on erroneous, missing or out-of-range values and logical inconsistencies; and the identification of various factors which may influence the quality and integrity of the study. Descriptive or hypothesis driven statistical analysis plan is most defined in separate document in addition to study protocol and to registry protocol. Changes to the pre-specified statistical analysis should be reflected by an amendment to the study protocol. All changes should be presented in the study report. In this context, it is advisable the use of the following key documents to have a well-documented and methodologically robust process: study protocol, statistical analysis plan and data management plan. Therefore, any raw data transformation and data curation methods must be detailed in standardized documents and processes; if the data involves a third party (e.g., data vendor), coordination with that third party on relevant aspects should be carried out.	(EUROPEAN MEDICINES AGENCY, 2018; 2021; FOOD AND DRUG ADMINISTRATION, 2021b; 2022a; GLIKLICH; LEAVY, 2020)
Regulatory requirement framework Does the study meet regulatory requirements (data collection, monitoring, Good Clinical Practices)?	The Health Authority should consider how to examine the regulatory requirements that are applied to data from randomized clinical trials that are integrated into the health care system and observational studies when they are intended to generate RWE for regulatory decision-making. For example, the use of risk-based and central monitoring for clinical trials that are integrated into the health care system. An early decision to be made when designing a RWE study is the source with respect to the collection method: secondary data collection, where the data for the study are already available and extracted from a dataset, and primary data collection, where data are generated for study capturing information of interest directly from patients as they come to the attention of the investigator. In some specific study designs (e.g ambispective cohort), both data collection methods could be combined. This choice has implications for safety reporting and should be clearly specified in the study protocol.	(EUROPEAN MEDICINES AGENCY, 2018; 2021; EUROPEAN NETWORK OF CENTRES FOR PHARMACOEPIDEMIOLOGY AND PHARMACOVIGILANCE, 2022; FOOD AND DRUG ADMINISTRATION, 2018; INTERNATIONAL SOCIETY OF PHARMACOEPIDEMIOLOGY, 2015; STROBE, 2022)

Point of analysis	Data Source	Reference guide or guidelines
	Application and documentation submitted	
	The study protocol should follow the recommendations of standard quality checklists, such as STROBE and Good Pharmacoepidemiology Practices (GPP), or other local templates applicable. The protocol should also provide details on mechanisms put in place to identify and collect missing data as well as strategies to handle with missing data and, if in case of prospective designs, minimize the number of patients lost to follow up. The protocol should provide a sample size calculation based on the primary objective and endpoints and the feasibility of attaining this sample size within the RWD source should also be assessed using conservative assumptions (or a previous feasibility analysis), both in terms of number of patients (considering the inclusion and exclusion criteria) and in terms of duration of follow-up based on assumptions for losses to follow-up. The structure and content of the study protocol should follow the existing regulatory requirements and should apply the best methodological standards, including if applicable those described by the ENCePP Guide on Methodological Standards in Pharmacoepidemiology.	
Procedures Standardization Is there a standardization in the diagnostic criteria? Outcome measures? Patient follow-up procedures?	To efficiently process RWD and submit it for evaluation to health regulatory agency, appropriate data standards are necessary. A data standard is a set of rules about how a particular type of data should be structured, defined, formatted, or exchanged between computer systems. Data standards make submissions predictable and consistent and have a form that an information technology system or a scientific tool can use. To work with RWD across multiple sources, data may need to be put into a common format, sometimes referred to as a common data model (CDM), with common representation (terminologies, vocabularies, coding schemes). Also, in real-world settings the definitions may vary by each data source and therapeutic area. Therefore, the use of clear and direct definitions used for diagnosis, outcomes, patients´ follow-up and others must be very well described in the protocol, to have a reproducible result. As data availability will vary according to data source, it is advisable that those definitions are aligned with published literature and local HCP specialists.	(FOOD AND DRUG ADMINISTRATION, 2018; 2021b)
RWD quality and applicability in external controls How to guarantee RWD quality and applicability in external controls?	External controls (e.g., historical controls) are a possible type of control arm in an adequate and well-controlled study. In the past, the external controls have used data from past traditional clinical trials, but increasingly, RWD has been used as the basis for external controls. Using external controls has limitations, including difficulties in reliably selecting a comparable population because of potential changes in medical practice, lack of standardized diagnostic criteria or equivalent outcome measures, and variability in follow-up procedures. These potential sources of bias and confounding must be considered carefully in external controls using RWE. Furthermore, hybrid RWE and clinical trial placebo/standard of care control arms and pragmatic trial data may emerge as new approaches and must be described in detail, including comparability, ability to aggregate or pool data, prespecified sensitivity analyses, and sample size considerations.	(FOOD AND DRUG ADMINISTRATION, 2018)
Characteristics of the real-world data source that can improve the chance of a valid result What are the characteristics of the real-world data source that can improve the chance of a valid result?	Understanding all aspects of the RWD source: data source used (including how the data are collected, and what biases are involved), quality of the data (including degree of the missingness in key variables and reasons for the missingness); the study design, protocol and the statistical analyses plan and whether that answers the research question; the outcomes used for the study, and whether these are relevant and have been validated, whether the study results are plausible and generalizable to the population of interest. • Analysis of the availability in the registry or other RWD source of the core data elements needed for the planned study period (as availability of data elements may vary over time), including relevant confounding and effect-modifying variables, whether they are mapped to any standard terminologies (e.g., MedDRA, SNOMED-CT) or common data model (e.g OMOP CDM), the frequency of their recording and the capacity to collect any additional data elements or introduce additional data collection methods if necessary; • Analysis of the quality, completeness and timeliness of the available data elements needed for the study, including information on missing data and possible data imputations, risk of duplicate data for the same patient, results of any verification or validation performed (e.g. through an audit), analysis of the differences between several registries available in the network and their possible impact on data integration, description of the methods applied for data linkage as applicable, and possible interoperability measures that can be adopted. **Data quality includes four main components.** • Consistency: the formats and definitions of the variables are consistent over time, across all centers within a registry and across all registries within a network of registries. • Completeness: patient enrolment is maximized; patient attrition is minimized and complete information on a core data set is recorded for all eligible patients with minimization of missing data. • Accuracy: the data available in the registry or other RWD source are a valid representation of patient information available to the health care professional, e.g., data available in medical charts or laboratory test results; Where the registry data are a compilation or duplication of electronic medical records at the point of care, accuracy should rely on a check of the extraction and uploading procedure. • Timeliness: there is a timely recording and reporting of data and data updates, based on their intended use in compliance with an agreed procedure.	(EUROPEAN MEDICINES AGENCY, 2021; GLIKLICH; LEAVY, 2020; MIKSAD; ABERNETHY, 2018; NATIONAL INSTITUTE OF HEALTHCARE EXCELLENCE, 2021)

Point of analysis	Data Source	Reference guide or guidelines
	Application and documentation submitted	
RWE characteristics (study design and analysis) that can improve the chance of a valid result What are the RWE characteristics of the study design and analysis that can improve the chance of a valid result?	• Description of processes in place for the identification of adverse events and prompt reporting of suspected adverse reactions occurring during treatments, and capacity to introduce additional processes for their collection and reporting if needed. • Study size estimation and analysis of the time needed to complete patient recruitment for the clinical study by providing available data on the number of centers involved in the registry(-ies) or other RWD source, numbers of registered patients and active patients, number of new patients enrolled per month/year, number of patients exposed to the medicinal product(s) of interest, duration of follow-up, missing data and losses to follow-up, need and possibility to obtain informed consent. • Evaluation of any potential information bias, selection bias with potential RWD sources due to the inclusion/exclusion criteria of centers (e.g., primary, secondary or tertiary care) and patients, potential time-related bias between and within registry(-ies), and potential bias due to loss to follow-up. • Evaluation of any potential confounding that may arise, especially if some data elements cannot be collected or measured. • Analytical issues that may arise based on the data characteristics and the study design. • Any data privacy issues, possible limitations in relation to informed consent and governance related issues such as data access, data sharing and funding source. • Overall evaluation of the suitability of the RWD data source (registry, EHR, or claims) for the specific study, considering any missing information on the above-mentioned aspects To adequately assess the results of a non-interventional or RWE study supporting a marketing application, the agency must be confident that particular data sources or databases were not selected, or that specific analyses were not conducted, to favor a certain conclusion. Therefore, the protocol and statistical analysis plan should be finalized prior to conducting the prespecified analyses listed in the protocol and statistical analysis plan. The sponsor should provide evidence that the protocol and statistical analysis plan were finalized prior to reviewing outcome data of a study and before performing the prespecified analyses. In addition, any revisions to the protocol should be date-stamped, and the rationale for each change should be provided. Sponsors should describe in the study protocol all the data sources accessed when designing the study, as well as results from feasibility evaluations or exploratory analyses of those data sources. Sponsors should provide a justification for selecting or excluding relevant data sources from the study. It is also recommended that sponsors generate audit trails in their datasets that can track access to, and analyses performed on relevant data sources. To ensure transparency regarding the study design, it is highly suggested that sponsors post their study protocols on a publicly available website, such as ClinicalTrials.gov or the web page for the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) for post-authorization studies.	(DREYER; BRYANT; VELENTGAS, 2016; EUROPEAN MEDICINES AGENCY, 2021; EUROPEAN NETWORK OF CENTRES FOR PHARMACOEPIDEMIOLOGY AND PHARMACOVIGILANCE, 2022; FOOD AND DRUG ADMINISTRATION, 2021a; HEALTH CANADA, 2019; HIGGINS JPT, 2022; RECORD ; STROBE, 2022)
	If certain RWD are owned and controlled by third parties, sponsors should have agreements in place with those parties to ensure that all relevant patient-level data can be provided to the agency and that source data necessary to verify the RWD are made available for inspection as applicable. Sponsors should ensure that RWD and associated programming codes and algorithms submitted to regulatory purposes are documented, well-annotated, and complete, which would allow the agency to replicate the study analysis using the same dataset and analytic approach.	
Which sensitivity analysis and statistical diagnosis should be pre-specified for observational studies using Real World Data to generate Real World Evidence for efficacy?	Sensitivity analyses should explore the robustness of estimates on primary objectives and analyses of interest to deviations from underlying assumptions and limitations in the data. The ENCePP Guide on Methodological Standards in Pharmacoepidemiology presents methods to address vias and adjust for confounding relevant to observational, RWE, or non-interventional studies.	(BAUMFELD ANDRE; REYNOLDS; CAUBEL; AZOULAY et al., 2020; EUROPEAN MEDICINES AGENCY, 2021)

Conclusions and future directions

This was the first suggestion of RWE Framework for regulatory submission proposed in the Brazilian context. This is a dynamic and evolving topic across the world and across regulatory agencies, therefore, it is highly recommended that this document is on constant review and update (if needed), in annual fashion based on published literature and practical examples (cases) and tailored to the Brazilian regulatory landscape.

References

BAKKER, E.; PLUESCHKE, K.; JONKER, C. J.; KURZ, X. et al. Contribution of Real-World Evidence in European Medicines Agency's Regulatory Decision Making. Clin Pharmacol Ther, Oct 17 2022.

BAUMFELD ANDRE, E.; REYNOLDS, R.; CAUBEL, P.; AZOULAY, L. et al. Trial designs using real-world data: The changing landscape of the regulatory approval process. Pharmacoepidemiology and Drug Safety, 29, n. 10, p. 1201-1212, 2020.

BOLISLIS, W. R.; FAY, M.; KUHLER, T. C. Use of Real-world Data for New Drug Applications and Line Extensions. Clin Ther, 42, n. 5, p. 926-938, May 2020.

DREYER, N. A.; BRYANT, A.; VELENTGAS, P. The GRACE Checklist: A Validated Assessment Tool for High Quality Observational Studies of Comparative Effectiveness. J Manag Care Spec Pharm, 22, n. 10, p. 1107-1113, Oct 2016.

EUROPEAN MEDICINES AGENCY. Big Data Steering Group (BDSG): 2020 report. : EMA 202.

EUROPEAN MEDICINES AGENCY. Discussion paper: Use of patient disease registries for regulatory purposes – methodological and operational considerations. : EMA 2018.

EUROPEAN MEDICINES AGENCY. Guideline on registry-based studies. : EMA 2021.

EUROPEAN NETWORK OF CENTRES FOR PHARMACOEPIDEMIOLOGY AND PHARMACOVIGILANCE. ENCePP Guide on Methodological Standards in Pharmacoepidemiology. 2022.

FOOD AND DRUG ADMINISTRATION. Real World Evidence Program. : FDA 2018.

FOOD AND DRUG ADMINISTRATION. Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biological Products Guidance for Industry. : FDA 2019.

FOOD AND DRUG ADMINISTRATION. Clinical Outcome Assessment (COA): Frequently Asked Questions. : FDA 2020.

FOOD AND DRUG ADMINISTRATION. Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Product. 2021a.

FOOD AND DRUG ADMINISTRATION. Data Standards for Drug and Biological Product Submissions Containing Real-World Data Guidance for Industry. : FDA 2021b.

FOOD AND DRUG ADMINISTRATION. https://www.fda.gov/drugs/news-events-human-drugs/fda-approves-new-use-transplant-drug-based-real-world-evidence. 2021c. Disponível em: https://www.fda.gov/drugs/news-events-human-drugs/fda-approves-new-use-transplant-drug-based-real-world-evidence.

FOOD AND DRUG ADMINISTRATION. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision Making for Drug and Biological Products. 2021d.

FOOD AND DRUG ADMINISTRATION. Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products Guidance for Industry. : FDA 2021e.

FOOD AND DRUG ADMINISTRATION. Study Data Standards Resources. 2022a.

FOOD AND DRUG ADMINISTRATION. Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products. : FDA 2022b.

FRANKLIN, J. M.; GLYNN, R. J.; MARTIN, D.; SCHNEEWEISS, S. Evaluating the Use of Nonrandomized Real-World Data Analyses for Regulatory Decision Making. Clin Pharmacol Ther, 105, n. 4, p. 867-877, Apr 2019.

FRANKLIN, J. M.; PLATT, R.; DREYER, N. A.; LONDON, A. J. et al. When Can Nonrandomized Studies Support Valid Inference Regarding Effectiveness or Safety of New Medical Treatments? Clin Pharmacol Ther, 111, n. 1, p. 108-115, Jan 2022.

GLIKLICH, R. E.; LEAVY, M. B. Assessing Real-World Data Quality: The Application of Patient Registry Quality Criteria to Real-World Data and Real-World Evidence. Ther Innov Regul Sci, 54, n. 2, p. 303-307, Mar 2020.

HEALTH CANADA. Elements of Real World Data/ Evidence Quality throughout the Prescription Drug Product Life Cycl. : Health Canada 2019.

HIGGINS JPT, T. J., CHANDLER J, CUMPSTON M, LI T, PAGE MJ, WELCH VA. Assessing risk of bias in a non-randomized study. In: Cochrane Handbook for Systematic Reviews of Interventions version 6.3: Cochrane, 2022.

INNOVATIVE MEDICINES INITIATIVE. GetReal initiative. 2020. Disponível em: https://www.imigetreal.eu/GetReal-Initiative.

INTERNATIONAL SOCIETY OF PHARMACOEPIDEMIOLOGY. Guidelines for Good Pharmacoepidemiology Practices (GPP). : ISPE 2015.

JOHNSTON, K. M.; LAKZADEH, P.; DONATO, B. M. K.; SZABO, S. M. Methods of sample size calculation in descriptive retrospective burden of illness studies. BMC Med Res Methodol, 19, n. 1, p. 9, Jan 9 2019.

MARCHENKO, O.; RUSSEK-COHEN, E.; LEVENSON, M.; ZINK, R. C. et al. Sources of Safety Data and Statistical Strategies for Design and Analysis: Real World Insights. Ther Innov Regul Sci, 52, n. 2, p. 170-186, Mar 2018.

MEDICINES & HEALTHCARE PRODUCTS REGULATORY AGENCY. MHRA guideline on randomised controlled trials using real-world data to support regulatory decisions. : MHRA 2021.

MIKSAD, R. A.; ABERNETHY, A. P. Harnessing the Power of Real-World Evidence (RWE): A Checklist to Ensure Regulatory-Grade Data Quality. Clin Pharmacol Ther, 103, n. 2, p. 202-205, Feb 2018.

NATIONAL INSTITUTE OF HEALTHCARE EXCELLENCE. Data and Analytics Methods and Standards Programme and Implementation Update. : NICE 2021.

PETRACCI, F.; GHAI, C.; PANGILINAN, A.; SUAREZ, L. A. et al. Use of real-world evidence for oncology clinical decision making in emerging economies. Future Oncol, 17, n. 22, p. 2951-2960, Aug 2021.

RECORD. Reporting of studies conducted using observational routinely-collected data.

STROBE. STROBE - Strengthening the reporting of observational studies in epidemiology. 2022. Disponível em: https://www.strobe-statement.org/.

WANG, S. V.; PINHEIRO, S.; HUA, W.; ARLETT, P. et al. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. BMJ, 372, p. m4856, Jan 12 2021.

XIA, A. D.; SCHAEFER, C. P.; SZENDE, A.; JAHN, E. et al. RWE Framework: An Interactive Visual Tool to Support a Real-World Evidence Study Design. Drugs Real World Outcomes, 6, n. 4, p. 193-203, Dec 2019.

Annex I. Example submissions using RWD and/or RWE

Type of Study	Product	Indication	FDA		EMA	
			Approval	Label Extension	Conditional Approval	Approval
Pragmatic	Invega Sustenna	Schizophrenia		✓(2018)		
External Comparators	Bavencio	Metastatic Merkell cell carcinoma	✓(2017) Accelerated		✓(2017)	
	Brineura	Infantile Batten Disease	✓(2017)			✓(2017)
	Yescarta	Diffuse large B-cell lympho-ma	✓(2017)			✓(2018)
	Kymriah	Diffuse large B-cell lympho-ma				✓(2018)
	Omegaven	Cholestasis associated with parenteral nutrition	✓(2018)			
	Blincyto	B-cell precursor acute lymphoblastic leukemia in the 1st/2nd complete remission with MRD ≥ 0.1%		✓(2018)		✓(2019)
	Zalmoxis	Adjuvant treatment in haploidentical hematopoietic stem cell transplantation (HSCT			✓(2016) *Interrupted in 2019	
Observational Studies	Tepadina	Pediatric class 3 beta thalassemia		✓(2017)		
	Lutathera	SSTR-positive (GEP-NETs)	✓(2018)			✓(2017)
	Ibrance	HR +, HER2- Advanced / metastatic breast cancer in men		✓(2019)		
	Soliris	Transfusion-independent nocturnal paroxysmal hemoglobinuria				✓(2017) Extension
	Prograf	Use in combination with other immunosuppressant drugs to prevent organ rejection in adult and pediatric patients receiving lung transplantation.	✓(1994)	✓(2021)		

Source: (BAUMFELD ANDRE; REYNOLDS; CAUBEL; AZOULAY et al., 2020; BOLISLIS; FAY; KUHLER, 2020; FOOD AND DRU

Contains Nonbinding Recommendation

Appendix: sample presentation of information to be included with submissions containing RWD/RWE

The table below represents an example of how sponsors and applicants can identify a submission containing real-world data (RWD)/real-world evidence (RWE) as part of their cover letter accompanying such submissions to FDA.

General Information

Generic/proprietary name of product: _____

Disease/medical condition: _____

Purposes of Using RWD/RWE as Part of the Submission (select all that apply)

☐ To support safety and/or effectiveness for a product not previously approved by FDA

☐ To support labeling changes for an approved product, including:

- ☐ Add or modify an indication
- ☐ Change dose, dose regimen, or route of administration
- ☐ Expand the labeled indication of the product to a new population
- ☐ Add comparative effectiveness information
- ☐ Add or modify safety information
- ☐ Other labeling change — specify: _____

☐ To support or satisfy a postmarketing requirement (PMR)/postmarketing commitment (PMC)

Study Designs Using RWD to Generate RWE (select all that apply)

☐ Randomized controlled trial with pragmatic elements and those using RWD to supplement a control arm

☐ Single-arm trial that uses RWD in an external control arm

☐ Non-interventional (observational) study

☐ Other study design — specify: _____

_____

RWD Sources Used to Generate RWE (select all that apply)

☐ Electronic health records data

☐ Medical claims data

☐ Product, disease, or other registry data

☐ Data from digital health technologies in non-research settings

☐ Other data sources (e.g., questionnaires) that can inform on health status — specify: _____

_____