



HAUTE AUTORITÉ DE SANTÉ

ASSESS

HEALTH TECHNOLOGIES


**METHODOLOGICAL
GUIDE**

Real-world studies for the assessment of medicinal products and medical devices

10 juin 2021

Descriptif de la publication

Titre	Real-world studies for the assessment of medicinal products and medical devices
Method of production	Methodological guide based on a literature review, an inventory of French and international guidelines and the opinions of methodology guidance group experts.
Objective(s)	This methodological guide aims to support real-world studies relating to health products assessed by HAS assessment committees. It aims to provide practical benchmarks relating to methodological aspects to optimise the level of evidence of these studies and confidence in their results.
Targets	Pharmaceutical companies and contract research organisations
Requester	Self-referral
Sponsor	Haute Autorité de santé (French National Authority for Health) (HAS)
Project management	Service évaluation des médicaments (Medicinal Product Evaluation Department)
Literature search	Documentation - Monitoring Department
Authors	<p>Medical, Economic and Public Health Evaluation Division (DEMESP) cross-disciplinary project team: including representatives from the Medicinal Product Evaluation Department (Judith Fernandez, Julie Taransaud, Mathilde Grande), the Medical Devices Evaluation Department (Corine Collignon, Camille Marguerite), the Public Health and Economic Evaluation Department (Anne-Line Couillerot-Peyrondet, Sophie Tchakamian) and the international mission (Irena Guzina, Chantal Guilhaume). An external project officer (Clémentine Nordon) was also recruited for this project.</p> <p>External experts' group: François Gueyffier, Antoine Vanier, Dominique Costagliola, Sarah Zohar, Pauline Chauvin, Driss Berdaï, Emmanuel Oger, Bruno Falissard, Jean-Luc Faillie, Muriel Vray, Benoit Dervaux, Julia Bonastre.</p>
Declaration of interests	The members of the working group communicate their public declarations of interest to the HAS. They may be consulted on the website https://dpi.sante.gouv.fr . They were analysed according to the analysis grid of the HAS guidelines for the declaration of interests and management of conflicts of interest. The interests declared by the working group members were considered to be compatible with their participation in this work.
Validation	Version dated 10 juin 2021
Updating	

This document and its bibliographic reference are available to download at www.has-sante.fr 
Haute Autorité de santé – Service communication et information
5 avenue du Stade de France – 93218 Saint-Denis La Plaine Cedex. Tél. : +33 (0)1 55 93 70 00
© Haute Autorité de santé – juin 2021 – ISBN : 978-2-11-162660-7

Sommaire

Introduction	4
1. Why implement a real-world study?	7
1.1. To describe the conditions of use	7
1.2. To measure the efficacy and risks related to the use of the health product in real-world conditions.	9
1.3. To estimate utility scores	11
1.4. To measure consumption of resources	12
1.5. To measure the organisational impact of the health product	13
2. How to conduct a real-world study for the (re)evaluation by the HAS of a medicinal product or medical device?	14
2.1. Draft a protocol, with the support of a scientific committee	14
2.2. Propose a study design consistent with the research questions identified	15
2.3. Use pre-existing data, especially from the SNDS (National health data system)	16
2.4. Collect good-quality data	16
2.5. Integrate patient reported outcome measures (PROMs)	17
2.6. Guarantee data transparency	17
3. International references and collaborations	19
3.1. Methodological standards	19
3.2. International collaborations and references	19
Table of appendices	21
Literature search	42
Bibliographic references	43
Abbreviations and acronyms	49

Introduction

Context

The Haute Autorité de santé (French National Authority for Health) (HAS) is tasked with scientifically evaluating the value of medicinal products and medical devices (MD)¹ in medical, economic and public health terms. This assessment is intended to inform the decision-making of public authorities with respect to the funding of the health products concerned by the French national health insurance system, as well as to determine the conditions for their correct use, their role in the prevention, diagnosis or treatment strategy, and their efficiency with a view to price negotiations.

Health product assessments by the HAS are performed by specialised Committees: the Transparency Committee (CT), the National Committee for the Evaluation of Medical Devices and Health Technologies (CNEDiMTS), the Technical Committee on Vaccinations (CTV), as well as the Economic and Public Health Evaluation Committee (CEESP) in certain cases. The assessment principles and scientific analysis methods are described in specific documents (1-5).

In the event of assessment requests with a view to reimbursement, these Committees primarily base their evaluation on clinical trials, which are essential to demonstrate the efficacy of the health product, in accordance with the principles of *evidence-based medicine*.

However, the HAS Committees have a long history of using “real-world data”, i.e., data concerning the use, efficacy or safety of a health product derived from sources other than conventional clinical trials. To carry out its mission, enrich and support the assessment of health products, the HAS regularly analyses observational data and also solicits real-world data complementary to clinical trials when these are essential for a re-evaluation. In fact, the strength of clinical trials is also their weakness. By way of illustration, the rigour of the experimental design, which is necessary to demonstrate efficacy, can lead to conditions that are sometimes far removed from clinical practice and call into question the transposability of the results to routine practice. Results obtained over short periods of time may not be compatible with the rapid evolution of technologies and strategies.

While assessment was historically based on clinical trials alone, recent developments show that this model is evolving in HAS assessments. New clinical trial designs in the context of the development of gene therapies and technologies incorporating artificial intelligence with machine learning processes generate uncertainties with respect to the roll-out of these innovative solutions in routine practice and increase expectations relative to real-world studies. The MD itself becomes a tool for collecting information that can be used for study purposes, as long as the regulations on the use of health data are respected (6-8)

These devices offer significant potential, which needs to be fulfilled in the medium and long term. It is now a question of monitoring the maintenance or optimising the performance of the health product in real life, identifying responder patients, understanding, and anticipating toxicities or risks and supporting the organisational impact.

Considering data collected under real conditions of use is therefore of crucial importance in the assessment of health products. By improving their collection, storage, analysis, and transparency, and more generally the confidence that can be placed in their results, the HAS is convinced that the relevance of the assessment of health products for the benefit of patients will only be reinforced.

¹ In this methodology guide, the term “medical device” (MD) covers the medical devices and health products stipulated in articles L. 5211-1 or L. 5221-1 of the French Public Health Code (CSP)

There are methodological challenges because the quality of real-world studies and the scientific validity of their results are crucially important for their optimal consideration in the assessment of health products by the HAS (9).

The HAS has therefore begun updating the methodology guide dedicated to real-world studies (10) within a context of:

- increased accessibility to health data, in particular via the expansion of the Health Data Hub and the evolution of analysis methods expanding the possibilities for real-world comparative studies;
- incorporation of patient and user perspectives in the evaluation of health products, since they have specific knowledge of their disease (in particular, via the collection of patients reported outcome/experience measures (1, 2, 11));
- consideration of the importance of having access to complementary data in addition to data from clinical trials thanks to real-world studies requested by the HAS or conducted on the initiative of the company responsible² for the medicinal product or MD (12, 13).

Guide objectives

The guide uses the terms « real world data » to describe evidence from observational studies carried out in routine practice.

This methodological guide aims to support and assist the implementation of real-world studies relating to health products, with a view to their evaluation by HAS assessment committees. It aims to provide practical benchmarks relating to methodological aspects to optimise the level of evidence of these studies and confidence in their results.

This guide therefore applies to all real-world studies on health products (medicinal products or medical devices), including those requested by HAS assessment committees (“post-registration studies”). It is aimed at all players involved in the design and implementation of real-world studies on health products: manufacturers, contract research companies, as well as professional organisations and academic teams, which are increasingly frequently involved, particularly in the context of public-private partnerships.

This guide focuses on:

- observational or real-world studies, which represent the majority of the additional studies requested by the HAS;
- pragmatic trials, insofar as they are part of a continuum between randomised clinical trials and observational studies.

This document will not tackle non-pragmatic interventional trials (conventional clinical trials), meta-analyses or network meta-analyses, although these may be requested in the context of post-registration studies. In addition, the methodological aspects of real-world data on epidemiology, the burden of the disease or data collected for the purpose of indirect comparisons will not be detailed in this guide, although this type of data can also contribute to HAS assessments.

The guide is divided into three chapters:

² In this guide, for the purposes of simplification, the term company may also refer to manufacturers and distributors.

- the first covers research questions that may be raised during the clinical development of a medicinal product or a medical device and which may justify the implementation of a real-world study;
- the second details the main HAS recommendations for the implementation of high-quality real-world studies;
- and the third specifies the international methodological references to be taken into account when conducting a real-world study.

For the practical arrangements concerning document exchanges with the HAS and internal procedures relative to post-registration studies, dedicated instruction sheets are available on the HAS website:

- [For medical devices](#)
- [For medicinal products](#)
- For health products that are also the subject of requests for additional data from the CEESP, the protocol may be discussed with the HAS' Economic and Public Health Evaluation Department (SEESP). The department may also be contacted for the submission of protocols specific to economic analyses or organisational studies.

This document does not provide a “ready-made formula” that can be applied in all circumstances, regardless of the nature of the request, and should be seen as a methodology guide. It is the manufacturer's responsibility to propose and adapt its own protocol to implement a relevant real-world study with a view to assessment of health products by the HAS. It is not a binding document.

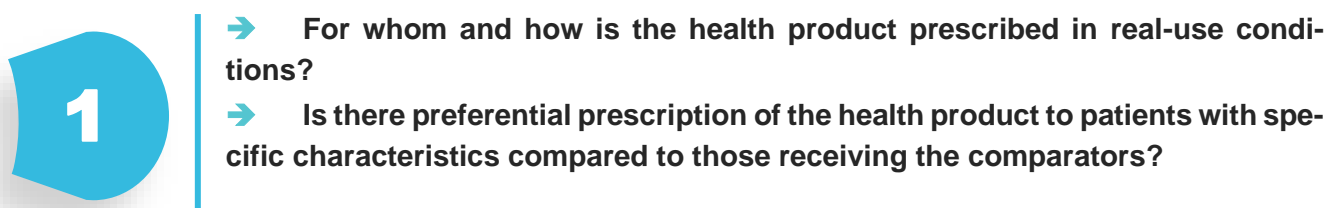
1. Why implement a real-world study?

This section details research questions that may be raised during the clinical development of a medicinal product or a medical device and which may justify the implementation of an observational study with a view to assessment of the health product by the HAS.

It primarily concerns post-marketing data generated in France, but it is highlighted that real-world data are also considered in the first assessments of health products by the HAS, in particular those generated in the context of early access programmes or temporary funding. In addition, it is pointed out that real-world data on epidemiology, the burden of the disease or data collected for the purpose of indirect comparisons can also contribute to HAS assessments.



1.1. To describe the conditions of use



Target population and beneficiary population

The potential difference between the target population (eligible for reimbursement in accordance with the HAS assessment) and the beneficiary population (receiving the health product) is one of the factors of uncertainty in the assessment of health products. There is often a difference between the characteristics of the participants included in conventional clinical trials – themselves often a subset of the target population – and those of patients receiving this health product in routine healthcare practice, i.e., the beneficiary population. In reality:

- the international investigation centres participating in conventional clinical trials are not always representative of French centres treating patients in routine practice; the patients they treat are also different;
- the eligibility criteria for participants in conventional clinical trials are stricter and more restrictive than those of the MA or the population eligible for reimbursement (14-16) or CE marking, raising questions about the representativeness of the population on which the health product efficacy assessment is made compared to the French target population. For example, patients who are very elderly, with comorbidities, taking other medicinal products, or an ongoing pregnancy, etc., are very often not eligible for clinical trials but included within the reimbursement scope;
- The sociodemographic characteristics of patients in clinical trials, particularly the male/female ratio of clinical trials, the age or the ethnicity of patients, do not necessarily reflect those observed in the patients actually receiving or using the health product. Consequently, these data do not always enable analysis of potential differences related to sociodemographic characteristics and/or investigation of the efficacy and safety of a health product on the less represented group (17);
- the patients actually receiving the treatment may do so outside the reimbursement scope (18-21).

Hence it may be useful to characterise the beneficiary population by identifying which patients are treated in routine practice, whether within the reimbursement scope or otherwise. A description of the care pathway – and of the centres treating the patients concerned – may also be expected.

In addition, in routine practice physicians do not prescribe a given health product randomly, as in a clinical trial. They may preferentially prescribe the new health product to certain patients rather than others, based on the patient's history, their previous and current treatments or prognostic factors (related, for example, to age, disease severity, frequency of side effects, etc.). This preferential prescription, which may also be due to the preferences of patients or caregivers, can be investigated in a routine practice study in order to better characterise the beneficiary population. In addition, this propensity of patients to receive a given treatment must be considered when investigating a causal relationship between the health product and the product effect measured in order to minimise bias (indication bias or selection bias).

Conditions of use

The difference between the conditions of use of the health product by patients or of its prescription by physicians between conventional clinical trials, on the one hand, and routine healthcare practice, on the other, represents another level of uncertainty.

When a health product is used by patients themselves, there may be a difference between the conditions of use of the health product (level of compliance and persistence, interruptions) in conventional clinical trials or supervised health product administration – particularly closely monitored and generally for a limited period of time – and use in routine practice. Patients with a low level of treatment compliance are often non-eligible for conventional clinical trials. In addition, during the conduct of the clinical trial, treatment compliance monitoring and improvement measures may be put in place (blood tests, patient reminders, scheduled visits). In contrast, in routine healthcare practice, patients will demonstrate a usual treatment compliance level. Poor treatment compliance may be related to patients themselves and/or the health product (inconvenience, no perceived efficacy or, conversely, a feeling of being better and of being cured, adverse effects, etc.). When the health product is used by a healthcare professional, there may be a difference in the conditions of use between routine practice professionals and clinical trial

professionals, who are more experienced. The impact of the environment, the expertise of user centres, the technical facilities or care organisation are therefore factors that need to be taken into account when analysing the transposability of the results of a clinical trial to routine practice.

As regards prescribers, there may be differences between health product prescribing conditions in conventional clinical trials and those in routine practice. These differences may concern the strength, the number of doses per day, the recommended treatment duration, treatment discontinuations, consideration of interactions with other medicinal products, contraindications and treatment monitoring recommendations. It should be noted that this difference may reflect the need for prescribers to adjust doses in order to optimise the treatment's efficacy/safety ratio – a need that could not be demonstrated during clinical trials.

To address these issues, it is necessary to describe the prescribing and use conditions of health products in routine practice.

1.2. To measure the efficacy and risks related to the use of the health product in real-world conditions.



- ➔ Will the efficacy assessed in clinical trials be the same in real-world conditions? Will the effect size demonstrated in clinical trials be observed in routine practice?
- ➔ Will the benefits and risks as demonstrated in clinical trials be transposable to other contexts: less experienced centres, French patients and healthcare system?
- ➔ Will the impact on quality of life or any other relevant measures for patients demonstrated in clinical trials be the same in real-world conditions?

Due to differences in population, follow-up duration, health product use and prescribing conditions (see above), there may be a difference between the effect of the health product as measured during conventional clinical trials and that actually observed in real-world conditions. This difference may concern:

- the expected beneficial effect in terms of morbidity and mortality (*efficacy-effectiveness gap*) (22);
- the impact of the product on the patient's quality of life or any other relevant measures for patients;
- or the risks and safety of use of the health product.

This difference primarily comes from the fact that certain factors related to the healthcare context or real-world use – not reflected in clinical trials – may modify the specific effect of the product. These context-based factors that modify the health product's specific effect may be related to the healthcare system (for example, level of access to care of patients), prescribing and use conditions (for example, doses actually used or medicinal product interactions, operator-dependent character for medical devices, organisation of care), and/or related to patient characteristics in routine practice (for example, more heterogeneous patients than in clinical trials in terms of comorbidities, age or treatment compliance) (23, 24).

These uncertainties may call into doubt the external validity of clinical trials (25, 26) and require the collection of observational data in real-world conditions.



2

- Will the efficacy of the health product relative to relevant comparators in the strategy in France be of the same magnitude as that demonstrated in clinical trials?
- Will the efficacy, safety of use of the product and/or any other relevant measures for patients compared to the reference strategy confirm the benefit assessed on the basis of non-comparative clinical trials?
- What will the effects of the health product be compared to relevant comparators not considered in the initial assessment, in particular those resulting from concomitant clinical developments?

The acceleration of clinical developments leading to the early marketing of health products on the basis of non-comparative clinical trials or the development of concomitant trials reinforces the need to analyse the relative efficacy and safety of a product to ensure its therapeutic value in real-world conditions of use. The question then arises as to the comparative efficacy of the health product of interest.

The question of the effect of the health product in relation to all relevant comparators also arises when assessing the efficiency of the health product; all relevant comparators may not be included in the economic analysis at the time of the initial assessment, particularly because of the data available. Indeed, the comparator of the health product of interest in a clinical trial may be a placebo, or an active treatment but that is of little relevance in the French healthcare context at the time of the assessment. In addition, indirect comparisons are sometimes submitted in reimbursement applications in order to take into account relevant comparators; however, the comparators retained in indirect comparisons may not be exhaustive with respect to the therapeutic strategy in France.

The objective of a real-world study may be to measure the comparative efficacy of the health product compared to the relevant comparators in the French context. In addition, real-world data may be generated or collected to serve as an external control group for a non-randomised clinical trial. In this case, the implementation of an external comparison must be anticipated and scheduled in advance in order to improve its robustness and ensure it is part of a deductive reasoning approach (27).



2

- **What will be the long-term effects and risks of the health product?**

At the end of a clinical trial, there is little or no knowledge of long-term or rare effects and risks. In fact, the effects of the health product on overall survival, progression-free survival or the onset of all types of events are estimated on the basis of trial durations that are generally inadequate to validate extrapolation hypotheses in economic models, in particular as regards treatment effect maintenance or implant longevity hypotheses. The objective of a real-world study may be to estimate the long-term effects of treatment to validate extrapolation hypotheses made at the time the health product was placed on the market or to quantify revision surgeries or complications. In this case, the aim will be to quantify the impact of these hypotheses on efficacy, safety and efficiency results. The collection of long-term efficacy and/or safety data may also be requested with a view to clinical re-evaluation of a health product, in particular when there are major uncertainties regarding the outcomes of patients/users.

2

→ What is the relevance of the product's effect for patients?

In conventional clinical trials, the primary endpoint must be objective, reproducible, well defined and measured in the same way by all the trial investigators, in order to avoid classification bias. This endpoint is usually based on an objectifiable clinical element (for example, hospitalisation, a clinical event such as myocardial infarction, etc.), or a biomarker directly related to the action mechanism of this element (for example, assay of a tumour marker, cholesterol level) (28). However, for patients, the benefit of a product does not lie solely in the clinical or biological improvement of their disease, especially as these aspects are not necessarily the most relevant from their perspective. Uncertainty about the health product in terms of its relevance to patients may therefore justify the implementation of a real-world study. This uncertainty goes beyond the straightforward perception of symptoms (e.g. pain, fatigue). These expectations and the consequences of the disease on their life may relate to the organisation of their everyday life, fear about the future, the burden of treatment, as well as many other aspects.

Relevant dimensions for patients have long been grouped together under the generic term of “health-related quality of life”, which considers numerous aspects relating to the individual's perception of quality of life affected by their illness or disability: physical, psychological and social aspects. However, other measures may be relevant for patients, such as satisfaction with their care, functional impairment or treatment compliance.

1.3. To estimate utility scores

3

→ What is the quality of life, evaluated using utility scores, of patients treated in real-world conditions?

Another level of uncertainty concerns evaluation of outcomes based on a health status achieved thanks to the health product and described by EQ-5D when these outcomes are evaluated using utility scores³. A lack of documentation of this criterion may be observed in the initial assessment.

In fact, at the time of first marketing, the measurement and evaluation of patients' health statuses may not be robust or may not comply with current recommendations (3), in particular when the pivotal clinical trial does not enable quality of life data to be obtained to evaluate health statuses. This is the case, for instance, when:

- the trial did not schedule the collection of health-related quality of life data using the recommended questionnaire (e.g. EQ-5D-5L in the current version of the HAS economic evaluation guide) (3);

³ In the context of economic evaluations, when health-related quality of life is identified as a major consequence of the health product being assessed, the health outcome used is quality-adjusted life years; health-related quality of life being measured using a utility score (see guide).

- the follow-up period in the pivotal trial, its population or missing data are not compatible with measurement of the utility score associated with certain health statuses;
- the clinical trial produces unexpected findings or with no significant effects on quality of life because the number of participants is too small (rare disease), or because it is impossible to collect data at the time of an event of interest (e.g. acute flare-ups or one-off event).

In these situations, the estimation of utility scores may be non-robust or biased. The literature review conducted to compensate for the lack of data from a trial does not always make it possible to identify a source that is appropriate to the situation and may lead, for example, in the context of the initial assessment, to the use of utility scores valued using a foreign weighting matrix or utility scores measured on a population other than the one being studied, or to the use of strong hypotheses with respect to the equivalence of different health statuses. Finally, intercurrent events, such as adverse events of interest, may not have been taken into account in terms of quality of life losses (disutility).

A real-world study may aim to robustly estimate the utility score of a particular health status, such as a longer-term outcome, or the disutility associated with an event of interest. In routine practice, this involves collecting the perception of quality of life associated with health statuses characteristic of the disease using the EQ-5D-5L questionnaire from a representative sample of the population actually treated in France in order to be able to evaluate them using the weighting matrix for the French population (3).

1.4. To measure consumption of resources



- **Is the measurement of resources consumed in routine practice different from that considered at the time the health product was placed on the market? Are any resources consumed that had not been identified, or, conversely, were there resources consumed in the pivotal trial but not observed in real-world conditions?**
- **If the population treated in real-world conditions is different from the target population considered at the time the health product was placed on the market or if healthcare practices or the context evolve, what is the impact of this difference on the measurement and evaluation of the resources consumed?**

At the time of first marketing, collection of the resources consumed may not cover the full range of resources that will actually be consumed in real-world conditions (3). Indeed, if the data was collected during a clinical trial, the cost items collected may be constrained by the protocol, in terms of scope of analysis (for example, hospital costs only), temporality (schedule constraints, number of scheduled visits), duration of follow-up (failure to take into account the cost associated with the evolution of the health status or disability if the follow-up duration is short) or perspective (for example, cost for the hospital only, and not for the patient or the community). The collection of resources consumed in a clinical trial may also have been carried out in several countries, which raises questions as regards the transposability of this measurement to the French context. The data may also be of poor quality because it is not recorded.

The measurement of resources consumed may also be based on French real-world studies, the collection of which may not be complete (absence of collection of certain cost items, missing data) or which concern an analysis population that is different to variable degrees from the one that will be treated in real-world conditions (beneficiary population).

1.5. To measure the organisational impact of the health product



5

- Are the impacts anticipated at the time a health product is placed on the market real and measurable by criteria in real-world conditions?
- Do the real-world conditions of use reveal organisational impacts that were not identified at the time the health product was placed on the market?

When a health product is placed on the market, organisational impacts are frequently claimed: impacts on the organisation of care, professional practices or patient care conditions. However, they are rarely documented or are often incompletely identified at the time of the initial assessment and only reported in a descriptive manner without any supporting evidence.

The lack of a structured framework for defining the effects of the health product on the healthcare system is probably one of the main reasons for this. In order to take better account of this aspect of the assessment of health products, it is first of all necessary to set out the relevant outlines. With this in mind, the HAS has drawn up a methodology guide (5) proposing a map designed to structure the way in which the organisational impacts of a health product can be identified and supported. To this end, the map proposes a classification made up of macro-criteria and criteria accompanied by examples of indicators.

A real-world study can therefore aim to identify or objectively measure the organisational impacts associated with the marketing of the health product, in particular through the map developed by the HAS.

2. How to conduct a real-world study for the (re)evaluation by the HAS of a medicinal product or medical device?



2.1. Draft a protocol, with the support of a scientific committee

The first step involves drafting a study protocol that must detail:

- the study rationale and objective;
- the hypothesis/hypotheses tested for comparative studies (in particular for calculation of the number of subjects required);
- the source and analysed populations;
- the data sources;
- the study method (definition and measurement of exposure and events of interest, bias minimisation and control methods, statistical analysis plan);
- the study procedures and quality control;
- the expected study limitations;
- ethical aspects.

Any protocol amendments made during implementation of the trial must be notified, duly justified and catalogued.

For drafting and validation of the protocol, it is recommended that a qualified and multidisciplinary scientific committee be put in place, which also includes patient experts and user association representatives.

In addition, an expert committee for the evaluation of events of interest (efficacy criteria and/or adverse events) can be set up.

For more information on the drafting of a real-world study protocol, refer to [appendix 1](#) relative to the international references.

2.2. Propose a study design consistent with the research questions identified

The study design should be determined on the basis of the research questions identified (see part 1 of the guide).

In response to the HAS request for additional data (“post-registration study”), it is stated that the implementation of several studies may be useful or necessary. It is the responsibility of the company or study sponsor to justify the study design chosen and its ability to respond adequately to the request for data made by the HAS. Finally, the methodological limitations and any expected biases should be discussed.

→ A real-world study to document the use of a health product in routine practice

Descriptive, non-comparative observational studies are the study design to be favoured in these situations. They make it possible to describe the beneficiary population, the way a health product is prescribed and used in real-world conditions of use (characteristics of patients and prescribers, therapeutic strategies, arrangements for starting or stopping treatment, potential misuse, etc.), the clinical evolution of exposed patients over time, the quality of life as perceived by the patient or the safety profile of the health product. In line with its 2020 prospective analysis report, the HAS recommends that sex should be taken into consideration when formulating the research project and that the female/male ratio of study patients should reflect that seen in routine practice (17). Furthermore, these studies provide a wealth of information on the context-based factors that potentially interact with the effect of the health product and patient outcomes, but do not enable a causal relationship to be determined.

For more information, refer to [appendix 2](#) relative to descriptive real-word studies and their methodological considerations.

→ A real-world study to compare the efficacy, safety and/or efficiency of a health product with another health product

Comparative studies are the design to be favoured in this case. In addition to being able to characterise the population that uses the health product, these studies can estimate efficacy, safety or any other relevant measure for patients in real-world conditions of use compared to its clinically relevant comparators. These studies can be real-world comparative studies, pragmatic trials or conventional clinical trials. In the absence of randomisation, the HAS recommends implementing measures to minimise indication bias, in particular using causal inference methods in an observational situation, such as adjustment, matching or weighting, when the conditions for applying these methods are met. In some cases, before-and-after comparison studies may be relevant to describe the impact of a change in care.

For more information, refer to [appendix 3](#) relative to comparative studies and their methodological considerations. This appendix also covers predictive modelling aimed at extrapolating present observations to another context.

Other study designs are possible depending on the research questions or data requests formulated by the HAS or other agencies.

2.3. Use pre-existing data, especially from the SNDS (National health data system)

Real-world studies can be based on a new data collection (*ad hoc* study or “primary data”), or on existing French or international databases (“secondary data”).

Although both these approaches are acceptable, the HAS recommends exploiting existing data with the aim of accelerating the availability of results, limiting the duplication of data collection and hence, ultimately, improving the collection of data. France has access to medico-administrative data covering its national territory, as well as cohorts in different diseases of good methodological quality, particularly in rare diseases. Access to this data has been facilitated in recent years and national systems will promote links between data producers and data users in the future (Health Data Hub, France Cohortes, etc.). Initiatives to facilitate the identification of existing databases are under way, such as the [programme of the Alliance pour la Recherche et l’Innovation des Industries de Santé \(ARIIS - French Alliance for research and innovation in the health industries\)](#) and the [Comité Stratégique de Filière Industries et Technologies de Santé \(Strategic Committee of the Health Industries and Technologies network\)](#).

The HAS also highlights the importance of anticipating the need for real-world data for the assessment of health products by setting up disease registries, particularly for rare diseases. These registries can be set up in the context of public/private partnerships and draw support from the rare diseases’ networks.

Finally, it is specified that when the HAS requests additional data (“post-registration study”), this does not necessarily imply the implementation of a *de novo* study. In fact, an observational or interventional study already set up on the initiative of the company, another agency, or a national council for healthcare professionals (e.g: PASS or PAES study for medicinal products) may be liable to address the HAS’ request. When the additional data requested by the HAS require the ad hoc collection of data, the HAS recommends designing this collection process to enable linkage to the SNDS with a view to reuse of this health data by other players. The HAS also recommends storing data generated by post-registration studies on the Health Data Hub.

For more information, refer to [appendix 4](#) relative to possible data sources and their methodological considerations. Refer to [appendix 5](#) for a particular focus on SNDS data with a specific section concerning algorithms for identification of diseases and possible links with other data sources.

2.4. Collect good-quality data

The production of good-quality data is one of the essential prerequisites to ensure study data are taken into consideration by HAS assessment committees. In line with international methodological standards, the following recommendations, in particular, are made:

- document the representativeness of centres, investigators and patients included in the study as well as possible. In accordance with regulatory requirements, the establishment of a non-inclusion registry where possible or the cross-referencing of inclusions with activity or sales data is recommended to justify the representativeness of the study;
- minimise patients lost to follow-up and missing data through adequate data monitoring and quality control. For ad hoc studies, on-site checks should be scheduled to verify the information collected, either on all files or on a random sample;

- unless otherwise requested, ensure that the study will document a beneficiary population that is representative of all the patients treated with the health product in real life;
- to support representativeness, compare the characteristics of the participating investigating centres with those of non-participating centres, those of the patients included with those not included in the study, and compare the characteristics of the patients according to their status (lost to follow-up or not, missing data or not concerning the main variables of interest).

For more information, refer to [appendix 1](#) relative to the international methodological references.

2.5. Integrate patient reported outcome measures (PROMs)

For patients, the benefit of a health product does not lie solely in the clinical or biological improvement of their disease. In fact, these aspects are not necessarily the most relevant from the patients' perspective. Consequently, the HAS encourages the use of outcome measures based on data collected directly by patients, as either secondary or primary endpoints, documenting quality of life, in particular. For post-registration studies, with a request from the HAS for quality of life data, it is recommended to:

1. Incorporate a self-questionnaire (*patient reported outcome measures* - PROMs) collecting data enabling analysis of patients' quality of life and/or all other relevant measures for patients. The validity and interpretation of this questionnaire in the disease investigated must be justified by a literature review. Hypotheses about which quality of life aspects are likely to be impacted first will need to be made in advance.
2. Schedule the use of a "Patient Global Impression of Change"-type question and/or a generic quality of life questionnaire in the absence of a specific self-questionnaire validated in the disease.
3. In accordance with the HAS economic evaluation guide (3) and in the specific context of a cost-utility analysis, favour the use of the EQ-5D-5L questionnaire⁴⁵.
4. The absence of use of PROMs must be systematically justified when submitting a post-registration study protocol.

For more information, refer to [appendix 6](#) relative to measures of interest for patients (*patient reported outcome measures*).

2.6. Guarantee data transparency

The principle of transparency is to systematically make public all information on the study, while ensuring personal data is protected. It helps to build confidence in the players using the data by providing information on the protocol and any amendments made, the data sources and analysis methods used (described in sufficient detail to allow replication), and the results of the study.

It is recommended that the study be registered in one or more public databases, such as the Épidémiologie-France portal (<https://epidemiologie-france.aviesan.fr/>), the ENCePP European EU-PAS registry (<http://www.encepp.eu/encepp/studiesDatabase.jsp>), the American ClinicalTrial.gov registry (<https://clinicaltrials.gov/>), the World Health Organisation's portal

⁴ For clinical assessment, it should nonetheless be noted that the EQ-5D should not be considered to be an instrument measuring the multiple aspects of quality of life.

⁵ For paediatric populations, the use of a generic measurement system developed and validated in children and adolescents is recommended (e.g.: HUI instrument).

(<https://apps.who.int/trialsearch/Default.aspx>) or the PROSPERO portal for literature reviews and meta-analyses (<https://www.crd.york.ac.uk/prospero/>).

As pointed out by the ENCePP code of conduct (29), the protocol must be finalised and wherever possible made public before the study begins. The objective is to be able to guarantee that the results of the study have been produced according to the protocol and the statistical analysis plan initially scheduled, and have not been influenced by analyses carried out during the study. In addition, the scientific publication of the protocol is also possible, and even encouraged by certain scientific journal publishers.

The HAS also recommends that the results be published in a scientific journal and encourages medicinal product and medical device manufacturers to use the Health Data Hub to store real-world study databases with a view to their reuse.

For more information on the publication of study information and data, refer to [appendix 1](#) relative to the international references.

3. International references and collaborations

3.1. Methodological standards

This methodology guide was drafted in accordance with international methodological references for the conduct of real-world studies. Compliance with methodological standards and good practices for the conduct of real-world studies are essential prerequisites to ensure study results are taken into consideration by HAS assessment committees.

In order to conduct a good-quality real-world study, it is therefore necessary to consult good practices for the implementation of studies in routine practice, from definition of the objective through to publication of the results. The main reference guides on protocol development, study conduct (including data collection, data management, quality control and ethical aspects) and on the drafting of result reports are listed in appendix 1 of this document.

3.2. International collaborations and references

Aware of the growth in real-world studies both in France and internationally, the HAS has developed links with other agencies competent in this field, in particular via the *European Network for Health technologies assessment* - EUnetHTA. Created in 2006, this network aims to improve health technology assessment practices and to establish effective and lasting collaboration between the various network members. The work carried out within the network has been on a voluntary basis and within the framework of joint actions, with financial support from the European Commission. The third joint action was launched in 2016, for a period of 5 years.

These cooperation actions are divided into several work packages (WP), including one specifically focusing on improving the quality of data generated during the development and life cycle of health technologies (WP5). This WP, which concerns, *early dialogues*, on the one hand, and *post-authorisation and post launch evidence generation*, on the other, is coordinated by the HAS as part of the joint action under way.

During previous joint actions, several methodological tools and documents were developed:

- criteria for the selection of technologies for which a request for additional evidence may be justified (30) ;
- a database (EVIDENT base), enabling requests for additional studies made by European HTA agencies to be listed (accessible to network partners only (<https://eunetha.eu/evident-database/>));
- two methodological documents (*position papers*) on the formulation of study requests, the first covering the formulation of questions that will be the subject of the additional study (31) and the second the most appropriate choice of study method to answer these questions (32);
- a study protocol template (33).

In 2019, EUnetHTA developed a specific tool for assessing the quality of the registries accessible online: REQueST (*Registry Evaluation and Quality Standards Tool*) (34). Its objective is to provide guidelines for analysing the quality of registries with a view to their use to assess health technologies. In REQueST, the term “registry” is defined as “an organised system that collects, analyses, and disseminates the data and information on a group of people defined by a particular disease, condition, exposure, or health-related service, and that serves predetermined scientific, clinical or/and public

health (policy) purposes”. The tool can therefore apply to any continuous collection of data in a defined population.

In addition, in the joint action currently under way, EUnetHTA has put in place pilot studies on pragmatic collaboration between HTA agencies for the joint definition of requirements in terms of additional data to be collected. Two types of collaboration exist:

- collaborations relating to a specific health product, following missing information identified during its assessment. These collaborations aim to (a) propose a common research question and a shared definition of the essential variables to be collected in this respect, and (b) specify the minimum requirements in terms of design and statistical aspects. The elements defined as a consequence are used as the basis for the implementation of post-registration studies on a national level. Three collaborations of this type were conducted up to 2020. The topics for collaboration (products) were proposed by the network's partner agencies, on the basis of missing information identified in their national assessments and included: an orphan medicinal product in a situation of long-term uncertainty, a medicinal product in metastatic breast cancer to document its use in real-world conditions, and left ventricular assist devices in end-stage heart failure.
- The qualification of registries, which consists in evaluating whether the data collected by the registry and the quality of its collection (evaluated via REQueST since 2019), correspond to the requirements of HTA agencies and can be used as a suitable source of real-world data.

Collaboration leads to the development of guidelines by participating agencies relative to the aspects discussed (variables collected, data quality). Two projects of this type were conducted on the European cystic fibrosis registry and the European bone marrow transplant registry. These two collaborative studies were conducted with the involvement the registry holders and, for the first one, in collaboration with the EMA, and are the subject of reports published on the EUnetHTA website (35).

The HAS also remains attentive to regulatory initiatives concerning real-world data relating to medicinal products and, in particular, the DARWIN EU network (36), as well as the latest European Medicines Agency guideline(37), for which public consultation is ongoing at the time of drafting of the present guide, relative to real-world studies.

Table of appendices

Annexe 1.	International methodological references	22
Annexe 2.	Descriptive real-world studies : methodological principles and consideration	24
Annexe 3.	Comparative real-word studies : methodological principles and consideration	27
Annexe 4.	Data sources	33
Annexe 5.	The Health Data Hub and the SNDS database	35
Annexe 6.	Patient-Reported Outcome Measures	39

Annexe 1. International methodological references

Themes	Sponsor, date	Title	Summary of content	Links, author
Observational studies	ISPE, 2016	Guidelines for good pharmacoepidemiology practice	Guidelines concerning the drafting of observational study protocols, study conduct, communication of results	ISPE, 2015 (38)
Analytical observational studies	STROBE, 2008	The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies	Check-list for drafting results reports for cohort studies, case-control studies and cross-sectional studies	von Elm, et al., 2008 (39)
Observational studies in the field of molecular epidemiology	STROBE-ME, 2011	The STrengthening the Reporting of OBservational studies in Epidemiology--Molecular Epidemiology (STROBE-ME): an extension of the STROBE Statement	Guidelines for epidemiological studies concerning biomarkers	Gallo et al., 2011 (40)
Observational studies Pragmatic trials	ENCePP, 2018	Guide on Methodological Standards in Pharmacoepidemiology (Revision 8)	Regularly updated inventory of methodological standards in pharmacoepidemiology or the design, implementation and communication of results	ENCePP, 2010 (41)
Observational studies	ENCePP, 2018	ENCePP Checklist for Study Protocols (Revision 4)	Check-list for the drafting of observational study protocols	ENCePP, 2018 (42)
Observational studies	ENCePP, 2018	Code of Conduct (Revision 4)	Good practice guidelines on the conduct of studies and the communication of results with the objective of ensuring transparency and scientific independence	ENCePP, 2018 (29)
Comparative effectiveness observational studies	ISPOR-ISPE, 2017	Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making	Good practice guidelines for the conduct and communication of comparative effectiveness observational studies	Berger et al., 2017 (43)
Comparative effectiveness observational studies	AHRQ, 2013	Developing a protocol for observational comparative effectiveness research: A User's Guide	Guidelines for the conduct and publication of comparative effectiveness observational studies	AHRQ, 2013 (44)
Pragmatic trials	CONSORT, 2010	CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials	Check-list for the drafting of clinical trial result reports (including pragmatic trials)	Schulz et al., 2010

				(45)
Medico-administrative database studies	RECORD, 2015	The REporting of studies Conducted using Observational Routinely-collected health Data	Check-list for the conduct of studies on routinely collected databases	Benchimol et al., 2015 (46)
Medico-administrative database studies	ISPOR-ISPE, 2017	Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0	Good practice guidelines for the conduct and communication of results from medico-administrative database studies	Wang et al. 2017 (47)
Medico-administrative database studies	Ateliers de Giens, 2019	Requests for post-registration studies (PRS), patient follow-up in actual practice: Changes in the role of databases	Methodological and procedural guidelines for the conduct of studies on hospital or research medico-administrative databases	Berdai et al., 2018 (48)
Observational studies for the assessment of medical devices	FDA, 2017	Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices.	Guidelines for the incorporation of real-world data in US regulatory decisions (FDA) concerning medical devices	FDA, 2017 (49)
Studies for the assessment of medical devices (not specifically observational studies)	EUnetHTA, 2015	Therapeutic medical devices guideline	Guidelines for the assessment of the comparative clinical efficacy of medical devices	EUnetHTA, 2015 (50)
Organised data collection system	EUnetHTA, 2020	The Registry Evaluation and Quality Standards Tool (REQueST®)	Tool to evaluate the quality of data collection systems in routine healthcare practice (including infrastructure, governance, funding, etc.).	EUnetHTA, 2020 (34)
Systematic review containing randomised or non-randomised studies	AMSTAR 2, 2017	A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both	Tool for the evaluation of systematic reviews of the literature containing randomised and non-randomised studies	Shea et al., 2017 (51)
Observational study meta-analyses	MOOSE, 2000	Meta-analysis of Observational Studies in Epidemiology	Check-list for the drafting of observational study meta-analyses	Stroup, et al., 2000 (52)
Observational studies	STaRT-RWE, 2020	STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies	Guidelines for the publication and presentation of observational studies	Wang et al., 2020 (53)

Annexe 2. Descriptive real-world studies : methodological principles and consideration

Description

There are two broad families of descriptive real-world studies depending on the need or otherwise to follow up patients over time:

- **cross-sectional studies**, which only concern a single data description time-point. These are useful when there is no need for follow-up over time and possible when the characteristics of the beneficiary population or the prescribing/use conditions are relatively stable over time.
- **longitudinal studies (particularly cohort studies)**, in which patients are followed up over time. These are useful for determining, for example, the duration of treatment, patient compliance, reasons for starting and stopping treatment, etc., or for describing a beneficiary population or adverse effects liable to change over time.

Important methodological considerations

In a descriptive real-world study, primary collection or secondary use of data are possible (54) (see appendices [4](#) and [5](#)), depending on the question asked and availability or otherwise of information of interest in a cohort or existing database. For example, descriptive real-world studies on medico-administrative databases ([such as the SNDS](#)) enable assessment of compliance via the dispensing of the treatment by retail or hospital pharmacies (if included on the “additional list”). However, they do not make it possible to assess the actual consumption of the health product, the patient's experience or health status or to determine the clinical monitoring carried out by the physician, excluding laboratory tests or radiological procedures. In the case of analyses relating to implantable medical devices, if several implants with the same LPP code have been used successively to treat different anatomical locations, the possibility of linking the information with the implant of interest must be checked so that a study based on hospitalisation databases can be envisaged (for example, in some cases, the operated side is identified separately; in other cases, and, in particular, when successive treatments of several anatomical levels are involved, the information is not available, which can be an obstacle to an assessment approach based on databases.

In the case of the *ad hoc* conduct of a study, certain methodological recommendations described hereafter need to be taken into consideration from the study planning stage. In the event of secondary data use, compliance with these same aspects needs to be verified.

Selection of the study population

The method for selecting the centres and patients (sampling) included in the real-world study must be the subject of particular attention because the objective will be to describe, as closely as possible to reality, the professionals and patients using the health product in routine practice. Before the study is carried out, several techniques can be used to reduce the risk of selection bias in the study:

- identification of the places of care that best reflect the places where the population of interest is treated (primary, secondary or tertiary care centres?) and possibly their respective proportions, or conversely the active list of these different types of prescribers with regard to the disease of interest. For example, if 20% of patients suffering from a particular disease or condition are treated in hospitals and 80% by general practitioners, it will be necessary to include enough general practitioners to find these percentages in the study;

- identification of possible prescribers or implanters of the health product in real-world conditions (characterisation of practitioners, discipline and practice setting). For example, if 50% of patients with a specific condition are treated by endocrinologists and 50% by rheumatologists, the investigating physicians will need to reflect this diversity.

In addition, to compensate for any selection bias related to the investigator and ensure investigators includes all eligible patients irrespective of their preferences, the consecutive inclusion of patients is necessary.

At the start of the statistical analysis, it is recommended that the presence of any selection bias be assessed by comparing the characteristics of centres and professionals having refused to take part in the study with those of the professionals included. As regards patients, it is sometimes possible to collect a few demographic data (age, sex) of patients having refused to take part in the study, in order to compare these with the patients included. The collection of patient characteristics refusing inclusion should be documented as far as possible.

In certain specific cases, such as in the context of a rare disease, exhaustiveness is important and the study should ideally include all patients affected by the disease. In order to do this, it will be necessary to check that all the departments liable to care for patients with the disease have been approached and that all patients from these departments have been included in the study. Several information sources may be necessary for this purpose.

It should be noted that a selection bias can never be totally excluded in a study requiring consent.

Data collected

→ Observational character of the study

In the context of ad hoc descriptive studies, the sponsor must ensure that the implementation of the study will not change the prescriber's habits and choices. Data collection (via a *Case Report Form*, for example) should enable the collection of relevant data for the study without, however, modifying medical practices.

Furthermore, it might seem suspicious if a study inclusion criterion involves the initiation of, or switch to, the intervention of interest for a significant proportion of patients. This would mean, by the very design of the study, encouraging prescribers to initiate this intervention in order to be able to include patients in the study (*seeding trial*) (55).

→ Level of detail and data granularity

Since the interest of a real-world study is often to provide information that is not available in clinical trials, particular attention must be paid to the scope and exhaustiveness of the information collected (characteristics of healthcare centres, prescribers, patients, health product prescribing and use methods).

To ensure that relevant data will be collected, it is useful to seek advice from clinical experts who are most likely to know the variables of interest to meet the objectives of the study. In the event of secondary data use, it must be ensured that the cohort (registry) or medico-administrative database contains data that is relevant to the study.

→ Data quality

Although patient follow-up is observational in these studies, it is imperative to limit the number of patients lost to follow-up and thus the risk of missing data, particularly with respect to measurement of

the event of primary interest (final patient status). When a patient is lost to follow-up and their final status is unknown, it is impossible to categorise them as having or not having had the main event of interest. If too many patients are lost to follow-up, the results of a study may be invalid or uninterpretable.

The longer the follow-up period, the greater the risk of missing data concerning the final status.

To minimise the risk of patients being lost to follow-up at the time of the study, consideration can be given to motivating patients to participate in the study throughout its duration, for example by sending reminders during the course of the study, using digital tools, or in any other way that should be described in the study protocol.

If the main event of interest is a death, the vital status of patients lost to follow-up must always be scrupulously researched using the civil status records of the patient's municipality of birth or by cross-checking with the French national registry of causes of death (Inserm – Centre d'épidémiologie sur les causes médicales de décès (Centre for the epidemiology of causes of death) – CépiDc).

Finally, in the event of statistical analysis, the number and percentage of patients lost to follow-up must be described and the characteristics of these patients must be compared with those of patients not lost to follow-up.

Annexe 3. Comparative real-word studies : methodological principles and consideration

Description

Cohort studies

Cohort follow-up enables comparative analyses to be conducted in order to determine whether patients “exposed” to the health product, compared to “unexposed” patients (or exposed to the comparator(s)), have a lower, equal or higher relative risk of the event considered, with the possibility of causal inference of the relationship (if appropriate biases, including confounding factors, are controlled) between these two elements.

→ *Case-control studies*

These are based on the constitution of a group of “case” patients presenting a disease/condition being studied (or an event linked to a disease: relapse, hospitalisation, etc.) and a group of “control” patients without this disease/condition. The principle of case-control studies is to compare the frequency of exposure of patients in the two groups. This type of study is suitable in cases where the disease being studied is rare and the exposure frequent. However, the constitution of the control group is a difficult process: the controls must be identical to the cases, except for the disease/condition in question, and must come from the same theoretical cohort as the cases. This type of study makes it possible to calculate an odds ratio (OR), which can approximate the relative risk of the event of interest occurring in exposed subjects compared to unexposed subjects, under certain conditions. This approximation is possible when the disease/condition is uncommon (<10%) and the relative risk value is not too high (<5%). There are a few specific case-control study scenarios:

– **nested case-control studies within a cohort**

The cases and controls are identified within the same cohort defined prior to the study. These studies reduce selection bias in the choice of controls, as well as recall bias, inherent in case-control studies. The major disadvantage of this type of study is the potential non-representativeness of the population of unaffected subjects (from which the controls will be randomly drawn) compared to the initial cohort, due to the loss of patients since the start of the cohort (death or lost to follow-up).

– **case cross-over studies**

This type of study is relevant when it can be assumed that there is a very short delay between the exposure and the event, such as the very short-term effect of a health product. Only patients having had the event of interest are included. The period directly preceding the event is called the “at-risk” period (during which the exposure or otherwise of each patient is measured) and the period more distantly preceding the event is called the “reference” period. Each patient is considered to be his/her own “control”. The exposure ratio during the at-risk period is compared with the exposure ratio during the reference period.

Important methodological considerations

Real-world comparative studies are a potential source of knowledge about the effect of treatments under less controlled conditions than those of conventional clinical trials (43, 56), especially versus comparators that have not been evaluated in conventional clinical trials.

They make it possible to answer questions that conventional clinical trials do not explore, such as the contextual elements that may interact with the treatment effect itself and alter its effectiveness in real

life. However, the implementation of the protocol, the analyses and the interpretation of the results must take into account possible systematic errors (bias) linked to the selection of patients, the measurement of the exposure or the event, or the presence of confounding factors in certain associations that are being sought to be highlighted (57, 58).

→ **Selection biases related to the patients included**

In addition to the selection bias risk described in the paragraph relative to descriptive real-world studies, other selection biases related to the patients included also exist, in particular:

- The **depletion of susceptibles bias**, which is a bias related to the past treatments of patients eligible for the study and which contributes to over-estimation of the benefit of a treatment or to under-estimation of the risk. Indeed, patients who have taken the treatment of interest in the past and stopped it (due to an adverse effect or lack of efficacy, for example) will not be counted as “exposed” to the treatment in question at the start of the study. In other words, the theoretical cohort of patients exposed to the treatment of interest has been “emptied” over time of patients for whom the treatment was not suitable. This risk of bias makes recourse to incident patients (“*new-user design*”), which consists of only including patients taking the medicinal product or using the medical device of interest for the first time, a relevant approach (59).
- The **Healthy user effect**, which corresponds to a positive correlation frequently observed between the quality or quantity of lifestyle behaviours and the propensity to accept or be compliant with a treatment. These behaviours improve their prognosis, unrelated to the treatment itself. It is therefore important to consider these behavioural factors in an observational study.

→ **Classification measurement bias**

In a cohort study, an **immortal time bias** corresponds to the incorporation in analyses of a “treatment exposure” time that would in reality include the period during which the patients in the “exposed” group had not yet received this treatment, i.e., the period between the event having led to the diagnosis (retrobulbar optic neuritis in multiple sclerosis, for example) and the actual initiation of treatment (anti-inflammatories). One of the measures to compensate for this bias is to perform analyses on a “person days” basis. This period should not be taken into account in the number of exposed person days, but added to the number of non-exposed “person days” (60, 61).

In case control studies, a **recall bias** corresponds to the fact that the persons having experienced the event of interest or their families are likely to better recall exposure afterwards than those not having experienced the event.

→ **Attrition bias and missing data**

In an observational study, the strategies put in place to encourage patients to participate right to the end of the study are not always optimal. The risk of losses to follow-up and missing data for the primary measure of interest is therefore greater than in a conventional clinical trial. Similarly, investigators may be less accurate or regular in collecting data - a risk that needs to be minimised.

The presence of a measurement bias related to this missing data depends on the frequency and the reason for the loss of information. When the data for the primary measurement of interest arrives completely randomly and without any relation to patient factors or the effect of the health product, the missing data is said to be “*Missing Completely At Random*”, a situation that is extremely rare. More

generally, the reason for the missing data can be explained by patient characteristics (*Missing At Random*, MAR), or can also be explained by the occurrence of the event of interest. In the latter case, the missing data is informative and described as being “*Not Missing At Random*” (NMAR).

The situation is further complicated if the frequency of missing data is not the same in the two intervention groups.

Ways to reduce the number of missing data include:

- during study conduct: a reduction in the burden related to study participation for patients in order to reduce attrition;
- during the analyses (62): if missing data is not too frequent and if the data is assumed to be MAR, multiple imputation of missing data is possible; for NMAR-type missing data, the imputation techniques are very limited; sensitivity analyses may be useful. A full-case analysis is not recommended, however, except for sensitivity analyses, alternatively considering patients lost to follow-up as treatment successes or failures.

→ **Indication bias and other confounding effects**

In a real-world study, and in contrast with conventional clinical trials, implementation of the study should not should not alter - or as little as possible - the usual course of care and the choice of health product prescribed (63). In routine healthcare practice, the choice made by prescribers between different possible care devices depends on factors linked to the patient (prognostic criteria, (64), comorbidities (65)), to the treatments (adverse effects, past treatments (66)), or to the prescriber's habits. As a result, and due to non-randomisation, these factors are not identical in patients receiving the health product of interest and those receiving the alternatives (comparators). Moreover, these differences tend to vary over time, adding a further level of complexity to the analyses.

Therefore, if the objective of the study is to make a causal inference about the link between “exposure” to one health product (versus another) and the occurrence of an event of interest, these confounding effects (indication bias and *channelling bias*) need to be taken into account in order to correctly measure the effect size and to be able to attribute this difference in effect to the treatment alone. When the weight of the confounding factors varies over time, this is known as *time-varying channelling bias* (67, 68).

Any potential confounding factors should be identified early in the planning of the study through clinical expertise and/or a literature review. It should be borne in mind that not all confounding factors are always observable or known and that their consideration is therefore frequently imperfect. Nevertheless, it is important that observable confounding factors be properly collected in the study, or present in the database used.

As regards statistical analyses, the methods for consideration of confounding factors include - but are not limited to - the use of:

- a multivariate model with matching and/or adjustment for confounding factors;
- a multivariate model with matching and/or adjustment for high-dimensional or simple propensity scores (69);
- a structural marginal model with *inverse-probability of treatment weighing* (70);
- instrumental variables
- double difference methods,
- regression discontinuity designs.

It should be noted that even if the study is comparative, the descriptive analyses are an essential step prior to any inferential analysis. Description of the different groups provides important information. Similarly, it may be instructive to visualise the distribution of propensity scores and the degree of overlap of the propensity score distribution curves (strategy of interest vs. comparator), which indicates the overall degree of similarity of the patients in the two groups with respect to the observed confounding factors, as well as the possibility or not of using the propensity score method.

Pragmatic trials

Description

Pragmatic trials are randomised clinical trials in which the traditional design elements of clinical trials have been made more flexible, in order to more accurately reflect routine practice (71). The concept of a pragmatic approach for clinical trials, as opposed to an experimental approach, was introduced in 1967 by Schwartz and Lellouch (72). The authors' argument was that a clinical trial can answer different questions that require different methodological approaches. The experimental approach is used to demonstrate the pharmacological efficacy of a medicinal product, having controlled all the other factors that might explain the results observed – hence the strict experimental framework. The pragmatic approach, on the other hand, aims to assess the efficacy of the treatment in usual treatment conditions (*effectiveness*). Although pragmatic trials are closer to routine practice, they reflect an experimentation and are not therefore considered to be observational studies. The design components of pragmatic trials have the following characteristics (73, 74):

- centres recruiting patients must reflect the usual places of care;
- the patient eligibility criteria must be broad and not strictly those of the marketing authorisation;
- the randomisation unit may be the patient or the place of care (*cluster-randomised trials*);
- the comparator should reflect the care alternatives and treatments usually prescribed in this indication;
- data collection should be as close as possible to conventional management of the disease/condition concerned and minimise the burden on the prescriber and, wherever possible, use electronic health records (75);
- the primary measure of interest must be as relevant as possible for the prescriber, for example that primarily used to adapt treatment, or for the patient, for example a PRO-type criterion.

The other components relative to the design of the pragmatic clinical trial must also be chosen to best reflect routine practice (monitoring of the intervention, follow-up frequency, treatment monitoring, etc.). The simplest design for a pragmatic trial consists in randomising participants into two intervention groups that will be compared. Other types of design also exist:

- **Cohort-multiple randomised trial:** the patients in an existing cohort are randomised, either to switch treatment (“intervention of interest”) or to stay with their current treatment (comparator group). Only patients randomised into the first group are asked to give their consent for inclusion in the trial, with the others already being followed up in an observational manner. This design reduces the burden of recruitment and consent for at least half of the patients and reduces the costs of the study. An important limitation related to this design is the possibility of refusal in the “intervention of interest” group and not in the other, resulting in a potential selection bias and a loss of statistical and methodological power.
-

- **Cluster-randomised trial:** the randomisation unit is the care centre and not the patient. The patients included in the trial will receive the treatment for which the centre caring for them was randomised. This design reduces contamination bias, insofar as blinding is not wanted in order to remain as close as possible to reality, in accordance with the principles of pragmatic trials (71). The major limitations of this design include: (a) a risk of selection bias linked to the fact that prescribers know which health product is scheduled to be initiated in their patients (76); (b) a loss of statistical power and (c) the possibility of a cluster effect that needs to be taken into consideration via interaction analyses.

Important considerations

Modification of design parameters intended to make a clinical trial more pragmatic may cause certain biases that need to be anticipated and minimised from the study planning stage and taken into account when performing the statistical analyses. These include selection biases related to patient and prescriber preferences, heterogeneity of patients and treatment effect with risk of dilution of the measured effect, measurement bias if the event of primary interest is subjective, etc. (64-66, 77, 78). More importantly, the question posed and answered by the pragmatic trial must be precisely formulated, in order to judge these risks in relation to what precisely is to be measured (“*estimand*”).

Predictive modelling

Predictive modelling aims to extrapolate observations from the present to another context (the future, another population, another country). Clinical trials are conducted in a specific and restrictive context compared to routine practice. For example:

- they are time-limited;
- they are not conducted across the whole country where the health product will be prescribed;
- the population analysed is not identical to the target population;
- the population analysed is not necessarily that in which the health product will have the best impact.

Artificial intelligence methods (for example, *machine learning* such as neural networks), or other statistical modelling methods may be used to supplement the results of a clinical trial or other types of study, in order to predict what the effect of the health product would be in the “new” context of interest (long-term effect, effect in France, effect in the target population, etc.).

In the pharmaceutical industry, these methods are sometimes used to predict the results of the phase 3 trial (79, 80) based on previously available data. The FDA and the EMA are encouraging reflection on the subject of simulations and *in silico* clinical trials to optimise the development of medicines and support their marketing.

Note:

- a distinction must be made between “simulated data” (the data does not exist, it is generated by computer) and simulation studies (regardless of the data, different scenarios are simulated, in order to optimise a prediction method and analyse its sensitivity to different constraints);
- what is known as “*big data*”, which corresponds to very large databases, with in the region of 100 million lines and columns, or more.

General principles

These methods (algorithms, *machine learning*, neural networks, etc.) can use frequentist or bayesian statistics. The models can be configured in a “supervised” manner (i.e. the variable to be predicted is defined in advance, the model seeking to explain this variable) or “unsupervised” manner (the method classifies subjects without any prior choice of the variable to be predicted).

For certain *machine learning* techniques, what is important is the capacity of the model to correctly predict what is happening in the “new” context of interest, using predictive power indicators. However, the interpretation of the coefficients of a model (for example, the regression coefficients) and the clinical meaning given to an observed association may be less important.

Methodological considerations

Machine learning or any other predictive statistic modelling method is conducted in several successive phases:

- the training phase, carried out on all or part of the training database (*training set: base used to specify the predictive model*): several models may be specified;
- the internal validation of the model using the training set;
- the test phase or external validation, conducted on all or another part of the database (*test set*) or on one or more external databases, to assess the model’s capacity to predict situations other than those used for the training phase, thereby making it possible assess any *overfitting* problems (see below), during this phase, the predictive power of the model is evaluated;
- the retraining phase.

Risks and methodological limits

- *Overfitting*: the predictive power of the model on the training set seems to be high, even almost perfect (*overfitting*), i.e., the model presents good or even excellent internal validity, but its external validity (capacity to predict test set events) is low. Therefore this model is not effective. Numerous statistical approaches can be used to specify models, taking into account the risk of overfitting.
- Extrapolation (*coverage of training*): the predictive model tested on another data source (another country, another time period) may prove unsatisfactory, so the model must constantly be adapted.

Annexe 4. Data sources

Primary data

Primary data collection means that the collection of data for the study was planned and carried out as part of an ad hoc study. This data can be collected by investigating physicians, clinical research associates from medical records, or patients (self-questionnaires).

The data collected forms a cohort, or a registry when the data collection is continuous and exhaustive in a geographically defined population (81).

Ad hoc data collection may be necessary if no existing data sources are able to meet the specific objectives of the study to be conducted.

Secondary data

Secondary use of data means that the data used for the study had already been collected for another purpose (82).

The main secondary data sources are:

- medico-administrative data from the SNDS (National health data system);
- data from routine healthcare:
- hospital information systems, in particular hospital data warehouses;
- public or private information systems in community medicine, particularly those based on prescribing aid software;
- data from other community healthcare players (pharmacies, medical biology laboratories and imaging centres);
- data from registries and cohorts;
- data produced by patients, either via connected self-measurement medical devices (heart rate, blood glucose, etc.) or via direct input of information by the patient into applications.

For several years now, the use of secondary data has played an important role in real-world studies, following the opening up of access to the SNDS and the progressive organisation of secondary use of academic cohort data. The organisation of the collection and exploitation of new health data is a growing field and its use for real-world data collection is expected to continue to expand (48).

Considerations relative to the choice of a secondary data source

When choosing a data source for a real-world study, it is important to check the following points (44, 48):

- Does the envisaged data source make it possible to meet the objective of the real-world study? In particular, are all measures of interest collected in the database? In other words, it is necessary to make sure that the variables available in the database enable the precise identification of: the disease or a subtype, relevant clinical elements, quality of life measures, variables measuring potential confounding factors, etc. This type of data is not present in SNDS databases.
- Is the data sought accessible indirectly through algorithms? Are the latter validated?
- Is the data sought accessible and of good quality? The quality of the data collected must be assessed, including metrological criteria such as accuracy, but also involving the assessment of data entry processes, the standardisation or otherwise of their values (including standards and dictionaries and their successive versions), the control processes (on raw files and prepared for exploitation in the real-world study), correction, reclassification, treatment of missing data, etc.;
- Is the representativeness of the data guaranteed?

- Is the timing of exposure to the health product and collection of the data of interest compatible with the timing of the real-world study?
- Is there sufficient transparency with respect to the funding source of the data source; are possible conflicts of interest defined?

Secondary use of academic cohorts and registries

Already established academic cohorts or registries represent a potentially very useful source of data for real-world studies. Several national cohorts have already been used in assessments of medicinal products by the Transparency Committee: the HEPATHER cohort⁶ in re-evaluation opinions relative to direct-acting antivirals in the treatment of hepatitis C, the OFSEP cohort⁷ for the re-evaluation of medicinal products for multiple sclerosis, the CKD Rein cohort⁸ for the assessment of tolvaptan in polycystic kidney disease, the RaDiCo cohort⁹ in idiopathic pulmonary fibrosis. The Epidémiologie-France portal proposes an online catalogue of the main French sources of health databases and studies. The catalogue of the French Health Data Hub is set to be expanded soon and to offer several health databases (see below).

The potential strengths of academic cohort studies and registries in the context of real-world studies are:

- the support of health professionals, which limits the risk of low participation;
- the wealth of clinical and biological data enabling specific questions to be answered that could not be answered with reimbursement or hospitalisation data for example;
- the collection of information over several years to assess the efficacy and long-term adverse effects of health products;
- the possibility of studying all treatments for a given indication, as opposed to studies focusing on the follow-up of a specific medicinal product or medical device;
- the scientific quality of national cohorts.

The main limitations are:

- the absence of important data for the monitoring of medicinal products and medical devices because they were not designed to meet the objective of the real-world study;
- the lack of representativeness for cohorts constituted from selected centres;
- access to data requires the establishment of a partnership between the academic team and the pharmaceutical companies; this is a lengthy process, which must be anticipated and organised;
- the quality of data collection (e.g. risk of missing data and selection bias) and which varies from one cohort to another.

⁶ ANRS CO22 HEPATHER, Treatment options in hepatitis B and C: a French national cohort

⁷ French multiple sclerosis observatory

⁸ Chronic Kidney Disease - Réseau Epidémiologie et Information en Néphrologie (Epidemiology and Information in Nephrology network)

⁹ Rare Disease Cohorts

Annexe 5. The Health Data Hub and the SNDS database

The Health Data Hub is the French health data platform (<https://www.health-data-hub.fr/>) created by the law of 24 July 2019 relative to the organisation and transformation of the health system. It is a public structure allowing project leaders to access health data in a very secure environment, so that they can contribute to finding solutions that improve people's health. Very precise procedures govern this access and, above all, this data, which does not contain the surnames, first names or social security numbers of individuals. The aim of the Health Data Hub is therefore to ensure that health data, which represents a genuine collective asset, can be used for the benefit of society. This data comes, in particular, from the main database of the French national health data system (SNDS) but also from healthcare facilities, Santé publique France, etc. These other sources are combined within a “catalogue”, i.e. a collection of non-exhaustive databases that will be built up progressively and iteratively, in partnership with those responsible for collecting the data concerned. The replication of these databases of interest at the level of a single system will make it possible to pool human, technological and financial investments in order to:

- Constantly enrich each of the databases with medico-administrative data (i.e. the main SNDS database);
- Aggregate, link, extract and make available targeted data to authorised users via the Health Data Hub technology platform;
- Increase the visibility of the database, with a potential impact on the data re-use rate;
- Offer secure and GDPR-compliant hosting.

This catalogue offers an opportunity to conduct research of public interest with strong impacts for society. The list of flows in the main database and the databases making up the catalogue is set out in a decree (to be published at the time of writing this guide).

It is recommended to consult Health Data Hub teams in order to obtain technical and regulatory support concerning access to the data catalogue and/or the design of a database with a view to its hosting and possible linking to the SNDS.

Focus on the French National health data system

The French National health data system (SNDS) was created by the law on the modernisation of the health system in January 2016. It is a medico-administrative database warehouse making it possible to link several health data sources from:

- the French health insurance inter-scheme IT system (SNIIRAM), the French national health insurance system database created in 1999 by the social security financing act;
- the French National hospital discharge database (PMSI) containing hospital data;
- INSERM's CépiDC database recording medical causes of death;
- and medico-social data from the regional disability centre (MDPH) information system.

The SNDS is thus divided into various individual databases and also includes a 1:97 representative sample (Échantillon Généraliste des Bénéficiaires (EGB - generalist beneficiary sample)) of the protected population, as well as thematic databases with aggregated data (Datamarts BO).

The purposes and scope of the SNDS have evolved over the years. Initially made available to state agencies, access to these databases has gradually been extended to the scientific community, subject to an access authorisation issued per project. Temporary access for research purposes is regulated. In general, it is necessary to obtain an authorisation from the French data protection authority (CNIL) to process personal health data. Prior to its decision, two institutions are involved in the procedure:

- The Health Data Hub, a one-stop shop for projects using health data.

- The Ethical and Scientific Committee for Research, Studies and Evaluations in the Field of Health (CESREES), an independent committee responsible for issuing an opinion on the methodology used, the need to use personal data, the relevance of such data with respect to the purpose of data processing and, where appropriate, the scientific quality of the project.

There are also so-called simplified access procedures, in which CNIL authorisation is not required, particularly if the planned processing complies with a reference methodology (known as an MR).

The main SNDS database gathers data on patients affiliated to different health insurance schemes and can be considered to be quasi-exhaustive with respect to data on the French population. After a pseudonymisation process (83), it is thus possible to access individual data on beneficiaries (year of birth, sex, department of residence, date and municipality of death, the notion of complementary universal health cover, primary care physician), their registration on the list of chronic conditions (ALD), data concerning hospital stays derived from PMSI data (medicine, surgery, obstetrics, psychiatry, follow-on care and rehabilitation, hospitalisation at home, with so-called principal, related and associated medical diagnoses, hospitalisation dates, therapeutic or imaging procedures carried out) and all care reimbursed in the community setting (medical consultations, nursing procedures, biological or imaging procedures, dispensing of reimbursed medicinal products or medical devices: inter-scheme consumption data, DCIR).

Benefits and limitations

The principal asset of the main SNDS database is the quasi-exhaustive nature of the data with respect to the French population. The granularity of the data is high. For example, the data available on medicines reimbursed by the national health insurance system includes: the name of the drug substance, the name of the proprietary medicinal product, the strength and the number of therapeutic units, the delivery form. However, it should be noted that by identifying dates of dispensing of reimbursed medicinal products or medical devices, the actual exposure to the health product can only be approximated. In reality, treatment compliance is not directly observable and the only indicators available are:

- the regularity of dispensing dates;
- the quantities of health products dispensed;
- the time between the first and last date of dispensing observed.

Similarly, the notion of clinical monitoring will only be able to exploit the dates on which biological or radiological procedures are performed, given that their results are not available.

Nonetheless, insofar as the SNDS is an inter-scheme database, the studies carried out on these databases can reasonably claim to be virtually exhaustive and representative of the French population. However, it should be borne in mind that the databases are not updated in real time, with the periodicity varying depending on the data sources. In addition, the incompressible time required for the regulatory stage and then for making the data available must be taken into account when conducting a study.

A second advantage of the SNDS is the absence of attrition bias and a low missing data rate. The issue of subjects lost to follow-up is crucial in observational studies, particularly in the event of long follow-up periods. With few missing data and collection of the vital status of individuals from the civil status registry of the patient's municipality of birth or by cross-checking with the CépiDc registry, the SNDS database is therefore particularly suitable for studies requiring long follow-up. Finally, reimbursement data provide precise information for an economic evaluation. However, healthcare receipts can sometimes be submitted for reimbursement several months after the procedure has been performed.

Some data not collected in the main SNDS database nonetheless represent major limitations to be taken into account when conducting studies. In particular, the absence of physical examination results (imaging, laboratory data, etc.) or paraclinical data (smoking, blood pressure, BMI, etc.), the absence

of reasons for consultation and the lack of social data (apart from affiliation to the supplementary universal healthcare coverage (CMUC) system) limit the identification of certain risk factors. In addition, the SNDS does not record data on the consumption of medicinal products dispensed during hospital stays (other than those on the “additional list”). Information on medical diagnoses is still limited to the coding of hospital stays carried out within the framework of activity-based pricing or registration on the chronic conditions list, with a risk of misclassification if a patient is eligible for several chronic conditions.

The following paragraphs concerning the use of health or disease event identification algorithms and multi-source studies provide some answers to the limitations mentioned above.

Finally, it should be noted that the use of the SNDS for epidemiological purposes requires in-depth knowledge and expertise of these databases.

The use of algorithms for the identification of diseases

The use of identification algorithms is necessary, for example, to define an endpoint or a variable considered to be a source of potential confusion or the presumed indication of the medicinal product or medical device of interest.

The Division for Strategy, Studies and Statistics - Department of Disease Studies of the French National Health Insurance Fund (CNAM) regularly publishes updates of identification algorithms based on data from the main SNDS database for diseases, health events or treatments that it uses for the mapping of diseases in its annual report or for the “disease” sheets” [published on the ameli.fr website](#). Their construction methodology can be discussed, but the main difficulty remains the validation of these algorithms, which requires the linking of data from the main SNDS database with another data source (cohort, registry, etc.), in particular to calculate data sensitivity and specificity.

The role of the ReDSiam network (SNIIRAM data network) is to develop, submit for critical debate and make available algorithms (Fosse-Edorh et al. (84), Goldberg et al. (85)).

Because of the importance of the issue, the Health Data Hub has launched a call for expressions of interest for the development of algorithms and their validation.

These algorithms have also been the basis for the development of an index measuring the severity of health status, with the value of an evaluation of their capacity to predict mortality in individuals over 65 years old affiliated to France’s general health insurance scheme (Constantinou et al. 2018 (86)).

Clinical studies or registries are an opportunity to test the algorithms made available, provided that they are linked to the main SNDS database (see below). Thus, for example, the evaluation of the performance of an algorithm for identifying bleeding in the context of oral anticoagulant therapy (Maura et al. (87)) demonstrated a low sensitivity and lack of homogeneity depending on the nature of the bleeding (Oger et al. (88)).

The algorithms used in post-registration studies should be detailed in the protocols submitted to the HAS.

The value of multi-source studies

Linking the main SNDS database with other data sources offers several advantages (Scailteux et al. (89)). Firstly, it allows the data in the main SNDS database to be supplemented with clinical or para-clinical data, in particular to take into account risk factors in statistical analyses. The CONSTANCES cohort (Goldberg et al. (90)), the CANARI study (91), the SACHA study (92) and the ISO-PSY study (93) are interesting examples of multi-source studies. In addition, linking the main SNDS database to other data sources can also enable validation of algorithms (Fuentes et al. (94), Oger et al. (88)), the more cost-effective conduct of long-term follow-up, for example following a randomised clinical trial, or

the collection, again at lower cost, of information that is useful for an economic analysis of data from a randomised clinical trial.

However, data linkage remains a complex undertaking (Pratt et al. (95)) despite the recent simplification of the process to request authorisation to use the registry number (NIR) by the CNIL (96). In the absence of an NIR, various indirect (probabilistic or semi-deterministic) linkage methods are proposed (Bounebache et al. (97)). Other avenues are being explored, bearing in mind the specific constraints of the exercise in which the two databases to be linked may be large and asymmetrical in size.

Furthermore, making it easier to link databases with the main SNDS database is one of the missions and objectives of the Health Data Hub, with the support of the French national health insurance fund and the national old age insurance fund.

Annexe 6. Patient-Reported Outcome Measures

Description and value

The outcome as perceived by the patient is known as the *Patient-Reported Outcome* (PRO). The instruments used to measure these PROs (Patient-Reported Outcome Measures - PROMs) are mainly self-administered questionnaires, which generally enable the quantitative evaluation of numerous concepts. These instruments measure a one-dimensional or multi-dimensional scale. Each scale consists of one or more items (questions) with a pre-specified response format. For a given scale, measurement of the dimension (e.g. level of fatigue) - a single quantitative value - is obtained by an algebraic transformation of the item responses: the measurement model.

PROMs differ from other types of assessable health outcomes in that the measurement is performed by the patient, without interpretation of responses by a health professional (98). Hence PROMs are an instrument of choice for measuring concepts from the patient's perspective (99), including subjective concepts (related to the thinking subject, such as pain, fatigue, anxiety, and more broadly health-related quality of life). If properly validated, a PROM has good metrological (psychometric) properties. In the context of real-world studies, they are a useful resource to enrich the assessment of the impact of health technologies from the patients' perspective. However, the use of these instruments is contingent on specific methodological issues.

Methodological considerations

Choice of concept to be measured

PROMs are frequently referred to as “quality of life scales”, which can sometimes be a misnomer. Quality of life (or health-related quality of life) is only one of the concepts measurable by a PRO. Some are assessments of health status: for example, symptom intensity, impairment, or functioning. The widely used EQ-5D questionnaire is an example. Other PROMs relate to patients' satisfaction with the care received. Still others address perceived health as the subjective perception of health status or social functioning. The latter instruments can then measure what is called health-related quality of life. Many so-called specific PROMs are developed for a given disease, while others are generic. One advantage of generic instruments is that they provide a measure to compare different populations or disease situations. Specific instruments often have better content validity (they cover the given condition more comprehensively) (100).

Finally, a PROM must be rigorously selected with regard to the concept that is to be measured. A taxonomy of assessable health concepts can be found in the WHO ICF classification (101), Wilson and Cleary's biopsychosocial model (102), or the “*Montreal Accord on Patient-Reported Outcomes*” (103). PROQOLID™ (<https://eprovide.mapi-trust.org/>), maintained by the *Mapi Research trust*, is a PROMs research database. The HAS panorama on PREMs and PROMs (104) as well as good practice guidelines (105) can also guide the choice of the self-questionnaire. Practical considerations should also guide the selection of the instrument, such as the length of the test or the cognitive load associated with the test (language level, response format, etc.) or compliance with the HAS methodology guides.

Quality of the tool used

PROMs usually aim to provide a quantitative measure that must have good psychometric qualities. There are two main psychometric properties:

5. validity: does the instrument measure what it is supposed to measure? (e.g. check that an instrument measuring depressive symptoms does not detect anxiety symptoms);

6. reliability: what is the value of the measurement? (e.g. check that an instrument gives the same value under identical conditions).

The verification of the psychometric properties of a PROM is specifically studied at the time of its creation and validation, then supplemented by subsequent work during routine use (106). In addition to the creation of the instrument and an exploratory test on a few subjects (accompanied by qualitative studies), this specific work involves a validation on a few hundred subjects and a statistical analysis of the psychometric properties. Reliability and validity are not assessed by a single indicator, but by multiple criteria, accumulating evidence as the instrument is used. The validity of an instrument has many sub-aspects: face validity, content validity, construct validity, structural validity, criterion validity etc. The international group COSMIN¹⁰ proposes a consensus-based taxonomy of psychometric properties (107). Only PROs that have demonstrated sufficient psychometric properties should be used. A systematic review of the literature, as well as a critical analysis of the evidence should be conducted before using a PROM.

Two specific aspects should be noted. It is necessary to make sure that the desired PROM has been validated in the language of use. Translation of a PROM into a language other than its original language (cross-cultural adaptation) follows specific rules, including validation in the language of translation, as it is a process with a high risk of degrading psychometric properties through cross-cultural variability (108) (<https://euroqol.org/publications/user-guides/>). Care should be taken to ensure a match between the target population of the real-world study and the PROM validation population. Among other things, validation in a hospital population does not guarantee that use in the community setting will be appropriate. It is preferable to use the data from the real-world study to double-check the psychometric properties of the tool and confirm appropriate use.

Practical use

While the paper format remains predominant, electronic-format PROMs (ePROs) are increasingly being used (109). In addition to the logistical advantages (no need for manual data entry), ePROs offer the advantage of easier administration in various contexts (e.g. at home via an electronic interface). The response device can be adapted to disabilities (e.g. visual). Some electronic instruments, known as CAT (Computerized Adaptive Testing) instruments, enable adaptation of the proposed items (in number and content) from a pre-designed bank in order to measure the concept of interest with the least number of items possible (110). Nevertheless, these CAT instruments require the use of a complex and appropriate measurement model, requiring the modelling of item responses.

In order to obtain a valid and reliable measure of the concept of interest, it is essential to use an instrument in strict accordance with its proposed measurement model. There are several types of measurement models. The most widely used measurement model is “Classical Test Theory” (CTT). The measure is often the simple sum of items on a scale (sometimes weighted). This easily obtained measure is called the score and, subject to good psychometric properties, is a good ordinal measure (it allows subjects to be ordered on the scale) of the concept of interest (111). Other measurement models are known as latent variable models (a variable that cannot be observed but that is assumed to explain the responses of several items, e.g. the quality of life concept): “Rasch Measurement Theory” (RMT), “Item Response Theory” (IRT) or “Structural Equation Modelling” (SEM) (112, 113). These models, at the same cost as numerical modelling, make it possible to obtain a measurement with more interesting properties, such as the interval property (each deviation of one unit on the scale represents the same quantity) or to model complex structures. Furthermore, for a given scale, no items should be removed from the scale (except for CAT instruments). For a multi-domain PROM, it is only permissible to estimate an overall measure in

¹⁰ COnsensus-based Standards for the selection of health Measurement INstruments

addition to the measurement profile of each of the scales if the structure of the instrument allows for this and there is a measurement model for this value.

Results and interpretation

Interpreting the results of PROMs (especially assessing the effect size), whether it is an individual estimate, an inter-group difference, or a change, is sometimes not easy. Some PROMs have been administered in large samples and there are therefore distributions that can be described as population norms. This is the case for SF-36, a quality of life measure, for which the distribution in many populations is available. In this case, a comparison of a subject or groups against the population norm can help with interpretation (114). This is also the case for the EQ-5D instrument, which is a multi-attribute instrument to be considered for economic evaluation (115, 116).

In addition, the a priori use of a Responder Definition (RD) is common. This is a threshold of change over a predetermined period of time that can be interpreted as a benefit of treatment. This classifies each patient in a sample as having experienced a benefit or not (98). This RD can be obtained from multiple perspectives: clinical according to the patient (a value that associates change on the PRO scale with an outcome such as disease severity or prognosis) or professional (the minimum change that will be considered a significant benefit). A frequently used perspective is to assess the Minimal Importance Difference (MID) of a PRO: i.e., the smallest change that the subject perceives as significant (117). Some instruments are therefore accompanied by MID estimates, which can be an aid to interpretation.

As the response to items is the result of a variable interpretative process, there is increasing evidence that PROMs do not stand up to the hypothesis of measurement invariance (the stability of the measurement scale, either over time or across different populations). Over time, due to psychological adaptation mechanisms to health events, patients' perception of the concepts being measured may change: the consequence is the response shift phenomenon (e.g. being "moderately tired" does not correspond to the same level of fatigue before and after chemotherapy (118)). Systematic differences in the level of responses to certain items of a PRO may exist between different populations (e.g. by age categories, or by gender) at the same concept level. For the same level of fatigue, older subjects may respond differently to a question than younger subjects, due to different interpretation of the question. This is known as "Differential item Functioning" (DIF) (119)). These invariance violations may invalidate the simple comparison of scores as correctly representing variations in the level of the concept of interest. These violations can be taken into account and enrich the interpretation of measurement variations via the use of appropriate methods and questionnaires.

Literature search

A literature search was conducted in the Medline, Embase and Science Direct databases from January 2013 to April 2020, in English and French, using the following search equation: *"Product Surveillance, Postmarketing"[Mesh] AND ("Drug Approval"[Majr] OR "Device Approval"[Majr]) OR (post-approval study or post-approval studies OR real world study design or real world studies design OR real life study design or real life studies design OR post-authorization studies or post-authorization study OR observational stud* design) Field: Title*

281 references were obtained

Bibliographic references

1. Haute Autorité de Santé. Doctrine de la commission de la transparence. Principes d'évaluation de la CT relatifs aux médicaments en vue de leur accès au remboursement. Saint-Denis La Plaine: HAS; 2018. https://www.has-sante.fr/upload/docs/application/pdf/2018-10/doctrine_10102018.pdf
2. Haute Autorité de Santé. Principes d'évaluation de la CNEDiMTS relatifs aux dispositifs médicaux à usage individuel en vue de leur accès au remboursement. Saint-Denis La Plaine: HAS; 2019. https://www.has-sante.fr/upload/docs/application/pdf/2017-11/principes_devaluation_de_la_cnedimts-v4-161117.pdf
3. Haute Autorité de Santé. Choix méthodologiques pour l'évaluation économique à la HAS. Saint-Denis La Plaine: HAS; 2019. https://www.has-sante.fr/upload/docs/application/pdf/2020-07/guide_methodologique_evaluation_economique_has_2020_vf.pdf
4. Haute Autorité de Santé. Choix méthodologiques pour l'analyse de l'impact budgétaire à la HAS. Saint-Denis La Plaine: HAS; 2016. https://www.has-sante.fr/upload/docs/application/pdf/2016-12/guide_methodologique_choix_methodologiques_pour_l_analyse_de_limpact_budgetaire_a_la_has_.pdf
5. Haute Autorité de Santé. Cartographie des impacts organisationnels pour l'évaluation des technologies de santé. Saint-Denis La Plaine: HAS; 2020. https://www.has-sante.fr/upload/docs/application/pdf/2020-12/guide_methodologique_impacts_organisationnels.pdf
6. Parlement Européen, Conseil de l'Union Européenne. Règlement (UE) 2016/679 du parlement européen et du conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données) Journal Officiel de l'Union Européenne 2016; L 119-1(04/05).
7. Parlement Européen, Conseil de l'Union Européenne. Rectificatif au règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données). Journal Officiel de l'Union Européenne 2018 ;L 119-1(23/05).
8. Agence du numérique en santé. Hébergement des données de santé. Labels et certifications [En ligne] 2021. <https://esante.gouv.fr/labels-certifications/hebergement-des-donnees-de-sante>
9. Conseil national de l'Industrie, Levy-Bachelot L, Mollo L, Josserean A, Bacry E, Combes S, *et al.* Programme «intelligence Artificielle et santé» du Comité Stratégique de Filière Industries et Technologies de Santé. Paris: Alliance Recherche et Innovation Industries Santé (ARRIS); 2021. https://gt2.ariis.fr/wp-content/uploads/2021/03/Recommandations-du-GT-Usage-des-donnees-en-vie-reelle-dans-la-prise-de-decision_vf.pdf
10. Haute Autorité de Santé. Les études post-inscription sur les technologies de santé (médicaments, dispositifs médicaux et actes) Principes et méthodes. Saint-Denis La Plaine: HAS; 2011. https://www.has-sante.fr/upload/docs/application/pdf/2012-01/etudes_post_inscription_technologies_sante.pdf
11. Haute Autorité de Santé. Projet stratégique 2019 – 2024. Saint-Denis La Plaine: HAS; 2018. https://www.has-sante.fr/upload/docs/application/pdf/2018-11/projet_strategique_2019-2024.pdf
12. Makady A, Ham RT, de Boer A, Hillege H, Klungel O, Goettsch W. Policies for use of real-world data in health technology assessment (hta): a comparative study of six HTA Agencies. *Value Health* 2017;20(4):520-32. <http://dx.doi.org/10.1016/j.jval.2016.12.003>
13. Makady A, van Veelen A, Jonsson P, Moseley O, D'Andon A, de Boer A, *et al.* Using real-world Data in Health Technology Assessment (HTA) practice: A comparative study of five HTA agencies. *Pharmacoeconomics* 2018;36(3):359-68. <http://dx.doi.org/10.1007/s40273-017-0596-z>
14. Hordijk-Trion M, Lenzen M, Wijns W, de Jaegere P, Simoons ML, Scholte OP, Reimer WJ, *et al.* Patients enrolled in coronary intervention trials are not representative of patients in clinical practice: results from the Euro Heart Survey on Coronary Revascularization. *Eur Heart J* 2006;27(6):671-8. <http://dx.doi.org/10.1093/eurheartj/ehi731>
15. Lewis JH, Kilgore ML, Goldman DP, Trimble EL, Kaplan R, Montello MJ, *et al.* Participation of patients 65 years of age or older in cancer clinical trials. *J Clin Oncol* 2003;21(7):1383-9.
16. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA* 2007;297(11):1233-40. <http://dx.doi.org/10.1001/jama.297.11.1233>
17. Haute Autorité de Santé. Sexe, genre et santé. Rapport d'analyse prospective 2020. Saint-Denis La Plaine: HAS; 2020. https://www.has-sante.fr/upload/docs/application/pdf/2020-12/rapport_analyse_prospective_2020.pdf
18. Gor KA, Shah KN, Joshi PB, Joshi HM, Rana DA, Malhotra SD. Off-label drugs use in neurology outpatient department: A prospective study at a tertiary care teaching hospital. *Perspect Clin Res* 2020;11(1):31-6. http://dx.doi.org/10.4103/picr.PICR_117_18 10.4103/picr.PICR_117_18. Epub 2019 Apr 26.
19. Rebordosa C, Houben E, Laugesen K, Bothner U, Montonen J, Aguado J, *et al.* No evidence of off-label use of olodaterol and indacaterol in Denmark, France, and the Netherlands: A Drug Utilization Study. *Sci Rep* 2020;10(1):586. <http://dx.doi.org/10.1038/s41598-019-57397-5> 10.1038/s41598-019-57397-5.
20. Stafford RS. Regulating off-label drug use--rethinking the role of the FDA. *N Engl J Med* 2008;358(14):1427-9. <http://dx.doi.org/10.1056/NEJMp0802107> 10.1056/NEJMp0802107.
21. Briasoulis A, Gao Y, Inampudi C, Alvarez P, Asleh R, Chrischilles E, *et al.* Characteristics and outcomes in

- patients with atrial fibrillation receiving direct oral anticoagulants in off-label doses. *BMC Cardiovasc Disord* 2020;20(1):42.
<http://dx.doi.org/10.1186/s12872-020-01340-4>
10.1186/s12872-020-01340-4.
22. Nordon C, Karcher H, Groenwold RH, Ankarfeldt MZ, Pichler F, Chevrou-Severac H, *et al*. The "efficacy-effectiveness gap": Historical background and current conceptualization. *Value Health* 2016;19(1):75-81.
<http://dx.doi.org/10.1016/j.jval.2015.09.2938>
23. Chassang S, Snowberg E, Seymour B, Bowles C. Accounting for behavior in treatment effects: new applications for blind trials. *PLoS ONE* 2015;10(6):e0127227.
<http://dx.doi.org/10.1371/journal.pone.0127227>
10.1371/journal.pone.0127227. eCollection 2015.
24. Nordon C, Bovagnet T, Belger M, Jimenez J, Olivares R, Chevrou-Severac H, *et al*. Trial exclusion criteria and their impact on the estimation of antipsychotic drugs effect: A case study using the SOHO database. *Schizophr Res* 2018;193:146-53.
<http://dx.doi.org/10.1016/j.schres.2017.07.031>
10.1016/j.schres.2017.07.031. Epub 2017 Jul 13.
25. Rothwell PM. Factors that can affect the external validity of randomised controlled trials. *PLoS Clin Trials* 2006;1(1):e9.
<http://dx.doi.org/10.1371/journal.pctr.0010009>
26. Kennedy-Martin T, Curtis S, Faries D, Robinson S, Johnston J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials* 2015;16:495.
<http://dx.doi.org/10.1186/s13063-015-1023-4>
10.1186/s13063-015-1023-4.
27. Haute Autorité de Santé, Commission de la Transparence. Doctrine de la Commission de la Transparence. Principes d'évaluation de la CT relatifs aux médicaments en vue de leur accès au remboursement. Saint-Denis La Plaine: HAS; 2020.
https://www.has-sante.fr/upload/docs/application/pdf/2021-03/doctrine_ct.pdf
28. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. *Stat Med* 2012;31(25):2973-84.
<http://dx.doi.org/10.1002/sim.5403>
29. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Code of conduct. London: ENCEPP; 2018.
http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml
30. European Network for Health Technology Assessment, Haute Autorité de Santé. Criteria to select and prioritize health technologies for additional evidence generation. London: EUnetHTA; 2012.
31. European Network for Health Technology Assessment, Haute Autorité de Santé, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Position paper on how to best formulate research recommendations for primary research arising from HTA reports. London: EUnetHTA; 2015.
32. European Network for Health Technology Assessment, Haute Autorité de Santé, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Position paper on how to decide on the appropriate study design for primary research arising from HTA reports. London: EUnetHTA; 2015.
33. European Network for Health Technology Assessment. Core protocol pilot for additional evidence generation. London: EUnetHTA; 2019.
https://eunetha.eu/wp-content/uploads/2019/10/EUnetHTA_Core-protocol-Pilot-for-AEG_0.pdf
34. European Network for Health Technology Assessment. REQueST® Tool and its vision paper [En ligne]. London: EUnetHTA; 2020.
<https://eunetha.eu/wp-content/uploads/2019/10/EUnetHTA-Position-Paper-on-research-recommendations.pdf>
35. European Network for Health Technology Assessment. EUnetHTA PLEG activities [En ligne] 2018.
<https://eunetha.eu/pleg/>
36. European Medicines Agency, Arlett P. DARWIN EU (Data Analytics and Real World Interrogation Network). PCWP and HCPWP Data workshop 23 September 2020. Amsterdam: EMA; 2020.
https://www.ema.europa.eu/en/documents/presentation/presentation-proposal-darwin-eu-data-analytics-real-world-interrogation-network-parlett-ema_en.pdf
37. European Medicines Agency. Guideline on registry-based studies - launch of public consultation [En ligne] 2020.
<https://www.ema.europa.eu/en/news/guideline-registry-based-studies-launch-public-consultation>
38. International Society for Pharmacoepidemiology. Guidelines for Good Pharmacoepidemiology Practices (GPP). Bethesda: ISPE; 2015.
39. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, *et al*. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61(4):344-9.
<http://dx.doi.org/10.1016/j.jclinepi.2007.11.008>
10.1016/j.jclinepi.2007.11.008.
40. Gallo V, Egger M, McCormack V, Farmer PB, Ioannidis JP, Kirsch-Volders M, *et al*. Strengthening the Reporting of Observational studies in Epidemiology--Molecular Epidemiology (STROBE-ME): an extension of the STROBE Statement. *PLoS Med* 2011;8(10):e1001117.
<http://dx.doi.org/10.1371/journal.pmed.1001117>
41. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. Guide on methodological standards in Pharmacoepidemiology (revision 8) : ENCEPP; 2010.
42. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). ENCePP Checklist for study protocols (Revision 4). London: ENCEPP; 2018.
http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml
43. Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, *et al*. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE special task force on real-world evidence in health care decision making. *Value Health* 2017;20(8):1003-8.
<http://dx.doi.org/10.1016/j.jval.2017.08.3019>
10.1016/j.jval.2017.08.3019. Epub 2017 Sep 15.

44. Agency for Healthcare Research and Quality, Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM. Developing a protocol for observational comparative effectiveness research. A user's guide. Rockville: AHRQ; 2013.
<https://effectivehealthcare.ahrq.gov/sites/default/files/related-files/user-guide-observational-cer-130113.pdf>
45. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332.
<http://dx.doi.org/10.1136/bmj.c332>
10.1136/bmj.c332.
46. Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, *et al.* The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med* 2015;12(10):e1001885.
<http://dx.doi.org/10.1371/journal.pmed.1001885>
47. Wang SV, Schneeweiss S, Berger ML, Brown J, de Vries F, Douglas I, *et al.* Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1.0. *Pharmacoepidemiol Drug Saf* 2017;26(9):1018-32.
<http://dx.doi.org/10.1002/pds.4295>
48. Berdai D, Thomas-Delecout F, Szwarcensztein K, d'Andon A, Collignon C, Comet D, *et al.* Requests for post-registration studies (PRS), patients follow-up in actual practice: Changes in the role of databases. *Therapie* 2018;73(1):13-24.
<http://dx.doi.org/10.1016/j.therap.2017.12.008>
49. Food and Drug Administration, Center for Devices and Radiological Health, Center for Biologics Evaluation and Research. Use of real-world evidence to support regulatory decision-making for medical devices. Guidance for Industry and Food and Drug Administration Staff. Rockville: FDA; 2017.
<https://www.fda.gov/media/99447/download>
50. European Network for Health Technology Assessment. Therapeutic medical devices. Guideline. London: EUnetHTA; 2015.
https://eunetha.eu/wp-content/uploads/2018/01/Therapeutic-medical-devices_Guideline_Final-Nov-2015.pdf
51. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, *et al.* AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.
<http://dx.doi.org/10.1136/bmj.j4008>
52. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, *et al.* Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283(15):2008-12.
<http://dx.doi.org/10.1001/jama.283.15.2008>
53. Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, *et al.* STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ* 2021;372:m4856.
<http://dx.doi.org/10.1136/bmj.m4856>
54. Carroll R, Ramagopalan SV, Cid-Ruzafa J, Lambrelli D, McDonald L. An analysis of characteristics of post-authorisation studies registered on the ENCePP EU PAS Register. *F1000Res* 2017;6:1447.
<http://dx.doi.org/10.12688/f1000research.12198.2>
10.12688/f1000research.12198.2. eCollection 2017.
55. Hill KP, Ross JS, Egilman DS, Krumholz HM. The ADVANTAGE seeding trial: a review of internal documents. *Ann Intern Med* 2008;149(4):251-8.
<http://dx.doi.org/10.7326/0003-4819-149-4-200808190-00006>
10.7326/0003-4819-149-4-200808190-00006.
56. Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, *et al.* Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Pharmacoepidemiol Drug Saf* 2017;26(9):1033-9.
<http://dx.doi.org/10.1002/pds.4297>
57. Strom BL, Kimmel SE, Hennessy S. *Pharmacoepidemiology*. 6th ed. : Wiley-Blackwell; 2019.
58. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia: Lippincott Williams and Wilkins; 2012.
59. Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep* 2015;2(4):221-8.
<http://dx.doi.org/10.1007/s40471-015-0053-5>
10.1007/s40471-015-0053-5. Epub 2015 Sep 30.
60. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ* 2010;340:b5087.
<http://dx.doi.org/10.1136/bmj.b5087>
10.1136/bmj.b5087.
61. Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med* 2019;25(10):1601-6.
<http://dx.doi.org/10.1038/s41591-019-0597-x>
62. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
<http://dx.doi.org/10.1136/bmj.b2393>
63. Concato J, Lawler EV, Lew RA, Gaziano JM, Aslan M, Huang GD. Observational methods in comparative effectiveness research. *Am J Med* 2010;123(12 Suppl 1):e16-23.
<http://dx.doi.org/10.1016/j.amjmed.2010.10.004>
64. Candlish J, Pate A, Sperrin M, van Staa T. Evaluation of biases present in the cohort multiple randomised controlled trial design: a simulation study. *BMC Med Res Methodol* 2017;17(1):17.
<http://dx.doi.org/10.1186/s12874-017-0295-7>
65. Pate A, Candlish J, Sperrin M, Van Staa TP. Cohort Multiple Randomised controlled trials (cmRCT) design: efficient but biased? A simulation study to evaluate the feasibility of the Cluster cmRCT design. *BMC Med Res Methodol* 2016;16(1):109.
<http://dx.doi.org/10.1186/s12874-016-0208-1>
66. Eichner FA, Groenwold RHH, Grobbee DE, Oude Rengerink K. Systematic review showed that stepped-wedge cluster randomized trials often did not reach their planned sample size. *J Clin Epidemiol* 2019;107:89-100.
<http://dx.doi.org/10.1016/j.jclinepi.2018.11.013>

67. Reams BD, O'Malley CD, Critchlow CW, Lauffenburger JC, Brookhart MA. Changing patterns of use of osteoporosis medications in the years after launch: implications for comparative effectiveness research. *Pharmacoepidemiol Drug Saf* 2014;23(3):251-60. <http://dx.doi.org/10.1002/pds.3545>
10.1002/pds.3545. Epub 2013 Nov 25.
68. Ankarfeldt MZ, Thorsted BL, Groenwold RH, Adalsteinsson E, Ali MS, Klungel OH. Assessment of channeling bias among initiators of glucose-lowering drugs: A UK cohort study. *Clin Epidemiol* 2017;9:19-30. <http://dx.doi.org/10.2147/CLEP.S124054>
69. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology* 2009;20(4):512-22. <http://dx.doi.org/10.1097/EDE.0b013e3181a663cc>
10.1097/EDE.0b013e3181a663cc.
70. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11(5):550-60. <http://dx.doi.org/10.1097/00001648-200009000-00011>
10.1097/00001648-200009000-00011.
71. Ford I, Norrie J. Pragmatic Trials. *N Engl J Med* 2016;375(5):454-63. <http://dx.doi.org/10.1056/NEJMra1510059>
10.1056/NEJMra1510059.
72. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967;20(8):637-48. [http://dx.doi.org/10.1016/0021-9681\(67\)90041-0](http://dx.doi.org/10.1016/0021-9681(67)90041-0)
73. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furburg CD, Altman DG, *et al*. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol* 2009;62(5):464-75. <http://dx.doi.org/10.1016/j.jclinepi.2008.12.011>
74. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147. <http://dx.doi.org/10.1136/bmj.h2147>
75. Staa TP, Goldacre B, Gulliford M, Cassell J, Pirmohamed M, Taweel A, *et al*. Pragmatic randomised trials using routine electronic health records: putting them to the test. *BMJ* 2012;344:e55. <http://dx.doi.org/10.1136/bmj.e55>
10.1136/bmj.e55.
76. Giraudeau B, Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med* 2009;6(5):e1000065. <http://dx.doi.org/10.1371/journal.pmed.1000065>
10.1371/journal.pmed.1000065. Epub 2009 May 5.
77. Califf RM. Pragmatic clinical trials: Emerging challenges and new roles for statisticians. *Clinical trials (London, England)* 2016;13(5):471-7. <http://dx.doi.org/10.1177/1740774516656944>
78. Cook AJ, DeLong E, Murray DM, Vollmer WM, Heagerty PJ. Statistical lessons learned for designing cluster randomized pragmatic clinical trials from the NIH Health Care Systems Collaboratory Biostatistics and Design Core. *Clinical trials (London, England)* 2016;13(5):504-12. <http://dx.doi.org/10.1177/1740774516646578>
79. Ezzati A, Lipton RB. Machine Learning Predictive models can improve efficacy of clinical trials for alzheimer's disease. *J Alzheimer Dis*. 2020;74(1):55-63. <http://dx.doi.org/10.3233/jad-190822>
80. Adam G, Rampásek L, Safikhani Z, Smirnov P, Haibe-Kains B, Goldenberg A. Machine learning approaches to drug response prediction: challenges and recent progress. *NPJ Precision Oncology* 2020;4:19. <http://dx.doi.org/10.1038/s41698-020-0122-1>
81. Direction générale de l'offre de soins. H07 : Les registres à caractère épidémiologique. Paris: Ministère des solidarités et de la santé; 2019. https://solidarites-sante.gouv.fr/IMG/pdf/dgso_h07_fiche_mig_registres_2019.pdf
82. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005;58(4):323-37. <http://dx.doi.org/10.1016/j.jclinepi.2004.10.012>
10.1016/j.jclinepi.2004.10.012.
83. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, *et al*. Value of a national administrative database to guide public decisions: From the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. *Rev Epidemiol Sante Publique* 2017;65 Suppl 4:S149-S67. <http://dx.doi.org/10.1016/j.respe.2017.05.004>
84. Fosse-Edorh S, Rigou A, Morin S, Fezeu L, Mandereau-Bruno L, Fagot-Campagna A. Algorithmes basés sur les données médico-administratives dans le champ des maladies endocriniennes, nutritionnelles et métaboliques, et en particulier du diabète. *Rev Epidemiol Sante Publique* 2017;65 Suppl 4:S168-S73. <http://dx.doi.org/10.1016/j.respe.2017.05.001>
85. Goldberg M, Carton M, Doussin A, Fagot-Campagna A, Heyndrickx E, Lemaître M, *et al*. Le réseau REDSIAM (Réseau données Sniiram) – Spécial REDSIAM. *Rev Epidemiol Sante Publique* 2017;65 Suppl 4:S144-S8. <http://dx.doi.org/10.1016/j.respe.2017.06.001>
86. Constantinou P, Tuppin P, Fagot-Campagna A, Gastaldi-Ménager C, Schellevis FG, Pelletier-Fleury N. Two morbidity indices developed in a nationwide population permitted performing outcome-specific severity adjustment. *J Clin Epidemiol* 2018;103:60-70. <http://dx.doi.org/10.1016/j.jclinepi.2018.07.003>
87. Maura G, Blotière PO, Bouillon K, Billionnet C, Ricordeau P, Alla F, *et al*. Comparison of the short-term risk of bleeding and arterial thromboembolic events in nonvalvular atrial fibrillation patients newly treated with dabigatran or rivaroxaban versus vitamin K antagonists: a French nationwide propensity-matched cohort study. *Circulation* 2015;132(13):1252-60. <http://dx.doi.org/10.1161/circulationaha.115.015710>
88. Oger E, Botrel MA, Juchault C, Bouget J. Sensitivity and specificity of an algorithm based on medico-administrative data to identify hospitalized patients with major bleeding presenting to an emergency department. *BMC Med Res Methodol* 2019;19(1):194. <http://dx.doi.org/10.1186/s12874-019-0841-6>
89. Scailteux LM, Droitcourt C, Balusson F, Nowak E, Kerbrat S, Dupuy A, *et al*. French administrative health care database (SNDS): The value of its enrichment. *Thérapie* 2019;74(2):215-23.

<http://dx.doi.org/10.1016/j.therap.2018.09.072>

90. Goldberg M, Carton M, Descatha A, Leclerc A, Roquelaure Y, Santin G, *et al.* CONSTANCES: a general prospective population-based cohort for occupational and environmental epidemiology: cohort profile. *Occup Environ Med* 2017;74(1):66-71.

<http://dx.doi.org/10.1136/oemed-2016-103678>

91. Scailteux LM, Balusson F, Vincendeau S, Rioux-Leclercq N, Nowak E. Rationale and design of the CANARI study: a case-control study investigating the association between prostate cancer and 5-alpha-reductase inhibitors for symptomatic benign prostate hypertrophy by linking SNIIRAM and pathology laboratories in a specific region in France. *Fundam Clin Pharmacol* 2018;32(1):120-9.

<http://dx.doi.org/10.1111/fcp.12320>

92. Bouget J, Balusson F, Scailteux LM, Maignan M, Roy PM, L'Her E, *et al.* Major bleeding with antithrombotic agents: a 2012-2015 study using the French nationwide Health Insurance database linked to emergency department records within five areas - rationale and design of SACHA study. *Fundam Clin Pharmacol* 2019;33(4):443-62.

<http://dx.doi.org/10.1111/fcp.12444>

93. Agence nationale de sécurité du médicament et des produits de santé. Rapport scientifique de fin d'étude Etude «ISOPSY2». Saint-Denis : ANSM; 2020.

https://www.ansm.sante.fr/var/ansm_site/storage/original/application/bd8e92bf4eff92e1c0543708dcdddbbc.pdf

94. Fuentes S, Cosson E, Mandereau-Bruno L, Fagot-Campagna A, Bernillon P, Goldberg M, *et al.* Identifying diabetes cases in health administrative databases: a validation study based on a large French cohort. *Int J Public Health* 2019;64(3):441-50.

<http://dx.doi.org/10.1007/s00038-018-1186-3>

10.1007/s00038-018-1186-3. Epub 2018 Dec 4.

95. Pratt NL, Mack CD, Meyer AM, Davis KJ, Hammill BG, Hamp C, *et al.* Data linkage in pharmacoepidemiology: A call for rigorous evaluation and reporting. *Pharmacoepidemiol Drug Safety* 2020;29(1):9-17.

<http://dx.doi.org/10.1002/pds.4924>

96. Commission nationale informatique et libertés. Modalités de circulation du NIR pour la recherche en sante aux fins d'appariement de données avec le SNDS. Guide pratique. Paris: CNIL; 2020.

https://www.cnil.fr/sites/default/files/atoms/files/guide_pratique_circuits_nir_recherche_en_sante.pdf

97. Bounebach SK, Quantin C, Benzenine E, Obozinski G, Rey G. Revue bibliographique des méthodes de couplage des bases de données : Applications et perspectives dans le cas des données de santé publique *J Soc Fr Stat* 2018;159 :79-123.

98. Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH). Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. Silver Springs: U.S. Department of Health and Human Services; 2009.

<https://www.fda.gov/media/77832/download>

99. Ziegelstein RC. Personomics. *JAMA internal medicine* 2015;175(6):888-9.

<http://dx.doi.org/10.1001/jamainternmed.2015.0861>

100. Fayers PM, Machin D. Quality of life: The assessment, analysis and interpretation of patient-reported outcomes. 2nd ed. Chichester: Wiley; 2013.

101. World Health Organization. How to use the ICF. A practical manual for using the International Classification of Functioning, Disability and Health (ICF). Exposure draft for comment. Geneva: WHO; 2013.

<https://www.who.int/classifications/drafticfpracticalmanual.pdf>

102. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA* 1995;273(1):59-65.

103. Mayo NE, Figueiredo S, Ahmed S, Bartlett SJ. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 2: terminology proposed to measure what matters in health. *J Clin Epidemiol* 2017;89:119-24.

<http://dx.doi.org/10.1016/j.jclinepi.2017.04.013>

104. Haute Autorité de Santé. Indicateurs de qualité perçue par le patient de type PROMS et PREMS – Panorama d'expériences étrangères et principaux enseignements. Note de cadrage. Saint-Denis La Plaine: HAS; 2020.

https://www.has-sante.fr/upload/docs/application/pdf/2020-05/panorama-proms-prems_note_de_cadrage.pdf

105. Haute Autorité de Santé. Guide du parcours de soins. Bronchopneumopathie chronique obstructive. Saint-Denis La Plaine: HAS; 2019.

https://www.has-sante.fr/upload/docs/application/pdf/2020-01/app_323_guide_bpco_actu_2019_vf.pdf

106. Falissard B. Mesurer la subjectivité en santé. Perspective méthodologique et statistique. Paris: Elsevier Masson; 2008.

https://www.elsevier-masson.fr/mesurer-la-subjectivite-en-sante-9782294703171.html?gclid=EAlaIqobChMlob3l9OaT7wIVCJ7VCh3sCw99EAAYASAAEglijyFD_BwE&gclid=aw.ds

107. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, *et al.* The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63(7):737-45.

<http://dx.doi.org/10.1016/j.jclinepi.2010.02.006>

108. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25(24):3186-91.

<http://dx.doi.org/10.1097/00007632-200012150-00014>

109. Egeter J, Hüfner K, Sztankay M, Holzner B, Sperner-Unterwieser B. Implementation of an electronic routine outcome monitoring at an inpatient unit for psychosomatic medicine. *J Psychosom Res* 2018;105:64-71.

<http://dx.doi.org/10.1016/j.jpsychores.2017.12.009>

110. Rose M, Bjorner JB, Fischer F, Anatchkova M, Gandek B, Klapp BF, *et al.* Computerized adaptive testing--ready for ambulatory monitoring? *Psychosom Med* 2012;74(4):338-48.

<http://dx.doi.org/10.1097/PSY.0b013e3182547392>

111. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. A practical guide. Cambridge: Cambridge University Press; 2011.

https://assets.cambridge.org/9780521118200/frontmatter/9780521118200_frontmatter.pdf

112. de Ayala RJ. The theory and practice of item response theory. New York: The Guilford Press; 2008.

113. Brown TA. Confirmatory factor analysis for applied research. 2nd ed. New York: The Guilford Press; 2015.
114. Leplège A, Ecosse E, Pouchot J, Coste J, Perneger T. Le questionnaire MOS SF-36 : manuel de l'utilisateur et guide d'interprétation des scores. Paris: Editions ESTEM; 2001.
115. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990;16(3):199-208.
[http://dx.doi.org/10.1016/0168-8510\(90\)90421-9](http://dx.doi.org/10.1016/0168-8510(90)90421-9)
116. Szende A, Janssen B, Cabases J. Self-Reported Population Health: An international perspective based on EQ-5D. Dordrecht NL: Springer; 2014.
https://www.ncbi.nlm.nih.gov/books/NBK500356/pdf/Bookshelf_NBK500356.pdf
117. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61(2):102-9.
<http://dx.doi.org/10.1016/j.jclinepi.2007.03.012>
118. Sprangers MAG, Sajobi T, Vanier A, Mayo NE, Sawatzky R, Lix LM, *et al.* Response shift in results of patient-reported outcome measures: a commentary to the response shift-in sync working group initiative. Quality Life Res.2021.
<http://dx.doi.org/10.1007/s11136-020-02747-4>
119. Teresi JA, Fleishman JA. Differential item functioning and health assessment. Qual Life Res. 2007;16 Suppl 1:33-42.
<http://dx.doi.org/10.1007/s11136-007-9184-6>

Abbreviations and acronyms

MA	Marketing authorisation
CEESP	Commission d'évaluation économique et de santé publique (Commission for Economic and Public Health Evaluation)
CNEDiMETS	Commission nationale d'évaluation des dispositifs médicaux et des technologies de santé (Medical Device and Health Technology Evaluation Committee)
CT	Transparency Committee
MD	Medical Device
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EUnethTA	European Network for Health Technology Assessment
HAS	Haute Autorité de santé (French National Authority for Health)
ISPE	International Society of Pharmacoepidemiology
MAR	Missing At Random
NMAR	Not Missing At Random
PAES	Post-authorisation efficacy studies
PASS	Post-authorisation safety studies
PROM	Patient Reported Outcome Measure
REQUEST	REQueST Tool and its vision paper
SNDS	Système national des données de santé (French National health data system)

Retrouvez tous nos travaux sur
www.has-sante.fr

