Methodology for the clinical development of medical devices
# Description of the publication

## Title

Methodology for the clinical development of medical devices

## Work method

Methodology guide drawn up based on a literature review, an inventory of French and international guidelines and the opinions of steering group and review group experts.

The working method includes the following different steps:

- definition of the project by the steering group;
- literature search: monitoring of research conducted in 2013 and new research in complementary fields (methodologies specifically developed for small target populations, *in silico* methods, studies constructed around medico-administrative databases);
- selection and analysis of the literature;
- drafting of the guide: updating of the 2013 guide and drafting of new sections;
- review of the guide by the review group;
- finalising of the guide by the steering group, taking into account the review group’s comments;
- examination by the Medical Device and Health Technology Evaluation Committee (CNEDIMTS) with a view to validation of the final version of the updated methodology guide.

## Purpose(s)

The purpose of this methodology guide is to support medical device development. It aims to provide practical benchmarks relating to methodological aspects in order to optimise the level of evidence of different types of studies and increase confidence in their results.

## Targets concerned

The updated methodology guide is aimed at manufacturers, research structures, project leaders and healthcare professionals involved in the clinical development of a medical device and who are planning to submit an application for registration for reimbursement of the medical device in question to the CNEDIMTS. Patients are also concerned, since this guide is intended to promote the collection of evidence, whatever the endpoint used (clinical, quality of life, organisational).

## Requester

Self-referral

## Sponsor(s)

*Haute Autorité de santé* (French National Authority for Health) (HAS)

## Project management

*Service évaluation des dispositifs* (SED - HAS Medical Device Assessment Department)

## Literature search

Documentation Department

## Authors

Steering group: SED project team for HAS: Aurélie GAREL PACULL, Emmanuelle FOUTEAU, Corinne COLLIGNON; members of the documentation department: Sophie DESPEYROUX, Juliette CHAZARENG; members of the CNEDIMTS: Dominique COSTAGLIOLA, Pascal SELLIER, Sarah ZOHAR.

Review group: Alain BERNARD, Michel CUCHERAT, Pascal GIRAUX, French National Medical Technologies Industry Union (SNITEM).

With the participation of Éric DARVOY (diagrams and illustrations), Laura DRUART (graphic chart), Isabelle DUBRAIL and Amara HRUSTIC (secretarial services).

Reminder of participants in the first version in 2013: project group: Isabelle FOURNEL (project officer); SED: Michel VANEAU; Documentation Department managers: Aurélien DANCOSINE, Laurence FRIGERE; secretarial services: Sandrine BOUVET, Fadela CHEBILI. Working group: Alain BERNARD, Michel CUCHERAT, Isabelle FOURNEL, Hubert GALMICHE, Pascal GIRAUX, Bernard GUILLOT, Jacques MACHECOURT, Patrick MAISON. Review group: Michèle MORIN-SUROCCA, Patrice NONY, Françoise ROUDOT-THORAVAL, Sophie STAMENKOVIC.
HAS would like to thank all the participants listed above.

**Declaration of interests**
The members of the **steering** group communicated their public declarations of interest to HAS. They may be viewed on the website [https://dpi.sante.gouv.fr](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 steering group members were considered to be compatible with their participation in this work.

**Validation**
Version dated 8 June 2021

**Updating**
## Contents

1. **Introduction**  
   
2. **Objective**  
   
3. **Context**  
   3.1. Double-blind, randomised controlled trials  
   3.2. Clinical development phases  
   
4. **Methodological specificities for the clinical development of medical devices**  
   4.1. Timing the assessment  
   4.2. Eligible population and recruitment  
   4.3. Acceptability  
   4.4. Randomisation  
   4.5. Blinding  
   4.6. Choice of control or comparator group  
   4.7. Factors related to operator experience  
      4.7.1. Learning curve  
      4.7.2. Volume of activity  
   4.8. Adjustment factors  
   4.9. Types of analysis  
      4.9.1. Intention-to-treat analysis  
      4.9.2. Per protocol analysis  
   
5. **Experimental studies**  
   5.1. Achieving blinding or compensating for the absence of blinding  
   5.2. Zelen’s design or randomised consent design trial  
   5.3. Expertise-based randomised controlled trials  
   5.4. Tracker trial designs  
   5.5. Cluster randomised trials  
   5.6. Crossover trials  
   5.7. Single Case Experimental Design (SCED) studies [New]  
   5.8. *In silico* studies [New]  
   5.9. Types of analysis  
      5.9.1. Bayesian methods [New]  
      5.9.2. Sequential trials  
      5.9.3. Adaptive trials  
   
6. **Real-world studies or observational studies [New]**  
   6.1. General principles [New]
6.2. Comparative observational studies specific to prospective data collection 45
6.3. Comprehensive cohort studies based on patient preference 48
6.4. Studies based on existing data [New] 50
6.5. Indirect comparisons based on single-arm studies [New] 52
6.6. Types of analysis
   6.6.1. Adjustment methods and propensity score [New] 53
   6.6.2. Instrumental variables 55

7. In practice 58

Table of appendices 60

Bibliographic references 70

Abbreviations and acronyms 81

Glossary 82
1. Introduction

The Medical Device and Health Technology Evaluation Committee (CNEDiMTS) is the Haute Autorité de Santé (HAS - French National Authority for Health) committee which specifically evaluates medical devices (MDs) and other health products with a view to their reimbursement by the French National Health Insurance scheme (Article L 165-1 of the French Social Security Code). It plays an advisory role to decision-makers, recommending the reimbursement of MDs or not (inclusion on the list of products and services qualifying for reimbursement or LPPR, changes in the conditions of inclusion, renewal), helping to determine the conditions for their proper use and their role in the therapeutic, diagnostic or disability compensation strategy.

The CNEDiMTS can only begin its scientific assessment task once CE marking has been obtained (Article R 165-4 of the French Social Security Code). Its assessment is therefore complementary to that implemented for CE marking: beyond demonstrating performance and safety, it focuses on assessing the benefits of the MD for the patient and for public health (actual clinical benefit), along with its place in the therapeutic arsenal available in France (clinical added value).

Assessment of the benefit of MDs must be based on conclusive clinical studies. CNEDiMTS opinions take into account the scientific and medical context at the time of the assessment, along with the available clinical data, primarily from clinical trials, at the time of the application for registration for reimbursement. While randomised, controlled clinical trials are generally considered to be the gold standard methodology to demonstrate the efficacy of a health product, in line with the principles of Evidence Based Medicine, these are sometimes difficult to implement for medical devices. The now well-identified specificities of the sector, such as the rapid pace of development, its operator-dependent or organisation of care-associated nature, and the sometimes very small target populations, imply the development of appropriate methods to provide evidence. Finally, the arrival of technologies incorporating artificial intelligence, the development of access to real-world health data, increasingly rapid technological advances and the dynamic nature of the medical device sector mean that assessment methods need to be adapted, combining an assessment based on appropriate robust methodologies with the context of the technology being assessed.

It is for these reasons, in a constantly evolving context, that the committee wanted to update its guide published in 2013 and initially entitled “Methodological choices for the clinical development of medical devices”. To clarify the objectives of this guide, which is a toolbox to enable companies to build their development plan, the committee has renamed it “Methodology for the clinical development of medical devices”. This updated guide complements the support that has already been made available to companies by HAS for several years. In particular, for complex development programmes, companies may request an early dialogue to discuss a clinical study before its implementation.

More generally, this guide is aimed at manufacturers, research structures, CROs, project leaders and healthcare professionals involved in the clinical development of a medical device or a health technology\(^1\) and who are planning to submit an application for registration for reimbursement of the medical device in question to the CNEDiMTS. Patients are also concerned, since this guide is intended to promote the collection of evidence, whatever the endpoint used (clinical, quality of life, organisational). It aims to provide practical benchmarks relating to methodological aspects in order to optimise

\(^1\) The rest of the text uses the terms medical device and health technology, but the information in this guide is applicable to all technologies falling with the scope of assessment by the CNEDiMTS.
the level of evidence of these studies and increase confidence in their results. The guide presents a review of current methods (advantages, disadvantages).
2. Objective

The aim of this methodology guide, which supplements documents relating to the medical device assessment process in France (1) and the CNEDIMTS assessment principles (2), is to identify the methods to be used (types of clinical studies and analysis designs) when randomisation and/or blinding are impossible to put in place:

- specifying their limitations;
- proposing application examples where possible.

This document should be considered as a methodological aid and is not binding. It does not provide a “ready-made recipe” that can be applied in all circumstances. The diversity in terms of the technologies assessed, the contexts in which they are used and the target populations means that it is not possible to define a single assessment approach. It is up to the sponsor to choose the clinical study method, depending on the nature of the application, the type of device and the target population, as well as the therapeutic arsenal available and the potential place of the technology in the therapeutic strategy for the disease or condition concerned.
3. Context

3.1. Double-blind, randomised controlled trials

Although the conditions of the study design may differ from routine clinical practice, double-blind, randomised controlled trials remain the essential reference for the assessment of any health product. That is because only randomisation and double-blind comparison make it possible to make the two groups studied comparable, with the only difference being the strategy allocated, and thus to attribute the differences observed in the two groups in terms of clinical findings, to the strategies studied (health products, procedures, etc.) in a given treatment regimen. In other words, this method formally enables the results to be attributed to the technology being assessed. Double-blinding enables the evaluation to be free of biases related to the subjectivity of the follow-up, the assessment of outcome measures, etc.

A risk-of-bias assessment tool has been created to assess the potential for bias in randomised trials: the RoB 2.0 tool provides a framework of reference that can be used to evaluate the risk of bias in the results of any type of randomised trial. The RoB 2.0 is intended for application to individual randomised trials, randomised parallel trials, randomised cluster trials and randomised crossover trials (3).

The following five domains are used in the risk-of-bias assessment of individual randomised trials:
- bias arising from the randomisation process;
- bias due to deviations from the intended procedures;
- bias due to missing outcome data;
- bias in measurement of the outcome;
- bias in selection of the reported result.

The choice of a methodology other than a randomised controlled trial must always be explained and justified by the manufacturer (situations such as concomitant developments, special populations for whom extrapolation of efficacy can be made on the basis of observational data, etc.).

The clinical assessment methods presented in this document should be considered when the respective effects of two products, or a medical device (or health technology) and a placebo, are to be compared. These methods are not an indicator of the clinical relevance of the outcome measures used for this purpose. In addition, they can only be applied to data of sufficient quality. Therefore, the studies to be carried out must benefit from a set of measures designed to ensure this (precise protocol filed in a registry and/or published before the start of the study, selection and training of investigators, appropriate data collection method, monitoring, quality control and requests for correction, etc.).

3.2. Clinical development phases

A lack of good quality data is a significant barrier to the assessment of medical devices and health technologies by assessment agencies and decision makers (4), (5).

The clinical assessment of a new medical device’s efficacy, which is the subject of this document, is conducted after the preclinical phase and feasibility studies. The study protocol and clinical data for first-time use in humans are of considerable importance in this context.

Feasibility studies are proposed immediately following the preclinical phase. These help fine-tune the technique and determine the appropriate efficacy endpoints. Depending on the context, one or
more studies may be necessary to answer different questions. In the majority of cases, the most appropriate type of methodology is a prospective non-comparative study.

The choice of efficacy threshold is crucial. Before envisaging studies including a large number of patients, the project leader or manufacturer must verify that their new MD is promising in terms of efficacy. The first step therefore involves choosing an efficacy endpoint based on data in the literature or supported by expert opinion, and determining the probabilities of efficacy and inefficacy. These probabilities will enable rules to be defined for stopping or continuing the study, while including the fewest number of patients possible.

Feasibility studies are useful for (6):

- Selecting patients who will benefit from the new medical device
  This step involves clarifying the clinical forms of the disease or condition and the characteristics of patients who may benefit from the MD.

- Fine-tuning the technique
  For implantable MDs, the implantation technique must be fine-tuned, describing the different steps of surgical procedures, as well as the technical facilities and personnel required. This can only be done in the context of a clinical study. Although the implantation technique may continue to be improved after this step, this should not result in delays setting up a study to demonstrate the clinical benefit.

- Measuring efficacy
  A feasibility study is a prerequisite for constructing the hypothesis to be tested and for calculating the sample size in a comparative study. At this stage of development, determining the primary endpoint is essential/fundamental because it is this endpoint that will enable the efficacy of the MD to be measured. The **endpoint must be clinically relevant** in terms of the disease/condition and the intended action, such as a reduction in mortality or in a clinically measurable complication. If an **intermediate or surrogate endpoint is used, this must have been validated.**

- Complications and risks
  Unlike pharmacological studies that do not involve any medical procedures, two types of adverse events may be reported where MDs are concerned: those directly related to the MD and those related to the implantation or surgical technique. Another objective of feasibility studies is to assess the main complications, which will be documented at every stage, in order to determine what the risk-benefit ratio will be.

These steps are indispensable and provide essential information for subsequent demonstrations of efficacy via randomised controlled trials.
4. Methodological specificities for the clinical development of medical devices

Some methodological principles that are intrinsic to randomised trials on pharmacological treatments may be more difficult to apply when assessing medical devices and health technologies. The problems related to the implementation of double-blind randomised controlled trials on medical devices are detailed below (7-10).

4.1. Timing the assessment

Choosing the most appropriate time in the life cycle of an MD to perform its clinical assessment is an important point in its clinical development. It is preferable to give priority to clinical assessment as early as possible, taking into account potential evolutions in the technology when defining the study protocol, as long as these do not modify its main function. In addition, the new European regulation 2017/745 relative to medical devices imposes reinforcement of the prerequisites necessary to obtain CE marking and of the requirements with respect to the level of the risk-benefit demonstration. Finally, once an MD is widely distributed, it is difficult to get physicians to support a study protocol, because a technique already in use is often empirically considered to be effective (11).

4.2. Eligible population and recruitment

The small size of the eligible population may also be a specific feature of studies on medical devices. In fact, the target population may be small, with a technology potentially only concerning a few hundred patients in some cases (12). In this situation, a traditional parallel-group trial may be more difficult to implement due to the complexity of recruiting patients and its cost.

The choice of study population is important (13). If the selection of eligible patients is too strict, the risk-benefit ratio for the device will be optimised, but the study’s external validity will be more limited. On the other hand, a broader selection can facilitate recruitment and make it easier to generalise from the results, but may fail to define the population most likely to benefit from the new medical device. The choice of inclusion and exclusion criteria is one of the key aspects of the study protocol.

4.3. Acceptability

The acceptability of the study to patients plays a decisive role in its feasibility. Obtaining patient consent is a prerequisite for conducting a clinical trial. During the phase when patients are informed before giving their consent, they must be provided with clear, documented and reliable information. In the absence of patient consent, the feasibility of the study is called into question. Where there are reasons to believe that the risk-benefit ratio between the treatments may be different, both patients and healthcare professionals may prefer a given procedure and refuse to take part in the trial. That is particularly true for new surgical methods that may be used in emergency situations or in the field of paediatrics.

Some patients may have a preference for a given procedure and refuse to be randomised. There may also be issues related to the acceptability of the study to healthcare professionals if the latter are
firmly convinced that the technique they normally use is the best strategy (14). These considerations may impede patient recruitment and make randomisation difficult.

4.4. Randomisation

Randomisation techniques considered to be appropriate include the use of random number tables and computer generation of the randomisation group (15). It is essential that a centralised randomisation process be used. The use of envelopes, including opaque and sealed envelopes, does not guarantee the random nature of the treatment received in open-label trials. These envelopes may be opened and, depending on the treatment offered, physicians may refuse to apply it to their patient, for example. Irrespective of the randomisation method, its procedures must be described in the protocol.

Although it is crucial, the randomisation method is described and adequate in barely half of all non-drug trials, with adequate allocation concealment in less than a quarter of cases (16, 17). Several factors could explain the absence of randomisation in some trials. Firstly, a randomised trial is more expensive than a case series (18). Secondly, randomisation may be judged to be impossible from the outset, usually for practical reasons (patient or healthcare professional preference for a new potentially effective treatment, etc.). More rarely, comparison with an invasive treatment led some authors not to randomise, but this argument is highly questionable (19). The literature actually provides numerous examples revealing that such trials are indeed possible in the majority of cases, including to assess invasive treatments versus non-invasive treatments, such as revascularisation by angioplasty or coronary bypass versus medical treatment (20, 24). Furthermore, comparison with an invasive treatment is preferable to the distribution of medical devices without any evidence of their efficacy and without a robust assessment of their risk-benefit ratio (4, 21).

4.5. Blinding

Blinding is an important element in clinical trials, because it can reduce classification or measurement bias related to the physician’s or patient’s subjectivity. A crucial element of blinding is that it must be impossible to distinguish between the treatments compared. Blinding may involve all or only some of the participants in the care chain: the patient, the physician administering the treatment, the person collecting the endpoint or the statistician. Traditionally, blinding concerns the participants or healthcare professionals caring for patients. That is the usual definition of blinding. The term “double-blind” is used when neither the patient nor the clinician knows which treatment the patient has received. Blinding is complete when the procedure is simulated in the control group with no contact between the operator (usually a surgeon) and the team responsible for the patient’s follow-up. Despite the practical difficulties for interventional treatments, some trials have been conducted with complete blinding, involving an identical placebo, a sham procedure, and no contact with the surgeons.

Boutron et al. summarised the different blinding methods used in non-pharmacological trials published in journals with a high impact factor (22). Blinding may be complete, partial or only concern assessment of the endpoint.

Blinding is more often impossible in non-pharmacological trials than in pharmacological trials (23), for ethical or practical reasons (4). In these situations, in order to evaluate the efficacy of a non-pharmacological treatment as objectively as possible, alternatives have been developed.
In some cases, it is impossible to blind these key individuals. The absence of blinding may lead to biases such as selection, follow-up, attrition or measurement bias, calling the internal validity of a study into question (24, 25).

Reviews of studies having supported FDA approval of a high-risk medical device demonstrated that: in the orthopaedic field, between 2001 and 2015, 76% of studies were randomised, 60% were blinded, with 62% of these double-blind (26), and in the ENT field, between 2000 and 2014, 46% of studies were randomised and 43% were blinded (27). Seventy percent (70%) of the studies concerning therapeutic medical devices submitted to the Berlin ethics committee between 2010 and 2013 were randomised and 38% were blinded (16). Another literature review demonstrated that blinding concerned barely a quarter of patients, 6% of physicians and two thirds of assessors in non-pharmacological trials (17). That is particularly detrimental given that the physician’s influence was also much more marked than in pharmacological trials (17). Yet open-label trials overestimate the effect size by 14% compared to double-blind trials (28).

Table 1. Examples of cases where a blind study is impossible, as per (29)

<table>
<thead>
<tr>
<th>Patient blinding</th>
<th>Surgeon blinding</th>
<th>Follow-up team blinding</th>
<th>Blinded assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>− Appearance or perception of the device</td>
<td>− Surgical procedure</td>
<td>− Adverse effects which are specific and characteristic of one of the treatments administered</td>
<td>− Practical organisation: an excessive number of visits or consultations limiting acceptability to the patient</td>
</tr>
<tr>
<td>− Scar that reveals the type of procedure (e.g. surgery versus non-invasive treatment, laparotomy versus laparoscopy, etc.)</td>
<td>− Appearance or manipulation of the device</td>
<td>− Dye or characteristic marks left by the device</td>
<td>− Problem of acceptability for the surgeon, and also for the patient, if the consultation is carried out only by an external person</td>
</tr>
<tr>
<td></td>
<td></td>
<td>− Ultrasound or radiological appearance suggestive of one of the treatments received</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>− Scar that reveals the type of procedure</td>
<td></td>
</tr>
</tbody>
</table>

4.6. Choice of control or comparator group

The choice of control or comparator group is crucial and an active comparator should be favoured where one exists (30).

This problem does not apply to trials in which the medical device assessed is added to the reference treatment (“on top" or “add on" trials). To validate the new technology intended to be added to the therapeutic strategy (or diagnostic or disability compensation strategy), the on top placebo regimen is used where there is no alternative.

The use of a placebo (or an inactive treatment) may also be acceptable in the following circumstances:

- where there is no effective treatment or procedure;
- where abstaining from a treatment or procedure with known efficacy will lead at worst to temporary discomfort or a delay in relieving symptoms;
- where comparison with an effective treatment or procedure would not provide scientifically reliable results, and administering a placebo does not add any significant risk of irreversible damage (30).
A literature review summarised the main surgical placebos used in non-pharmacological trials (22). For surgery and technical procedures, different methods have been reported, depending on the procedure. Thus, patients may be under general anaesthetic, or a surgical drape may be used to conceal the procedure. In some cases, the procedure was simulated by making an incision similar to that made in the treated group, or by injecting a placebo. In practice, so-called placebo surgery can be used in cases where there is no suitable comparator, i.e. an active reference treatment, and where it involves little risk (24, 31-37). It should be noted that it is also important to standardise pre-operative care (patients or equipment in the same position), perioperative care (duration of procedure, instruments, manipulation or care) and postoperative care. In other studies, the surgeon who performed the procedure is not involved in patient follow-up. This may be a solution in the absence of a possible placebo. Boutron et al. also reported the different placebos possible when using medical devices: placebo prostheses, identical but inactive medical devices, or active devices made ineffective, or simulated use of an MD (22, 38).

Table 2. Examples of placebos used in surgery and for medical devices, as per (22) and (29)

<table>
<thead>
<tr>
<th>Placebo</th>
<th>Surgical techniques</th>
<th>Intraoperative</th>
<th>Postoperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>During anaesthesia</td>
<td>Patients under general anaesthetic</td>
<td>Skin incisions to obtain a similar scar to the study procedure (e.g. coronary artery bypass surgery using an internal mammary artery versus skin incisions alone in patients with angina pectoris, arthroscopic surgery versus skin incisions alone in patients with osteoarthritis)</td>
<td>Scars concealed</td>
</tr>
<tr>
<td>Patients masked with a drape</td>
<td>Sham procedure (e.g. intracerebral injection of foetal cells versus skin incisions and abrasion of the external cortical bone of the cranium in patients with Parkinson’s disease)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.7. Factors related to operator experience

4.7.1. Learning curve

A particular feature of health technologies using medical devices is that the operator’s expertise has an impact on the results of the technique (4). Different levels of experience may lead to different levels of performance when carrying out procedures. A lack of experience may influence the result of the study, penalising the new treatment tested (performance bias) (39).

Therefore, the learning curve for operators must be taken into account when assessing surgical or interventional techniques. During the development of a new medical device, provision must be made for training and learning plans. In fact, the surgeons’ knowledge and their expertise are variability factors (24) that need to be taken into account, for example with a breaking-in period (40). During surgical interventions, the impact of learning may be assessed through indicators (such as the device implantation success rate, the technical success rate of the procedure, etc.).

The variability of the medical device effect according to the operator’s level of expertise should also be assessed. This variability is common and can be illustrated by numerous examples, such as in trials comparing angioplasty and thrombolysis in the treatment of myocardial infarction (41, 42). Therefore, there is a risk that assessment of a new technology versus a control may be unbalanced in favour of the control technique, because of the operator’s experience. An assessment performed too early risks reflecting complications related to learning the new procedure. The study should therefore incorporate the effect of learning, for example by recording training and experience (learning curve) (24). Hence, from a pragmatic point of view, this learning phase should be taken into account in the trial so that any benefit provided by the device or health technology can be accurately assessed.

4.7.2. Volume of activity

After the initial technical expertise is acquired, regular practise by the operator may have an impact. Therefore the annual volume of activity of the operator (and centre) must also be taken into account when assessing a new technology. Hence, almost 70% of studies found a significant association between favourable clinical results and the healthcare professional’s volume of activity (43). This association is even more marked for complex, rarely performed and high-risk surgical procedures (44).

4.8. Adjustment factors

In a comparative trial, and particularly a randomised one, there is no problem in terms of confounding because the other components are identical in both groups (all else being equal). In fact, the use of the health technology must be standardised because it is essential that it can be reproduced in the future use of the MD. But it can sometimes be difficult to separate the effects of health technologies from those of other components of the healthcare system. In reality, MDs are often used in conjunction with other interventions (surgical or diagnostic procedures, monitoring, etc.). Evaluating the effect of the MD itself on the endpoint can therefore be difficult. Thus, surgical interventions are not only dependent on the surgeon (preference, learning and volume of activity, etc.) but also on practices that may vary in different centres, the other members of the team (anaesthetist, nurse), on pre- and postoperative care and on the organisation of care. The use of connected technologies is leading to
major changes in the organisation of care, and even the care pathway, and raises questions about what should be assessed: the MD alone or the MD within a care organisation (4), (24)-(11).

It is important to very precisely define what needs to be assessed in the protocol, in advance and, if applicable, the supervision required for the use of the MD, as well as the organisation of care, in order to avoid biased results linked to types of practices. This standardisation must be extremely detailed, meticulously describing the procedure and the resources required, which must be included in the specifications (13). However, in cases where these practices are not reproducible, the aim should be to collect data on any differences between centres, rather than to standardise the organisation from the outset, so that the analysis can be adjusted for these factors.

4.9. Types of analysis

4.9.1. Intention-to-treat analysis

Intention-to-treat analysis consists in analysing patients in their randomisation group, irrespective of the treatment they actually received and irrespective of their outcome relative to the study. It should be favoured in superiority trials because it is the most conservative approach. It prevents the benefits of randomisation from being lost and limits effects that may not be due to chance but related to the intervention.

4.9.2. Per protocol analysis

Per protocol analysis consists in analysing patients according to the medical device they actually received rather than what was initially scheduled for their initial randomisation group. This strategy can lead to an increase in any difference in effect between the strategies compared. It should be favoured in non-inferiority trials.
5. Experimental studies

5.1. Achieving blinding or compensating for the absence of blinding

When it is impossible to blind healthcare professionals, a blind assessment of the endpoint is envisaged. This guarantees a neutral assessment of the endpoint. In this situation, the assessment is performed by assessors independent of the study who are blinded as to the medical device received. This assessment may also be centralised in the case of laboratory tests, radiological investigations, or excerpts from clinical examination (videos, photos, recorded interviews).

On the other hand, a blind assessment may fail:

- if patients know their randomisation group and need to meet the assessor;
- if the medical device is identifiable by the independent assessor;
- if, in practice, follow-up by the surgeon who performed the procedure cannot be avoided;
- if the documents for the assessment committee are sent by an investigator who knows which medical device has been received.

For surgical interventions, patient blinding can be achieved if the patient is under general anaesthetic or if the procedure is concealed using a drape (22).

In some cases, an adjudication committee independent of the investigators is put in place to check the endpoint. Blinding (or partial blinding) as to the study’s hypotheses is also described as an alternative (22). This involves partial information being given to the participants, who, for example, may not be informed of the existence of a standard treatment, of the study’s hypotheses (hypothesis that one treatment is superior), or of the randomisation outcome for patients under a modified Zelen’s design. In these situations, for ethical reasons patients must be advised that, for scientific purposes, they will not be informed of the study’s specific objective but will have access to all this information at the end of the study.

When blinding or alternatives to blinding are impossible, it is important to:

- favour the most objective efficacy endpoint possible;
- and to conduct a blind assessment of the endpoint (for example by an independent expert committee).

Figure 1 shows a decision tree to guarantee the best level of evidence as regards the endpoint.
5.2. Zelen’s design or randomised consent design trial

Context

In biomedical research and in conventional randomised trials in particular, the patient must be informed of the benefits and risks of participating in the study. Normally, randomisation only takes place after the patient's informed consent has been obtained. Consequently, patients who do not give consent are not included. In addition, some patients who initially consented may withdraw their consent once they know the results of randomisation. For example, a patient may prefer to receive a promising new medical device rather than the control treatment. In order to compensate for these inclusion issues, Zelen suggested only asking for patients’ consent for the new medical device and not for the standard treatment (45). This study design is very little used in practice due to methodological constraints such as selection bias, as well as specific ethical issues due to the lack of information given to the group randomised to receive standard treatment.
Principle

Zelen’s design involves randomising patients without first obtaining their informed consent. Only patients randomised to the new medical device group must sign an informed consent form. Patients who refuse will be given the standard treatment (46).

Zelen proposed three options.

– (1) If they are eligible, patients are randomised to one of the two following groups: group G1 where consent is not requested and in which patients receive the standard treatment and are given information about the treatment; and group G2 where consent is requested and the patients receive the investigational medical device (figure 2) (45).

– (2) When patients are randomised into group G2, they are asked which treatment they would prefer to receive. This approach should be favoured in the event of particularly incapacitating treatment (for example, prostate cancer) (figure 3) (45).

– (3) Patients are randomised to receive treatment A or treatment B. If they are randomised to the treatment A group, they are asked if they wish to receive treatment A. If they refuse, they receive treatment B or another treatment. If they are randomised to the treatment B group, they are asked if they wish to receive treatment B. If they refuse, they receive treatment A or another treatment. This "double-consent" alternative is particularly useful when there is no established standard treatment (figure 4) (47).

Figure 2. Pre-randomisation according to Zelen, case 1 (45)
Figure 3. Pre-randomisation according to Zelen, case 2 (45)

Eligible patient

Randomisation

Consent not requested G1

Standard procedure

Which procedure does the patient prefer to receive? G2

Standard procedure

New procedure

Figure 4. Double consent according to Zelen, case 3 (47)

Eligible patient

Randomisation

Do you want the standard procedure? G1

YES

Standard procedure

NO

New procedure

Do you want the new procedure? G2

YES

New procedure

NO

New procedure

Standard procedure
**Advantages**

Patients know which treatment they will receive before they give their consent, which prevents withdrawal of consent once treatment has been assigned. This study design requires patients to make a decision only when they are to receive the investigational medical device, which helps limit patient stress. Patient recruitment is facilitated, even when patients have a strong preference for one treatment over another. This type of study does not affect patients’ trust in their doctor, insofar as the physician only offers the patient one treatment. This study design can therefore facilitate inclusion of patients and, ultimately, enable more patients to be included than in a conventional trial (45), (48).

**Disadvantages**

A pre-randomised trial poses specific ethical problems, particularly in the event of the single consent design, where a standard treatment is available. This raises the issue of whether patients ought to know how their treatment has been chosen (49). As there is no informed consent in group G1, this design cannot be used in placebo-controlled trials. In addition, many clinical trials require frequent follow-up visits for data collection. These visits are generally more frequent than the usual follow-up for the condition. This does not pose a problem for patients in group G2, who have given their consent. However, it does pose a problem for patients in group G1 when they are informed of more frequent or more invasive visits for data collection and monitoring (50).

As the comparison is made between all patients randomised to group G1 and all patients randomised to group G2, the efficacy of the new medical device may be obscured if many patients in group G2 choose the standard treatment (50). Thus, comparing group G1 with group G2 irrespective of the treatment received dilutes the measurable effect of the new medical device (45). There is therefore a risk that it will not be possible to draw any conclusions.

Finally, a double-blind design is impossible (45) and selection bias is the biggest disadvantage of this strategy. In fact, selection bias is accentuated when patients with a poor prognosis are under-represented in the experimental group, because the refusal rate may be proportional to the severity of the condition². Overall, this type of design more closely resembles an observational study.

**Constraints**

Patients in the control arm must be informed that their data is being collected. It is also important to compare the characteristics of patients in group G1 with those of patients in group G2 receiving the standard treatment after refusing the new medical device, in order to identify any potential selection bias (45).

**Scope of application**

Very rarely used in France, this study design can be useful:

- when patients receiving the standard treatment do not require additional visits and when death is the only endpoint; as well as in the fields of surgery, interventional studies or medical devices;

---

² However, there is no under-representation in the control group since consent is not sought in this group.
for diseases where it is psychologically difficult for the patient to decide on treatment (oncology, palliative care, choosing treatment for one's child, etc.) (50, 51).

5.3. Expertise-based randomised controlled trials

Context

Conventional randomised controlled trials may be difficult to implement, particularly in the interventional field, due to issues of acceptability to the surgeon. Poor acceptability may hamper patient recruitment in a conventional clinical trial (14). Furthermore, surgeons may unconsciously perform the procedure more meticulously in patients randomised to have the intervention they prefer, or they may follow up patients in a different manner to healthcare professionals less experienced in the technique being assessed (14, 52). Consequently, the study protocol must be precise and adhered to rigorously.

These factors highlight the limitations of conventional randomised trials in the interventional field. To compensate for these limitations, expertise-based randomised controlled trials have been developed.

Principle

Unlike a conventional clinical trial, where patients are randomised to receive either intervention A or intervention B provided by the same team, an expertise-based randomised controlled trial involves randomising patients to a surgeon or team that specialises in a given intervention (figure 5) (14). Each surgeon or team only carries out one procedure, at which they are proficient and are experts. This assumes that their proficiency in the procedure has been verified in advance and is comparable in each group (39). This type of trial may have limited acceptability for patients, who must place their trust in two surgeons or two teams.
Figure 5. Diagram of randomisation in expertise-based randomised controlled trials, as per (39)

Advantages

- Improved internal validity

Expertise-based randomised controlled trials have better internal validity (52) for the following reasons.

- No differential expertise bias

Expertise-based randomised controlled trials can reduce the differential expertise bias that exists in conventional trials (53).
• Limitation of bias related to the absence of blinding
This type of trial can also limit differential follow-up bias related to the absence of blinding. In fact, as each surgeon only performs the procedure that they specialise in, the risk of differences between patients as regards the procedure and associated factors will be lower than in a conventional trial, where surgeons may follow up patients differently between the different groups (52). However, there is no guarantee that patients will be followed up and managed in the same way in both groups, which is why these practical details must be specified and standardised in the protocol before the start of the trial.

• Fewer deviations from the protocol
In order to optimise the efficacy assessment of a new medical device, patients must be given the treatment that they were randomised to receive. If this does not happen, these deviations from the protocol adversely affect the internal validity of the clinical trial. In a conventional clinical trial of treatments for tibial fractures (reaming vs no reaming), Devereaux et al. (52) demonstrated different protocol deviations between the two arms, with a higher proportion of patients randomised to the non-reamed procedure group (more technically challenging technique). Such deviations from the protocol were more frequent when surgeons rarely performed the procedure to which the patient was randomised than in the opposite case (54-56). Hence, the risk of differential deviation from the protocol is reduced in expertise-based trials, since surgeons only perform the procedure they are used to carrying out (52).

  – Improved feasibility
The feasibility of an expertise-based randomised trial is better than that of a conventional trial. Surgeons do not need to have been trained in both procedures and will be easier to recruit before the start of the clinical trial (52).

  – Better acceptability to surgeons
A cross-sectional survey showed that 58% of orthopaedic surgeons prefer to participate in expertise-based controlled trials, versus 17% for conventional randomised controlled trials. The surgeons’ preference was influenced by their expertise (14). In addition, there is improved ethical acceptability because surgeons only perform the procedure that they are used to carrying out and in which they are “experts” (39, 52). Overall, surgeons are less reluctant to take part in an expertise-based clinical trial than in a conventional trial (39).

Disadvantages
Blinding is not possible and there is a major risk of measurement bias (52). More specifically, it can sometimes be difficult to tell whether the superiority of one technique over another is genuine, or whether it is related to the expertise of the surgeon performing it (14). Consequently, extremely skilled surgeons must be recruited in each group. This can, in turn, lead to longer waiting times and a longer inclusion period than in a conventional trial (24, 53).
**Constraints**

This type of design requires at least one expert in each randomised intervention at each centre (14). In addition, the initial pre-inclusion consultation must be undertaken by a neutral person who determines patients’ eligibility (14).

Finally, for surgeons to be able to participate in the trial, they must have reached a sufficient level of skill (explanatory trials) or a plateau in the learning curve (pragmatic trials) in order to avoid differential bias related to the surgeon’s expertise (52). This assumes that the new intervention is not being evaluated too early, when there would be a risk of reflecting shortcomings in learning the technique rather than in the actual efficacy of the procedure.

**Scope of application**

Expertise-based trials are potentially useful in all situations where there is a recognised value in taking into account the existence of a learning curve. In fact, when using a new medical device requires experience to be acquired, an assessment performed too early could reflect the operator’s lack of experience rather than the actual effect of the new device. This type of trial may be used when the two techniques are very different and each require a learning curve – or when expertise in the technique substantially influences the outcome – and when they are performed by different types of healthcare professionals (57, 58).

This method may also be used for techniques that have already been disseminated and for which randomisation is no longer possible due to users’ convictions, or when practices differ between centres.

**5.4. Tracker trial designs**

**Context**

Medical devices are often marketed before in-depth studies have been conducted on their efficacy. In addition, they are prone to substantial technological change (59). Tracker trials were developed so that these technological changes or improvements to the procedure could be taken into account during the trial (24).

**Principle**

Changes during the trial are authorised, recorded and taken into account in the statistical analysis. Changes to the study design are also permitted (24).

**Advantages**

Tracker trials can adapt to changes in technology and in clinical practice. They enable early assessment, thus respecting the principle of balance, or “equipoise” (where there is no argument to recommend one treatment over another), and make maximum use of the different data available. This type of study can be used to rapidly assess new, innovative or potentially dangerous techniques (59).
Disadvantages

Analysis is complex because the operator’s experience and treatment developments must be taken into account. This type of trial uses more sophisticated methods than conventional trials. Flexible budgets are required in order to take into account the duration of the trial, in particular (59).

From a practical point of view (24, 59), they are difficult to organise, especially when the technology is evolving because most operators then have little experience (59).

There is little data in the literature on the specific analyses to be performed for this type of trial.

Constraints

Tracker trials must be flexible and include new treatments as and when they emerge. The protocol must be revised regularly in order to be able to integrate new arms for new or emerging treatments or, conversely, to remove some arms. All operators and centres must be included independently of their level of experience. However, this factor must be taken into account in the analysis (59).

Scope of application

There are currently few examples of this type of study available in the literature. It seems to be of most value in the preliminary phase, in order to reject dangerous technologies, and thus prior to the assessment of clinical efficacy, which is the subject of this guide.

5.5. Cluster randomised trials

Principle

Cluster trials involve randomising groups (clusters) of individuals (by centre, hospital or department) rather than randomising individuals directly (figure 6) (60). 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. In practice, the technique that a centre must use is selected randomly. For example, hospital A will systematically use technique 1, and hospital B will systematically use technique 2.
Two types of cluster randomised trial designs have been described: whole cluster trials, and cluster trials with active recruitment (62). In whole cluster randomised trials, one person in each centre is in charge of the cluster and the procedure to be tested is allocated to them.

**Advantages**

This design makes it possible to minimise contamination bias when blinding of treatment is not possible. Randomising clusters rather than individuals helps to prevent any contamination between two techniques available at the same centre (62-64). For example, in the use of a postoperative medical device, the risk of administering device A instead of device B is far lower when only one of the two is used in the hospital, than when both are available. It is also simpler from a logistical point of view. This experimental design is particularly useful for comparing overall or multimodal treatment strategies.
Disadvantages

This type of trial may lack power if the cluster effect is not taken into account when calculating the sample size (62, 64).

Because of their recruitment and follow-up methods, cluster randomised trials may compromise comparability between the groups. For example, some participants may be included in one cluster when they should have been allocated to the other. Without knowledge of which cluster each patient should have been included in, intention-to-treat analysis is difficult and jeopardises the comparability of the groups, which should be obtained at both individual level and cluster level in this type of trial (62). While in trials with individual randomisation units this risk is low because the number of randomisation units is high (frequently more than 100 or 200), this risk is high in trials with collective randomisation units where sometimes fewer than 10 units are randomised (hospitals, care structures or physicians). Obviously, the smaller the number of randomisation units available, the higher the risk of imbalance due to chance between the control and intervention groups (64).

Furthermore, there may be a difference in recruitment between clusters in cluster trials with active recruitment, in terms of the number of patients included and their characteristics (62). An important factor for validity is that all patients are analysed. Finally, there is a risk of contamination between clusters in the event of patient transfers from one department to another.

In whole cluster randomised trials, there may be empty clusters if the person in charge of the cluster refuses to participate just after finding out the result of randomisation (62). There is no statistical method to limit this bias. In addition, the randomisation unit (individuals or groups of individuals) has a much stronger influence in a cluster randomised trial than in a conventional trial (62).

In cluster randomised trials with active recruitment, there may also be inactive clusters if the persons in charge of them do not adhere to the result of randomisation. Moreover, because patients are recruited after randomisation, participants are selected by someone who knows which group they will be randomised to. Patients consent not to randomisation but to participation in a predetermined group; this may result in selection bias and differential recruitment (62). Differential recruitment may be considered equivalent to non-response bias, but the number of eligible participants is unknown and therefore difficult to estimate (62).

Cluster randomised trials also raise ethical issues because they require two levels of consent: healthcare professionals give their consent, and then patients are informed that they are taking part in a study (65). Hence, patients have no choice in their treatment. In trials where a new medical device is compared with the standard treatment, it is not necessary to inform patients in the latter group, since their treatment is the one usually offered. However, where there are additional outcome measures to evaluate efficacy, patients must have been informed in advance and have given their consent. This therefore compromises patient blinding (65). In the absence of blinding, it is not uncommon in cluster studies for hospitals, departments or physicians who are assigned an intervention that does not suit them (e.g. if they are randomised to the control group) to ultimately refuse to participate in the study, exacerbating the risk of imbalance between the groups (64).

Constraints

When calculating the sample size, it is necessary to take into account the cluster effect, which increases the size of the sample required (62). The statistical analysis should also take this cluster effect into account (62).
In fact, this type of study has a hierarchical data structure: a cluster level and an individual level. It is necessary to take into consideration the fact that individuals within the same cluster tend to resemble each other more than individuals in different clusters.

For whole cluster randomised trials, it is necessary to ensure that the person in charge of each cluster adheres to the study protocol before randomisation (62).

In order to limit selection bias in cluster randomised trials with active recruitment, Puffer et al. (66) suggested that all participants should be identified and included before a cluster is randomised. This strategy, which prevents both empty clusters and selection bias, should be imperative but it is very difficult to ensure. In practice, it cannot be systematically implemented for logistical reasons. Other strategies have been suggested, such as blinding recruiters to the allocation group or randomising clusters only once the first participant has been included (index case). The latter option avoids empty clusters but cannot prevent differential recruitment (62).

**Scope of application**

Cluster randomised trials are primarily used for organisational interventions, behavioural interventions or health promotion programmes (64). Insofar as they are likely to improve feasibility or practical aspects, they could be considered for medical devices, or for interventional techniques or surgical techniques. For example, they have already been used to assess different hospital hygiene strategies (51, 67).

### 5.6. Crossover trials

**Principle**

In contrast with parallel group trials, where each patient receives only one treatment, crossover trials involve randomising the sequence in which each treatment is administered (figure 7) (68). Each patient therefore receives both treatments (69), which are thus compared in the same patient during different periods (70).
Advantages

As subjects are their own controls, with the same precision and with measurements correlated between periods, a smaller sample size is required (70), which is particularly useful in the case of rare diseases. In addition, patients may express their preference for one or other of the treatments (69). Finally, the two groups are perfectly comparable because each subject receives both treatments.

Disadvantages

The main concern with crossover trials is the risk of the so-called “carry-over” effect. This effect results from the fact that the first treatment administered may have a residual effect during the second period. Another disadvantage is that crossover trials are not suitable when the endpoint studied is irreversible (recovery or death) (69) or where there is a learning curve. If the measurements conducted in the same patients are not correlated between the periods, the power of a crossover trial is two times weaker than a parallel group trial with an equal number of subjects.

The duration of the treatments being compared is longer than in a parallel group trial, thus leading to a greater risk of loss to follow-up. Measurement of the endpoint may be modified by a “learning” effect for certain assessment tests (measurement bias).
Constraints

All the biases must be discussed (68). Crossover trials work on the assumption that the disease studied is stable, and that there is no residual effect from the first treatment when the second treatment is administered. The implementation of a washout period (time without treatment or withdrawal phase) between two treatment administration periods reduces this risk (69).

The effect of the treatments must not be irreversible so that the subjects are in the same state at the beginning of the second period as they were at the start of the first period.

The disease must not evolve significantly between the two periods. For example, it must not resolve (or worsen) spontaneously before the end of the second period.

There must be no interference between the administration order of treatments and their effect. The effect of a treatment must be the same whether it is administered first or second. To verify this point, it is necessary to test for any interaction between the treatment and the period (69) or any carry-over effect. If there is an interaction, only the data from the first period will be analysed.

From a statistical point of view, specific tests for matched data must be used to take into account the fact that each subject is their own control.

Specific crossover scenarios

Two treatments may be administered simultaneously to the same patient. For example, two topical treatments can be applied simultaneously to two randomly selected skin sites (Figure 8). This type of protocol assumes that there is no possible interaction between the two application sites, such as systemic diffusion of the product administered.

Figure 8. Specific scenario of a crossover trial where two treatments are applied simultaneously

Assessment of eligibility

Randomisation

Site 1: Procedure A  Site 1: Procedure B

Site 2: Procedure B  Site 2: Procedure A

Scope of application

Crossover trials may be used in the assessment of medical devices intended for the treatment of stable chronic diseases, devices where the effects quickly disappear when their use is discontinued and those with a rapid-onset effect. Crossover cluster trials also exist (71).
5.7. Single Case Experimental Design (SCED) studies [New]

Context

Some medical devices are indicated for a very small target population (from a few dozen to a few hundred patients), making it difficult to recruit a large sample of patients for studies aimed at demonstrating their efficacy. Single Case Experimental Design or SCED studies are a set of study designs capable of adapting to this recruitment issue. They also aim to overcome methodological biases related to the heterogeneous profile of the patients and the absence of a control group.

This type of study, which has been used for some fifty years in the fields of education and psychology, is currently being explored in the field of physical medicine and rehabilitation, whether to demonstrate the value of rehabilitation interventions or of certain medical devices, in clinical practice to determine the appropriate management for a given patient or as part of experimental studies for the purpose of generalising the results.

Principle

The SCED methodology is based on three principles (72):

- the intensive and prospective study of a single person or a small group of subjects over time, with each patient serving as their own control;
- repeated and frequent measurement of outcomes throughout every phase of the study;
- sequential and randomised introduction of the intervention during the study.

SCED studies include at least two phases:

- the baseline phase (designated by the letter A), during which the performance of each patient is the subject of several measurements. This phase enables identification of a trend line;
- the treatment or intervention phase (designated by the letter B), during which the repeated measurements continue.

The intervention is introduced sequentially and in a randomised manner for each subject. The sequential inclusion of patients in the protocol aims to control the absence of a retest effect or spontaneous recovery during the baseline phase.

The study design may include more than two phases, with alternating baseline phases and intervention phases (ABAB design). It is not necessary for interventions to be identical.

Three main types of SCED study have been described (72):

- ABAB introduction/withdrawal trials, also known as multiple N-of-1 trials, for cases in which the intervention has an immediate, on/off effect with no carryover, with a short wash-out;
- SCED trials with alternating treatments, for cases in which the intervention has an immediate on/off effect with a short wash-out and when the aim is to study at least two different interventions;
- multiple baseline SCED trials: across subjects (at least 3 subjects), across different settings/contexts (for example, first in a rehabilitation centre, then at home, then at school) or across behaviours (sequential introduction of several modules of an interaction targeting different behaviours). This type of SCED applies when the effect of the intervention
is not immediate and is long-lasting, and the change induced by the intervention is theoretically slow.

Figure 9. Choice of SCED, as per (72)

Measurement of the outcome measure is repeated at least three times (often more) before introducing the intervention, then repeated during the intervention, immediately afterwards and after a longer period, the same number of times during each phase of the study. To be able to reach a conclusion with respect to the intervention’s efficacy, the study design must schedule at least three demonstrations of the effect (replications) at three different points in time. This “three demonstrations” recommendations has no formal basis but is a conceptual standard found in the literature covering the methodology of SCED studies (73).
Advantages
Replication of the demonstration of effect and randomisation of the time of introduction of the intervention and the order of patients contribute to the good internal validity of this type of study. Some authors recommend at least five measurements per phase to reinforce the internal validity and the statistical validity of the result (73, 74).

As each subject is their own control, this type of design means that there is no need for subject matching. In a SCED study, the power derives from the number of repeated measurements and not the number of patients included.

Disadvantages
The question of the generalisation of the results obtained with this type of trial is a key issue that must be specifically considered when designing the study. In order to reinforce the external validity of a SCED study, it is recommended that the study be replicated: after the first patients, the same protocol should be reapplied, in an identical manner, on the same number of new patients, at least three times in several different centres (73, 74).

Constraints
The main constraint is the need to define an appropriate endpoint to assess the effect of the intervention. The SCED methodology generally requires that clinicians create a relevant ad hoc outcome measure that can be reliably measured numerous times.

The choice of the duration of each phase is also an important factor for the success of the study.

A variety of tools have been developed for the design and critical analysis of this type of study:
- the What Works Clearinghouse recommendations for the technical documentation of SCEDs (73);
- a practical guide specific to SCED designed to assess intervention effectiveness in rehabilitation (72);
- the RoBiNT (Risk of Bias in N-of-1 Trial scale) scale to assess the methodological quality of a SCED study (75).

Provided that these rules of study design and execution (procedural compliance) are respected, and even if the examples of applications in a health technology assessment context are still limited, SCED studies currently offer the prospect of studies of good methodological quality with few patients.

Scope of application
In particular, SCED studies could be useful for assessing the efficacy of assistive devices, orthotics and external prostheses in the field of physical medicine and rehabilitation, as well as for assessing other types of MD intended to compensate for deficiencies, such as hearing aids.
5.8. *In silico* studies [New]

**Context**

*In silico* studies are defined as “the use of individualized computer simulation in the development or regulatory evaluation of a medicinal product, medical device, or medical intervention” (11, 76)\(^3\). Therefore they are models used as feasibility studies to guide development strategies. They do not replace clinical trials for demonstration of the efficacy of a health product in humans.

**Principle**

*In silico* studies group together a very broad range of numerical methods that use mathematical models to simulate the effect of a health product on a disease/condition and on a virtual patient population with the help of computer tools. They require data obtained from *in vivo* or *in vitro* studies or from observations in clinical studies, and link them by statistical correlations to structural information. The quality of these studies therefore depends on the quality of the database used (77, 78).

**Advantages**

*In silico* feasibility studies are useful to help define the development strategy for health technologies when a pathophysiological model exists because they are inexpensive and rapid to conduct.

**Disadvantages**

Despite efforts to collect homogeneous data, the validity of these studies cannot be verified, in particular because these *in silico* studies are mostly based on preliminary studies with high risks of bias.

**Constraints**

Currently, no *in silico* study has been able to replace the entire clinical development and health product assessment process for a whole organism. A physiological model of the disease is sometimes required. A combination of *in silico*, *in vitro* and *in vivo* methods remains necessary. These trials are not intended to replace clinical trials in humans but make it possible to optimise traditional clinical trials for the clinical assessment of medical devices\(^4\)(79).

*In silico* studies are currently reserved for feasibility studies and play a guiding role very early on in the health product development process. The aim of this approach is to optimise conventional *in vitro* and *in vivo* studies. For example, it can be used to explore hypotheses using computer simulation, before launching conventional clinical trials, in order to increase the likelihood of their success or identify failures at an earlier stage. They make it possible to fine-tune the clinical development strategy and answer different questions depending on the stage of development: proof of concept, optimisation of clinical studies, decision analysis during development, for example identification of the best

---

\(^3\) How simulation can transform regulatory pathways: [https://www.fda.gov/science-research/about-science-research-fda/how-simulation-can-transform-regulatory-pathways](https://www.fda.gov/science-research/about-science-research-fda/how-simulation-can-transform-regulatory-pathways)

indication, target subpopulation(s) in diseases for which it is difficult to recruit patients, with priority given to rare or paediatric diseases (80, 81).

**Scope of application**

*In silico* studies cannot be used as assessment studies since it is impossible to know their validity. They assist in the construction of studies. In this context, *in silico* model development and validation are encouraged by the FDA and the EMA, in order to identify potential opportunities for the extrapolation of their findings for the clinical assessment of health products.

*In silico* studies can be considered for the following situations (82, 83):
- informing clinical trial designs;
- identifying the most relevant target population for a study;
- predicting the safety of the medical device.

For medical devices, these studies are sometimes used to assess fluid dynamics to predict how blood or other bodily fluids move inside and around the device being tested, or for structural finite element analysis to make sure that the forces exchanged between the body and the medical device will not cause any harm.

At the present time, these trials are among the tools available for the clinical development of health products, used to complement existing tools in order to optimise the clinical development phase of a health product where a pathophysiological model exists. They are currently reserved for feasibility studies only.

### 5.9. Types of analysis

#### 5.9.1. Bayesian methods [New]

**Context**

Bayesian inference is an alternative to the frequentist approach for the analysis of studies (randomised or observational). Bayesian analysis methods can be used in all indications and phases and are not limited to small study populations or sequential trials (84).

In conventional frequentist randomised trials, data from previous studies is used only at the time of study design (i.e. are not directly taken into account in the inference). Thereafter, only information collected during the study is used. Previous results can also be taken into account formally when a meta-analysis is conducted.

**The Bayesian approach takes into account existing information (whatever its nature) and the information collected by the current trial and, above all, makes it possible to determine probabilities** (85). The major point to watch is the calibration of the prior distribution, which must be non-informative, accompanied by a sensitivity study.

In 2010 the FDA guidance considered the use of the Bayesian approach in the field of medical devices to be one of the possible methods, particularly when the potential number of participants in trials is limited (86).
Principle
The Bayesian approach makes it possible to calculate a distribution of the probability of the treatment’s efficacy based on study data and prior distribution. Prior distribution can be informative and influence the final result. In theory, an informative prior distribution makes it possible to take into account prior knowledge about the effect of the device, but runs the risk of strongly and arbitrarily influencing the results of the study. For this reason, it is preferable and expected that non-informative distribution be used.

This involves using existing information (not necessarily chronologically related), which may be provided by the literature, cohorts or other clinical trials. When this information is used, a sensitivity study, which must incorporate a genuinely non-informative prior and a pessimistic prior, is necessary to assess the influence of the choice of prior distribution on the potential results of the trial.

It should be noted that, through the structure of Bayesian inference, the prior distribution, combined with the likelihood (the clinical trial data being part of the experiment), gives the posterior distribution based on which conclusions are drawn.

In a Bayesian clinical trial, the uncertainty about a parameter of interest is described according to a probability distribution, updated during the collection of information in the trial. The probabilities constituting the prior distribution are based on information available outside the trial (data from the literature, expert opinions, cohorts or other clinical trials). Based on Bayes’ theorem, the posterior probabilities are estimated from the current trial, conditional on the prior probabilities.

Bayesian estimation does not provide confidence intervals, but credible intervals based on the posterior distribution and a prior probability of treatment efficacy (85). Unlike the frequentist approach, there are no statistical tests, but results based on the posterior distribution and its credible intervals. It is thus possible to calculate the predictive probability distributions of the parameter of interest in order to reach a conclusion for the trial.

Advantages
The value of an MD must be demonstrated by a study itself. Bayesian methods supplement the information by providing the probability that the medical device may be effective conditional on external knowledge and trial observations. They therefore provide more information for decision making. Similar to a meta-analysis, the use of valid previous data increases the information and precision provided by the trial. In some cases, this method enables a reduction in the sample size, either through the provision of external information or by using adaptive trials. Another advantage lies in the greater flexibility of these techniques in response to changes during a trial.

Thus, the Bayesian method makes it possible to use all available information (both past and present) for a given technology. It also makes it possible to extrapolate performance from one version of the MD to another. Finally, by using information from outside the trial, it can reduce the sample size and hence the time and resources needed for a clinical trial (85, 87).

Disadvantages
The major disadvantage is the risk of poor calibration of the prior distribution, thereby influencing the final result of the trial. This can be avoided by conducting a sensitivity study on the choice of priors before use.
Constraints
Bayesian analyses are based on the construction of prior probability distribution(s) based on information of potentially variable nature and quality. It is therefore important to use methodologists who have expertise in this type of analysis. This upstream work is particularly important.

Scope of application
While the so-called “frequentist” approach of conventional trials remains favoured in the biomedical literature, more and more trials using Bayesian methods are tending to be published, particularly in the field of cardiology (88-94).

5.9.2. Sequential trials
The aim of an interim analysis is to be able to stop a trial early if one of the treatments evaluated proves to be more effective, or if there is a safety problem in one of the arms. To avoid increasing the risk of wrongly concluding that a difference exists where there is none (inflating the alpha risk), appropriate methods must be used.

Different analysis strategies can be used for interim analyses (which must have been scheduled before the start of the study). These strategies include sequential trials.

Context
Sequential trials were developed to avoid giving patients a less effective treatment for longer than is necessary (95).

Principle
The different stages are planned independently of any data previously collected (96).

Various sequential analysis methods have been described:

- the traditional sequential method based on the analysis of pairs of patients (little used nowadays) (97);
- sequential group analyses every n subjects (also little used nowadays) (98-101);
- flexible methods that make it possible to choose when to perform analyses (57-58, 79, 80, (102);
- and continuous sequential analyses, such as the sequential probability ratio test and the triangular test (103-105), (106), (105, 107).

Sequential trials are randomised controlled trials in which the study results from patients included are examined before randomising new patients, and a decision is made as to whether to continue the trial or not. The aim is to avoid ethical problems with continuing the study if the difference between treatments turns out to be more significant than was expected (108).

These analyses are based on graphic methods. The boundaries are determined before the trial on the basis of the alternative hypothesis, alpha risk and beta risk. Patients continue to be included if the line remains between the two boundaries. If the upper boundary is crossed, the null hypothesis is rejected, whereas if the lower boundary is crossed, the null hypothesis is not rejected (95). In both these cases, new inclusions are stopped.
The figures below show the graph-based rules for deciding whether to stop or continue inclusions.

**Figure 10.** Graph-based decision-making rules for the triangular test (closed design), as per (107, 109)

\[ Z \text{ (variance)} \]

The maximum number of analyses is limited because the zone for continuing inclusions is closed.

**Figure 11.** Graph-based decision-making rules for the sequential probability ratio test (open design)
Advantages
This type of interim analysis allows analyses to be repeated without inflating the alpha risk (avoiding wrongly concluding that there is a difference when none exists), while not decreasing the power. From an ethical point of view, these methods have the advantage of being able to stop a trial based on the results of previous inclusions, and to authorise the early conclusion of the study (108). Furthermore, the mean sample size in this type of trial is lower than in conventional trials, independent of the effect size or the power (95, 110).

Disadvantages
Sequential methods have practical limitations. The rule for stopping inclusions is based on just one endpoint. It is therefore necessary to find a single endpoint on which treatment will be assessed (108). Stopping the trial early leads to a risk of having insufficient power for the secondary endpoints, or influencing measurement of the frequency of adverse effects, which would be extremely harmful. Therefore, if researchers are interested in the frequency of adverse effects, sufficient power is needed (95). Finally, the number of subjects is unknown at the start, and the confidence intervals are wider than those obtained with a conventional trial (95).

The boundary approach presupposes that the time from inclusion to measuring the endpoint is short. If a lot of patients are included in a short period, the use of data obtained from only some of them to decide to stop inclusions is not satisfactory (95). This will lead not to a reduction in the sample size but to a reduction in the duration of the trial. The boundaries are defined on the basis of prior knowledge and cannot be changed on the basis of data collected from the sequential analyses (95).

Constraints
The results at each stage must be confidential.

Consequently, an independent data monitoring committee needs to be formed, which will decide on changes to the number of subjects required and on whether inclusions should be continued or stopped, without giving detailed information on the analysis results to people involved in the trial (investigators, assessors, patients) (95). There must be regular inclusions to be able to comply with the necessary frequency of analyses, and good-quality follow-up to avoid delays in updating data. Procedures enabling patients to be monitored regularly with prompt submission of information are therefore required. Long-term follow-up after the trial has been stopped is also essential in order to take into account/measure any adverse effects.

The use of specific methods to calculate confidence intervals (103, 104) and the correction of assessment bias is also required.

Scope of application
Currently, these analysis techniques are particularly useful for rare or orphan diseases and in paediatrics. More broadly, they may enable a treatment strategy that proves to be ineffective, implying a loss of opportunity for patients exposed to it, to be stopped earlier.
5.9.3. Adaptive trials

Context

The objective of these adaptive trials is to maximise the total number of patients given the best treatment. Thus, at the start of the trial patients are randomised with a ratio of 1:1 to each arm, then, as the results start to show a difference between the treatments, this ratio is modified in favour of the group receiving the treatment that seems to be more effective (108).

Adaptive trials should not be confused with trials where the allocation ratio is adapted on the basis of the preliminary results of the trial (111).

Principle

Like sequential trials, adaptive trials are based on the use of interim analyses. But in this type of trial, the study design may be changed on the basis of the results observed during interim analyses (95) without increasing the alpha risk (112). All the information collected during the different stages is used, and not just that from the current stage (96). The allocation of a treatment to a patient depends on the results obtained in previous patients (112).

If changes are made, a new phase of the trial starts, and the analysis concerns the data accumulated in this new phase, and no longer all of the data. The various adaptations possible are a reassessment of the sample size, adaptation of endpoints, and the addition or deletion of interim analyses (95).

Adaptive trials may consist of two phases or several phases (113). Two-phase adaptive trials involve initially selecting a sample of size N1. Depending on the results of this initial phase, either the study ends, or it enters the second phase. Adaptive trials in several phases are an extension of two-phase adaptive trials.

Advantages

Adaptive trials offer significant flexibility. Furthermore, all the information collected is used for the analysis (95, 96). In addition, it has been shown that this study design makes it possible to correctly treat a larger number of patients in comparison with sequential trials and randomised trials (108, 114).

Disadvantages

Adaptive trials have more logistical constraints than other types of trials (96). Great care is also required as regards re-assessments of the sample size.

Practical difficulties may be encountered, but can be avoided by agreeing to a trial with several phases, where each phase uses a different ratio for randomisation between the groups (108). As with sequential trials, a simple and unequivocal efficacy endpoint needs to be defined (108). Van der Lee raised the issue of how to define rules for stopping the trial or how to interpret the results if the primary endpoint has been changed during the study (95).

The internal validity of adaptive trials has also been called into question. In fact, according to Armitage, this type of study is not strictly speaking randomised and has the same disadvantages as historical controls (97, 108).
Because of these disadvantages, adaptive trials have been judged to be of little benefit in comparison with sequential trials (108, 115, 116).

**Constraints**

As with other methods that use interim analyses, the results of adaptive trials can be distorted if investigators have knowledge of the trial's interim data. In addition, such bias cannot be corrected by statistical adjustment, which jeopardises the interpretation of the results.

Consequently, an independent data monitoring committee needs to be formed, which will decide on changes to the sample size and on whether inclusions should be continued or stopped, without giving detailed information on the analysis results to people involved in the trial (investigators, assessors, patients) (95, 117).

Finally, it is crucial for the credibility of the final conclusions that hypotheses and adaptations made as the trial is conducted are recorded prospectively (95).

**Scope of application**

Adaptive trials may be useful for operator-dependent techniques, in surgery, in studies of interventional or surgical techniques that require a learning or fine-tuning phase, and in the assessment of medical devices.
6. Real-world studies or observational studies [New]

6.1. General principles [New]

Context

The methodological term for this type of study is “observational studies” because, in theory, a population or a phenomenon is simply “observed” without intervening, or intervening as little as possible, in their natural evolution. This type of data, which is not collected in an experimental setting but generated during routine care without any intervention in the usual patient care arrangements, is also referred to as “real-world data” which therefore a priori reflects routine practice.

“Real-world studies” or observational studies are of obvious interest as complementary studies to experimental studies to collect data in real-life use situations. 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 very different 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.

Taking into account 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 relevance of the assessment of health products for the benefit of patients will only be reinforced (118). HAS has developed a guide specifically dedicated to “Real-world studies for the assessment of medicinal products and medical devices” (119, 120-121)

Principle

Real-world studies can address very different issues (clinical or epidemiological, assessment, evaluation of burden of disease, etc.). They can be used to compare groups of treated subjects with untreated subjects (or subjects treated with an alternative), in order to demonstrate the association between the treatment under consideration (device) and the evolution of the disease. They may make it possible to measure the impact of a device on morbidity and mortality or quality of life under less controlled conditions than traditional interventional clinical trials (122).

These anonymised real-world data may come from several sources: specific observational studies created for the circumstance based on ad hoc collection, studies drawing on existing databases, registries, data from connected devices or collected on the Web, or combinations of these collection methods (such as cohorts linked with data from medico-administrative databases), etc. (121).

In a comparison situation, it is absolutely essential to clarify the trial to be emulated ("The Target Trial") i.e. the hypothetical randomised controlled study that should ideally be conducted in order to meet the single causal objective identified (“emulation of the target trial”) (123-127).

The current methodological standards for real-world studies must take into account the following points:

- one single primary endpoint must be defined per study;
- inclusion and exclusion criteria should be carefully identified; and it is preferable to include patients who are new users of the device;
- it is preferable to choose an active comparator (the decision to treat excludes the time before treatment);
- the choice of variables to be taken into account is crucial (the choice of variables associated with the prognosis of the endpoint is more important than the variables differentiating the groups being compared); and the variables must be of good quality (standardisation, for example);
- checking for residual confounding bias should be systematic;
- the time for the start of follow-up (T0 or D0) must be adequately defined (risk of immortal time bias especially when compared to a placebo (“no specific treatment”));
- the analysis should preferably be carried out on an intention to treat (ITT) basis;
- the statistical analysis plan (SAP) should be prepared in advance of the analysis.

**Advantages**

In general, real-world studies may include patients who could not be included in randomised trials because they did not meet all the inclusion criteria (especially in the group not treated with the health technology, which is not beneficial in the case of analysis methods measuring a treated average/average treatment as well as matching). They can provide data on small populations that may not have been included or may have been under-represented in randomised clinical trials. They enable longer follow-up. They may reveal complications that could not have been observed in clinical trials. They increase the external validity of the results (119, 120).

If the conditions are comparable and the methodologies are well designed, clinical trials and real-world studies can lead to similar results. A Cochrane Collaboration study in 2014 concluded that the results found in well-conducted real-world studies may be comparable to results from randomised controlled trials, but not always. There are examples for which the results of real-world studies and randomised trials diverge (128-130).

**Disadvantages**

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 HAS. The data sources that can be used are diverse, in terms of their origin (medical records, medico-administrative databases, hospital/pharmaceutical activity data, registries, observational studies, adverse event registers, epidemiological, demographic and environmental data, patient reported outcomes, health surveys, social media, etc.) and their quality (131).

**Constraints**

The methodological choices to be made in real-world studies are crucially important to optimise the level of evidence of these studies, since the approaches may be different depending on the question asked (132, 133). There is no perfect method, and each has its advantages and disadvantages. The choice of method must be discussed and justified based on the context of the study (disease/condition and health product studied, administrative and regulatory constraints, data available in existing databases, etc.). In practice, all possible biases should be discussed at the study protocol design stage and the method adopted to control them should be presented. If it is difficult or impossible to control them, a sensitivity analysis, scheduled in the protocol, should
make it possible to determine in what direction and to what extent these biases could have modified the results and conclusions of the study (122). All biases should be assessed using appropriate tools (for example ROBINS-I) (134, 135).

Scope of application

HAS has also published a guide dedicated to real-world studies in order to optimise the level of evidence of these studies and confidence in their results and which details their scope of application.

Real-world studies primarily concern post-marketing studies in France and are mostly intended to support applications for indication extensions or inclusion renewals. For innovations in particular, since early data is often limited, the collection of such data is essential to combine early availability and assessment (136).

It is nonetheless pointed out that real-world data is also taken into account in the first assessments of health products by HAS, in particular those generated in the context of the national innovation funding mechanism or temporary funding of medical devices. In addition, 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.

Figure 12. Why implement a real-world study?, as per the HAS guide on “Real-world studies for the assessment of medicinal products and medical devices”

6.2. Comparative observational studies specific to prospective data collection

Context

Cohort studies are typically used for epidemiological research purposes. However, they can be used when randomised controlled trials cannot be considered for ethical reasons, for example, during the development cycle of a medical device (108), in the case of post-inclusion studies with a view to renewal of inclusion on the LPPR or in the context of indirect comparisons.
Principle

Cohort studies enable comparative analyses to be conducted in order to determine whether patients “exposed” to the health product, compared to “unexposed” patients (or exposed to comparators), have a lower, equal or higher relative risk of presenting the event considered with the possibility of causal inference of the relationship (if there is appropriate control of biases, in particular, control of confounding factors).

Cohort studies include a sample of patients with common characteristics, representative of the target population, who are followed up longitudinally in order to observe the occurrence of health events over time in this defined population. They compare the incidence of the outcome measure between a group receiving medical device A and a group not receiving medical device A (either receiving another intervention or receiving nothing). The distribution of patients in each of these groups is not determined by the investigator; it is observed data (137, 138).

Figure 13. Example of a prospective observational study comparing the occurrence of postoperative infections depending on the type of suture used

<table>
<thead>
<tr>
<th>Constitution of groups: Observation</th>
<th>Follow-up</th>
<th>Observation: occurrence of the endpoint?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suture A</td>
<td></td>
<td>Operating site infection?</td>
</tr>
<tr>
<td>Suture B</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Advantages

Cohort studies 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 efficacy in routine care practice. They can be used to create groups of exposed subjects of sufficient size and can also take into account time-related phenomena.

A prospective design makes it possible to plan collection of all the information required. If properly conducted, these studies can limit information bias, when exposure information is recorded prospectively and in all subjects at risk of exposure. They can also be used to collect specific information about the intervention (e.g. doses administered in an imaging device or duration of an intervention). In addition, it is possible to measure the incidence of an event or a relative risk and to take into account the time to onset of the event after the intervention.
**Disadvantages**

The main risks of real-world studies are selection biases, which occur when the probability of inclusion is influenced by the disease/condition and the exposure. (139-141). Their main limitation is the non-comparability of the groups: since exposed and unexposed patients are not - by definition - randomly selected, the different characteristics of the two groups may, at least partially, explain the observed differences (8, 142). So, while the objective would be to be comparable in terms of the variables liable to influence the primary endpoint, the main limitations are primarily related to the erosion of representativeness over time. In particular, depletion-of-susceptibles bias is a selection bias related to the past treatment history of eligible patients in the study. It can lead to an overestimation of the benefit of a device (or an underestimation 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 using the medical device of interest for the first time, a relevant approach (143, 144). 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 and the actual initiation of treatment. 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” (124).

The feasibility of this type of study should be discussed depending on the level of diffusion of the technology. In fact, in the case of new technology, if the product is very attractive it will be used almost exclusively in all patients (there will be no more untreated patients to serve as controls). As a result, these studies can be cumbersome and costly to implement (145, 146). In addition, for these studies, patient monitoring needs to be strengthened. Finally, since the data is collected within a particular health facility and health system, and that health system may have distinct practices, the analyses should be interpreted in the context of the health system in which the data was collected, and stratified by country in the case of international studies.

**Constraints**

The reliability of these studies is based on a rigorous methodology. The implementation of the protocol, the analyses and the interpretation of the results must take into account possible biases (145-150).

The presence of dedicated personnel to check the quality of the data should be systematic. Where follow-up is particularly long, procedures need to be put in place to limit the risk of loss to follow-up. The observation period must be defined. It is therefore recommended that the following points be observed (138, 151-155):

- identify all potential confounding factors when planning the study (through a causality graph);
- limit the number of patients lost to follow up and missing data as far as possible by minimising the burden of participation in the study for both patients and physicians, through adequate and sufficient data monitoring (including active investigation of the vital status of patients);
- and systematically provide for a register of non-inclusions;
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 routine care practice;

in the statistical analysis, in order to support representativeness, compare the characteristics of participating and non-participating investigators (also compare them at national level), 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). Depending on the case, use imputation techniques or sensitivity analyses and use methods to take into account confounding factors (e.g. multivariate model with matching and adjustment on confounding factors or on propensity scores);

carry out quality control, data management and audits.

Scope of application

Observational comparative studies have a clear value in generating evidence outside the context of RCTs and are a potential source of knowledge about the effect of medical devices under less controlled conditions than traditional clinical trials (156-160).

Where a randomised trial is not feasible, they can be used to compare very different interventions, such as surgical techniques, medical devices or others (with the limitations discussed above). Furthermore, cohort studies are mainly reserved for the analysis of relatively common diseases or conditions. The data from these studies can be used in a long-term assessment for a renewal of inclusion for example. Historical prospective cohort studies are inappropriate for evaluating new indications (9, 151, 161).

6.3. Comprehensive cohort studies based on patient preference

Context

In trials comparing a surgical treatment and a non-surgical treatment, patients may perceive the non-surgical treatment group as a less effective intervention, and refuse to participate in the study (162). Excluding patients with a strong preference for one or other of the treatments could weaken the external validity of the results. Studies with a comprehensive cohort design based on patient preference have been developed to overcome these obstacles.

Principle

All patients meeting the eligibility criteria for a trial are recruited, independently of whether they consent to randomisation (162). Patients who refuse to be randomised receive the treatment of their choice (figure 11).

5 Before using an imputation method, it should be checked that the missing data are due to chance alone.
Advantages
This type of study can improve the recruitment of patients participating in a clinical trial. In addition, information about the acceptability of treatments can be obtained, and the risk of selecting non-representative patients due to recruitment difficulties is reduced. Finally, it reconciles the advantages of randomised trials with the opportunity to test the association between patient preference and the endpoint (164).

Disadvantages
The allocation of patients to one procedure or another depending on their preference has the same limitations as those noted in observational studies: there is no guarantee that patients are comparable as regards variables that have not been measured. This type of study is also subject to selection bias.

Constraints
The randomisation variable is introduced into the model as an adjustment covariate.
Scope of application

This type of study can improve patient recruitment when it is difficult to obtain consent for a conventional randomised trial. It could be particularly useful for surgery, interventional procedures and medical devices.

6.4. Studies based on existing data [New]

Context

Already existing databases (“secondary data”), such as medico-administrative databases, can be used in studies with a very good cost/quality ratio since the data are already recorded.

The main secondary data sources are:

- medico-administrative data or other databases integrated in the Health Data Hub in France;
- 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;
- databases from other community healthcare players (pharmacies, medical biology laboratories and imaging centres);
- databases produced via connected self-measurement medical devices (heart rate, blood glucose, etc.) or via direct input of information by the patient into applications;
- databases from academic registries or cohorts.

The Health Data Hub and the French national health data system (SNDS)

The Health Data Hub is the French health data platform ([https://www.health-data-hub.fr/](https://www.health-data-hub.fr/)) created by the law of 24 July 2019 relative to the organisation and transformation of the health system. This data comes, in particular, from the main database of the French national health data system (SNDS created by the law of 26 January 2016 relative to modernisation of the health system), but also from healthcare facilities, Santé publique France, etc.6,7.

The SNDS gathers and makes available health information collected by public bodies. This merger of several databases currently involves three existing databases: the *système national d’information inter-régimes de l’Assurance maladie* (Sniiram - French health insurance inter-scheme IT system); data from hospitals and other healthcare facilities (*programme de médicalisation des systèmes d’information* – PMSI (French National hospital discharge database)); statistical data relative to causes of death (BCMD). Then, once they have been constituted, the SNDS will integrate two additional databases: “medico-social” data from *maisons départementales des personnes handicapées* (regional disability centres - MDPH); a representative sample of reimbursement data per beneficiary transmitted by mutual insurance companies.8

---

6 [https://www.snds.gouv.fr/SNDS/Accueil](https://www.snds.gouv.fr/SNDS/Accueil)
8 [https://www.snds.gouv.fr/SNDS/Composantes-du-SNDS](https://www.snds.gouv.fr/SNDS/Composantes-du-SNDS)
The SNDS contributes to research, studies, evaluation and innovation in the fields of health and medico-social care\(^9,10\) (165).

**Advantages**

The use of secondary data sources can speed up the availability of results, limit duplication of data collection and thus ultimately improve data collection. It can also foster public/private partnerships.

Medico-administrative data are population data and not a sample. They offer the possibility of follow-up over long periods of time in large populations, with no drop-outs during follow-up, and good consistency of coding. The missing data rate is low.

**Disadvantages**

The factor limiting the implementation of database studies is the existence of the data sought in existing databases. Existing items are predefined and therefore not very flexible. A lot of clinical information is not measured. There is no data concerning risk factors, quality of life, etc.

In addition, the data available in the SNDS is currently updated annually, which may delay access to health data for assessment purposes in a timeframe compatible with decision-making.

**Constraints**

The quality of the information entered into the databases is essential: quality indicators to document and evaluate these data are needed (166). Optimising the use of available databases is important (improving information collection, declarative registry of health databases, querying tool enabling actors to navigate between databases, etc.). The transparency of the methods used to analyse health data is essential (167). Wherever possible, a multibase study should be preferred (as opposed to using a single base implying a risk of incidental findings) since this enables testing of coherence between analyses and increases the reproducibility of the study results.

**Scope of application**

In France, the set-up of the national health data platform (GIP Health Data Hub) will expand the use of real-world data in compliance with the requirements of the General Data Protection Regulation (GDPR). The next step will be the development of data generated in the course of patient care (168).

In cases where the MD is connected to databases, it 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\(^11\). Data collection directly from patients is a fast-evolving field, either via connected self-measurement devices or through direct patient input into dedicated applications or platforms.

\(^{9}\) [https://documentation-snds.health-data-hub.fr/introduction/01-snds.html](https://documentation-snds.health-data-hub.fr/introduction/01-snds.html)

\(^{10}\) [https://www.cnil.fr/fr/snds-systeme-national-des-donnees-de-sante](https://www.cnil.fr/fr/snds-systeme-national-des-donnees-de-sante)

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)\(^\text{12}\) (118).

6.5. Indirect comparisons based on single-arm studies [New]

**Context**

External comparative studies should be reserved for situations where randomisation is not feasible. In the absence of direct comparison, an indirect comparison, conducted on the basis of defined and validated methodological principles, may be taken into account. The absence of direct comparison with a clinically relevant comparator must be justified by the company and may be accepted in certain situations (169-171).

**Principles**

The use of single-arm studies to demonstrate the benefit of a new medical device requires an implicit comparison that will necessarily involve an external comparison reference (e.g. historical comparison). The external comparison must be formalised in the study protocol with prior choice of the comparison reference as well as the formal comparison method: test against a standard, adjustment with individual data, synthetic control group or Matching-Adjusted Indirect Comparisons (MAIC) (172, 173). Possible external comparators may be a reference value, set by a “regulatory requirement” or derived from a systematic review of the literature using published data or individual data from a cohort or randomised trial.

**Advantages**

Single-arm studies with external comparisons can be used when the target population is small and inclusion is difficult.

**Disadvantages**

To be acceptable, these studies must be able to guarantee the absence of residual confounding factors. This comparison is limited by 1) the *post hoc* choice of the reference for the comparison, 2) the confounding bias, which necessarily requires an adjustment approach, and 3) other biases: selection, measurement and missing data among others.

For the control arm where retrospective data collection is implemented, the following limitations should be taken into account:

- modification over time of the therapeutic strategy, changing the disease history;
- increase in operator and patient experience over time;

\(^{12}\) [Link to ARIIS website](https://gt2.ariis.fr/sorienter-vers-les-donnees-pertinentes/la-forme-que-pourrait-prendre-un-tel-outil/)
different patient characteristics at inclusion between a historical control group and a contemporary group (151, 161).

Experience shows that finding a suitable comparator and taking all biases into account can be much more complicated and time-consuming than conducting a randomised trial.

**Constraints**

In order to limit bias, it is necessary to define the external control in advance and justify it, i.e. to ensure an unbiased choice of the reference value of the external control (based on a systematic review of reference studies, taking into account the statistical uncertainty of these reference values and retaining the most unfavourable value), identify all potential confounders (prognostic factors of the endpoints and modifiers of the treatment effect) by systematic review, take into account these confounders and potential biases, and perform sensitivity analyses.

**Scope of application**

Data from real-world studies can also be used, under certain conditions, as a control arm in an indirect comparison study, particularly where there is a registry or systematic data collection. However, this use is difficult when the target patients for a new MD differ substantially from those of the reference therapeutic strategies. In the field of MDs, single-arm studies can enable, for example, the assessment of the evolution of a device, the interest of which was initially demonstrated by a randomised clinical trial, or when the number of patients available is too small to carry out randomised clinical trials (174).

**6.6. Types of analysis**

**6.6.1. Adjustment methods and propensity score [New]**

**Context**

In observational studies, treatment is not allocated randomly and the investigator plays no role in allocating treatment. The investigator merely observes what happens between two groups, one receiving treatment A and the other receiving treatment B. There is no guarantee that the groups are comparable. The propensity score is a statistical method that aims to reduce prognostic differences between the compared groups associated with non-randomisation in observational studies (175).

**Principle**

The propensity score can be defined as the probability of a patient receiving treatment A rather than treatment B based on their observable baseline characteristics (propensity to receive a treatment based on their characteristics) (176).

The objective is to balance the distribution of covariates related to prognosis of the endpoint (such as age, gender, comorbidities and severity of disease) between the groups, in order to neutralise confounding factors (176, 177). Ideally, adjustment should be made for all confounding factors affecting the dataset and assessment of the primary endpoint.
The validity of this method is based on two assumptions: (1) the assumption of conditional independence on observable characteristics and (2) the common support condition assumption. The first assumption means that membership of the treatment group should not depend on the outcome conditional on the characteristics considered in the model. The second assumption is that the individuals in each comparison group are sufficiently similar for the comparison to be meaningful. This requires that the distribution of the propensity score of each of the two groups overlap sufficiently. The absence of common support would preclude any conclusion as to the causal effect of a treatment.

The propensity score method involves several steps:

- (1) estimate the propensity score;
- (2) verify the existence of a common support. The aim is to verify the existence of a common support (overlap) in order to ensure that it is possible, for each individual in the group with treatment A, to find at least one participant in the other group B with similar characteristics;
- (3) verify the balancing property. The quality of the balance between the groups on the variables associated with the outcome should be checked. The aim is not to compare the variables related to the endpoint between the two groups by means of a comparative analysis, but to check that there is not too great an imbalance on these variables. Verification of the balancing property assumes that the percentage of imbalance between the groups on the variables associated with the endpoint must be sufficiently reduced. The 10% threshold is used to judge whether the balance between the two groups is acceptable.

Various methods exist to take into account the propensity score in the analysis (142, 178-182):

- (1) regression model with propensity score adjustment;
- (2) stratification on the propensity score;
- (3) matching of subjects between the groups: the subjects receiving treatment A must be matched with similar subjects receiving treatment B. This involves forming pairs of subjects who have the closest possible propensity scores. The disadvantage of this approach is that it reduces the sample size to the number of matched subjects, resulting in lower power, and not all subjects are included in the analysis;
- (4) inverse probability of treatment weighting (IPTW)\(^{13}\), which aims to create a “pseudo-population”, allocating a weight to each subject belonging to the treatment A group and to the treatment B group. The advantage of this approach is that it retains the effect on the whole population. This is currently the most widely used method.

The propensity score method works best under three conditions (177):

- when the event studied is rare;
- when there are a lot of patients in each group;
- when a large number of covariates have been measured.

**Advantages**

The propensity score can neutralise known confounding factors for measured variables by making groups comparable for a given score.

---

\(^{13}\) Another method is standardised mortality ratio (SMR) weighting.
Disadvantages
The use of the propensity score does not guarantee that confounding will not remain, unlike a well-conducted randomised trial. Groups can only be balanced on known confounders and variables collected and included in the propensity score estimation (142, 177). When a large amount of data is missing for one of the covariates, the validity of the propensity score analysis may be difficult to interpret (151). There may also be over-adjustment on factors that are not confounders (colliders). The risk of over-adjustment is not specific to the propensity score; it is a problem with all adjustments. However, sensitivity analyses and diagnosis of residual confounding bias followed by recalibration can help to assess the robustness of the results (175, 183).

Constraints
The use of the propensity score is based on three strong assumptions that need to be verified:
1) conditional on the covariates, each subject has the same probability of being treated;
2) all confounding variables are identified, observed and measured;
3) the conditional independence of treatment assignment and the endpoints is verified.
Propensity score analysis must be planned in advance. All relevant variables to be collected must be specified in the protocol. The number of participants required should be calculated before the study starts, taking the propensity score analysis into account.
When the results are analysed, it should be verified that the distribution of the propensity score is comparable in both groups (177).

Scope of application
Propensity score analysis is very useful in observational studies when a randomised trial cannot be conducted due to medical practices, patient preference or the organisation of healthcare (154, 184, 185). This analysis strategy could be used for medical devices, in surgery and for interventional techniques.

6.6.2. Instrumental variables

Context
A complex technique to implement, instrumental variables have been widely used in econometrics for years, but they are difficult to use in the field of health (186). Instrumental variable analysis enables observational data to be exploited to estimate the efficacy of a treatment, even in the presence of unmeasured risk factors (187).

Principle
An instrumental variable is a variable that is strongly associated with the treatment indication (such as the prescriber’s preference), but which is not related to risk factors for the occurrence of an event and does not directly affect the endpoint (187) (figure 15).
In the instrumental variable approach, treatment is considered to be confounded by the indication. Thus, patients may be selected to receive one of the two treatments because of known or unknown prognostic factors. Unlike the propensity score, which aims to adjust for known confounding factors, the instrumental variable approach aims to eliminate bias resulting from unknown confounding factors (188).

The instrumental variable approach starts by identifying the instrumental variable (189).

**Advantages**

The instrumental variable approach reduces the impact of confounding factors in observational studies (189) and can take into account unmeasured variables (187, 188).

**Disadvantages**

Estimation of the treatment effect is based on strong hypotheses, which limit the use of this approach in practice and are also difficult to verify.

In addition, the effect of treatment may not be generalisable to a population of patients whose treatment status has not been determined by an instrumental variable. The treatment effect is estimated for “marginal” patients, those who are compliant (190). The effect on the general population may be different.

Overall, when there is a small difference in the probability of receiving a particular treatment between groups of patients defined using an instrumental variable, differences in the occurrence of the event due to the differential use of this treatment may be very small and, consequently, difficult to assess (142).

Finding valid instrumental variables is extremely difficult. In fact, most variables that have an effect on treatment can also have a direct effect on the event (191). In addition, a good instrumental variable estimator relies on large numbers of patients (191). Where numbers are small, the estimates are incorrect. Finally, interpretation is difficult, particularly when the treatment effect is heterogeneous. These disadvantages mean that this technique is difficult to implement.

**Constraints**

A good instrumental variable must satisfy three major assumptions (187):

---

**Figure 15. Instrumental variable, as per (187)**

![Diagram showing instrumental variable and its impact on confounding factors, treatment, and event](image)
it must have a strong relationship with the medical device indication; this must be assessed without bias;
- it must not correlate with any confounding factors (measured or unmeasured);
- it must not be related to the event directly, but only through the effect of the medical device.

**Scope of application**

Instrumental variables are frequently used in econometrics. However, the inherent constraints of these variables make it difficult to consider their use in the context of medical device assessment and hence they are very little used.
7. In practice

The choice of clinical evaluation plan depends entirely on the context: the action sought by the use of the medical device, the existence of alternatives, the size of the target population, etc. These different parameters must be taken into account when choosing the type of study to be implemented. On the basis of these elements, the rationale for the decision should be supported by an explicit scientific rationale. That is the key to an optimised development strategy.

For any study, the quality of the data collected or used is one of the essential prerequisites in order for its results to be taken into account by the CNEDiMTS in its medical device and health product assessments.

Irrespective of the type of study chosen, the principle of transparency - which is to systematically make public all information on clinical research, while ensuring personal data is protected - must be respected at all times. It helps to build confidence in the players using the data by providing information on the protocol and any amendments made to it, the data sources and analysis methods used (described in sufficient detail to enable replication), as well as the results of the study.

This is particularly important for the assessment of innovations likely to bring a real benefit to patients because a continuum in their assessment is often necessary between experimental data and real-world data.

Double-blind, randomised controlled trials remain the gold standard for the assessment of all health products.

If a randomised controlled trial cannot be conducted, then the choice of study type may be guided by various elements.

- The characteristics of the medical device
  - Substantially and rapidly evolving technology
    - Trial with adaptive randomisation
    - Bayesian methods
    - Possibly a tracker trial design
  - A small target population
    - Bayesian methods
    - Crossover trials, as long as the disease studied is stable and the endpoint can be repeated
    - Single Case Experimental Design (SCED) trials
    - Sequential trials
  - Potentially serious adverse events
    - Sequential trials

- Medical acceptability
  - Proficiency in the technique influences the result or the techniques are very different
    - Expertise-based randomised controlled trials
  - In some circumstances, cluster randomised trials
Acceptability to the patient

- The comparator is an invasive technique or the technique is in widespread use
- Zelen’s design or randomised consent trial
- Trials based on patient preference or comparative observational studies if randomisation is impossible
## Table of appendices

| Annexe 1. | Work method                               | 61 |
| Annexe 2. | Literature search                         | 62 |
| Annexe 3. | List of figures and tables                | 69 |
Annexe 1. Work method

The main stages are described below.

Framing phase
A roadmap was drafted and submitted to the CNEDiMTS for approval prior to publication on the HAS website.

Data collection and review
A systematic scientific literature review was conducted: the documentary search applied is detailed in Appendix 2.

Formation of the review group and guide review stage
A review group made up of methodologists was set up.

The guide was sent to the different members of the review group. Following the review stage, the various comments were reviewed by the steering committee for any amendments and validation of the definitive version of the guide.

Review by the CNEDiMTS
The finalised guide was reviewed and approved by the CNEDiMTS.

Document publication
The finalised guide is published on-line on the HAS website.
Annexe 2. Literature search

Method

The search targeted all subjects and all types of studies defined in agreement with the project manager and was restricted to publications in English and French.

The initial search covered the period from January 1999 or January 2013 to October 2019. Documentary monitoring was then conducted until the end of February 2021.

The following sources were queried:
- for the international literature: the Medline database;
- the Cochrane Library;
- websites publishing guidelines, technological or economic assessment reports;
- the websites of learned societies with expertise in the field studied.

This search was supplemented by the experts' bibliography and the references cited in the documents analysed.

Bibliographic databases

The bibliographic database search strategy is constructed using, for each subject, either thesaurus terms (descriptors), or free-text terms (from the title or the abstract). They are combined with the terms describing the study types.

Table 1 shows the Medline database search strategy.

Table 1. Literature search strategy in the Medline database (PubMed interface)

<table>
<thead>
<tr>
<th>Study type/subject</th>
<th>Period</th>
<th>Number of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methodological choices for the assessment of medical devices</td>
<td>01/2013 – 10/2019</td>
<td>6</td>
</tr>
<tr>
<td>Recommendations</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td>–6</td>
</tr>
</tbody>
</table>

((Bayes Theorem OR Research Design ! OR Models, Statistical !)/de maj OR (bayesian OR Bayes OR study design OR methodological OR statistic* OR strateg*)/ti OR premarket/ti,ab) AND ((Clinical Trials as Topic ! OR Comparative Effectiveness Research OR Meta-Analysis as Topic ! OR Practice Guidelines as Topic OR Device Approval ! OR...
Product Surveillance, Postmarketing ! OR Reproducibility of Results ! OR Sample Size)/de OR premarket/ti,ab) AND ((Equipment Safety OR Equipment and Supplies ! OR Biomedical Technology !)/de OR device*ti)

AND

Step 2
(recommendation* OR guideline* OR statement* OR consensus OR position paper)/ti OR (Health Planning Guidelines)/de OR (Practice Guideline OR Guideline OR Consensus Development Conference OR Consensus Development Conference, NIH)/pt

<table>
<thead>
<tr>
<th>Meta-analyses, systematic reviews</th>
<th>01/2013</th>
<th>73</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/2019</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 1

AND

Step 3
(metaanalys* OR meta-analys* OR meta analysis OR systematic review* OR systematic overview* OR systematic literature review* OR systematical review* OR systematical overview* OR systematical literature review* OR systematic literature search OR pooled analysis)/ti OR (Meta-Analysis OR Systematic Review)/pt OR Cochrane Database Syst Rev/so

<table>
<thead>
<tr>
<th>Randomised controlled trials</th>
<th>01/2013</th>
<th>46</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/2019</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 1

AND

Step 4
(random*/ti OR (Random Allocation OR Double-Blind Method OR Single-Blind Method OR Cross-Over Studies)/de OR Randomized Controlled Trial /pt

<table>
<thead>
<tr>
<th>Controlled trials</th>
<th>01/2013</th>
<th>67</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/2019</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 1

AND

Step 5
(random*/tia,b OR (Random Allocation OR Double-Blind Method OR Single-Blind Method OR Cross-Over Studies)/de OR (Randomized Controlled Trial ! OR Controlled Clinical Trial ! OR Multicentre Study)/pt

<table>
<thead>
<tr>
<th>Other types of studies</th>
<th>01/2013</th>
<th>116</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/2019</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
EXCEPT Step 2 OR Step 3 OR Step 4 OR Step 5

**Medical devices and assessment of biomedical technologies**

<table>
<thead>
<tr>
<th>All study types</th>
<th>01/2013</th>
<th>10/2019</th>
<th>45</th>
</tr>
</thead>
</table>

**Step 6**

(technology assessment/ti) AND (Equipment and Supplies !/de maj OR medical device*/ti)

**Assessment of devices and medico-administrative databases**

<table>
<thead>
<tr>
<th>Recommendations</th>
<th>01/2013</th>
<th>10/2019</th>
<th>2</th>
</tr>
</thead>
</table>

**Step 7**

(Equipment Safety OR Equipment and Supplies ! OR Biomedical Technology !)/de maj OR device*/ti

AND

(Clinical Trials as Topic! OR Meta-Analysis as Topic! OR Guidelines as Topic! OR Health Services Research ! OR Research Design! OR Process Assessment, Health Care OR Device Approval ! OR Product Surveillance, Postmarketing ! OR Probability !)/de maj OR (pre-market OR postmarket OR design OR designs)/ti OR propensity score/ti,ab

AND

((Registries ! OR Records ! OR Databases as Topic !)/de maj OR (database* OR register OR registry OR registries OR health record* OR medical record*)/ti OR (national database* OR administrative database* OR public regulatory database*)/ti, ab

AND

**Step 2**

Meta-analyses, systematic reviews

<table>
<thead>
<tr>
<th>01/2013</th>
<th>10/2019</th>
<th>9</th>
</tr>
</thead>
</table>

**Step 7 AND Step 3**

Randomised controlled trials

<table>
<thead>
<tr>
<th>01/2013</th>
<th>10/2019</th>
<th>16</th>
</tr>
</thead>
</table>

**Step 7 AND Step 4**

Controlled trials

<table>
<thead>
<tr>
<th>01/2013</th>
<th>10/2019</th>
<th>90</th>
</tr>
</thead>
</table>

**Step 7 AND Step 5**

Observational studies

| 01/2013 | 10/2019 | 46 |
### Step 7

**AND**

### Step 8

(cohort\* OR longitudinal stud\* OR follow-up stud\* OR prospective stud\* OR retrospective stud\*)/ti OR (Cohort Studies OR Longitudinal Studies OR Follow-Up Studies OR Prospective Studies OR Retrospective Studies)/de OR Observational Study/pt

<table>
<thead>
<tr>
<th>Other types of studies</th>
<th>01/2013</th>
<th>10/2019</th>
<th>-74</th>
</tr>
</thead>
</table>

### Step 7

**EXCEPT** Step 2 OR Step 3 OR Step 4 OR Step 5 OR Step 8

### Medical devices and *in silico* methods

<table>
<thead>
<tr>
<th>All study types</th>
<th>01/2013</th>
<th>10/2019</th>
<th>-66</th>
</tr>
</thead>
</table>

### Step 9

(Computer Simulation !/de maj OR (computer simulation\* OR computerized model\* OR computer model\* OR in silico*/ti,ab OR in silico*/ot)

AND

((Equipment Safety OR Equipment and Supplies ! OR Biomedical Technology!)/de OR device*/ti,ab)

AND

(Clinical Trials as Topic !/de OR Clinical Trial !/pt OR (clinical trial\* OR clinical stud*)/ti)

### Single-case experimental design: methodology

<table>
<thead>
<tr>
<th>All study types</th>
<th>01/2009</th>
<th>06/2019</th>
<th>-104</th>
</tr>
</thead>
</table>

### Step 10

(single-case experimental design\* OR single-case design\* OR SCED OR N-of-1\* OR small population group trial\* OR single subject experimental design OR single subject research design\* OR small N design\* OR multiple-case design\* OR single-case design\* OR single-systems design\*)/ti,ab OR (single-case experimental design\* OR single-case design\* OR SCED OR N-of-1\* OR small population group trial\* OR single subject experimental design OR single subject research design\* OR small N design\* OR multiple-case design\* OR single-case design\* OR single-systems design\*)/ot OR (small population*/ti AND clinical trial*/ti)

**AND**

### Step 11

(assess\* OR analysis\* OR effectiveness OR methods OR implementation OR guide OR bias OR evidence-based)/ti OR
method/ti,ab OR (Research Design ! OR Data Interpretation, Statistical)/de maj OR (Bias ! OR Evidence-Based Practice !/methods)/de

<table>
<thead>
<tr>
<th>Single-case experimental design and medical devices</th>
<th>All study types</th>
<th>01/1999</th>
<th>06/2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 12 Equipment and Supplies !/de OR medical device*/ti,ab AND Step 10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-case experimental design and physical medicine and rehabilitation</td>
<td>All study types</td>
<td>01/1999</td>
<td>06/2019</td>
</tr>
<tr>
<td>Step 13 Physical and Rehabilitation Medicine !/de OR (neurorehabilitation OR rehabilitation)/ti AND Step 10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

de: descriptor; de maj: major descriptor; *: truncation; ti: title; ab: abstract; pt: publication type; !: explode of generic term; so: journal title; ot: authors’ key words

**Sites consulted**

Latest consultation: October 2019

*Bibliothèque médicale Lemanissier*
*Catalogue et index des sites médicaux francophones – CISMeF*
*Comité d’évaluation et de diffusion des innovations technologiques – CEDIT*

Adelaide Health Technology Assessment – AHTA
Agency for Care Effectiveness
Agency for Healthcare Research and Quality – AHRQ
Alberta Health – HTA provincial reviews
Alberta Medical Association
Allied Health Evidence
American College of Physicians – ACP
Australian Clinical Practice Guideline
Australia and New Zealand Horizon Scanning Network
BMJ Best Practice
British Columbia guidelines
California Technology Assessment Forum
Campbell Collaboration
Canadian Agency for Drugs and Technologies in Health – CADTH
*Canadian Task Force on Preventive Health Care*
*Centers for Disease Control and Prevention – CDC*
Belgian Health Care Knowledge Centre – KCE
Centre for Clinical Effectiveness – CCE
*Centre for Effective Practice*
*Centre for Reviews and Dissemination databases*
CMA Infobase
Cochrane Library
Guidelines International Network – GIN
*Health Services Technology Assessment Text – HSTAT*
Institute for Clinical and Economic Review – ICER
Institute for Clinical Evaluative Sciences – ICES
*Institute for Clinical Systems Improvement – ICSI*
*Institute for Health Economics Alberta – IHE*
Institut national d’excellence en santé et en services sociaux – INESSS
International Network of Agencies for Health Technology Assessment – INAHTA
McGill University Health Centre
Malaysian Health Technology Assessment Section
Medical Services Advisory Committee – MSAC
*National Coordinating Centre for Health Technology Assessment – NCCHTA*
National Health and Medical Research Council – NHMRC
*National Health Services Evidence*
National Health Services Innovation Observatory
National Institute for Health and Clinical Excellence – NICE
New South Wales Agency for Clinical Innovation
New Zealand Guidelines Group – NZGG
Ontario Health Technology Advisory Committee – OHTAC
Public Health Agency of Canada
Scottish Health Technologies Group
Scottish Intercollegiate Guidelines Network – SIGN
*Singapore Ministry of Health*
*Tripdatabase*
U.S. Food and Drug Administration
U.S. Preventive Services Task Force
Veterans affairs, Dep. Of Defense Clinical practice guidelines
Veterans Affairs Evidence-based Synthesis Program

**Monitoring**

In addition, monitoring was carried out until the end of February 2021 in Medline, on the basis of the equations in table 1 and on the websites indicated above.
Annexe 3. List of figures and tables

Table of figures

Figure 1. Decision tree for limiting measurement bias in randomised trials 18
Figure 2. Pre-randomisation according to Zelen, case 1 (45) 19
Figure 3. Pre-randomisation according to Zelen, case 2 (45) 20
Figure 4. Double consent according to Zelen, case 3 (47) 20
Figure 5. Diagram of randomisation in expertise-based randomised controlled trials, as per (39) 23
Figure 6. Diagram of cluster randomised trials, as per (61) 27
Figure 7. The crossover principle 30
Figure 8. Specific scenario of a crossover trial where two treatments are applied simultaneously 31
Figure 9. Choice of SCED, as per (72) 33
Figure 10. Graph-based decision-making rules for the triangular test (closed design), as per (107, 109) 39
Figure 11. Graph-based decision-making rules for the sequential probability ratio test (open design) 39
Figure 12. Why implement a real-world study?, as per the HAS guide on “Real-world studies for the assessment of medicinal products and medical devices” 45
Figure 13. Example of a prospective observational study comparing the occurrence of postoperative infections depending on the type of suture used 46
Figure 14. Principle of studies based on patient preference or comprehensive cohort studies, as per (163) 49
Figure 15. Instrumental variable, as per (187) 56

Table of tables

Table 1. Examples of cases where a blind study is impossible, as per (29) 13
Table 2. Examples of placebos used in surgery and for medical devices, as per (22) and (29) 14


34. Cohen PJ. Failure to Conduct a Placebo-Controlled Trial May Be Unethical. The American Journal of Bioethics 2002;2(2).


considerations for cluster-randomized trials (RoB 2 CRT); 2020.
https://drive.google.com/file/d/1J4okZ1zMZwZf2LS8Y_3XYghAZfUsel/view


http://dx.doi.org/10.1056/NEJMra1510059

https://has-sante.fr/upload/docs/application/pdf/eval_interventions_ameliorer_pratiques_guide.pdf


http://dx.doi.org/10.1016/j.annder.2013.07.024

https://drive.google.com/file/d/18Ek-Yw8HYQsULa8Ldexp1yOhaFk0EMgF/view


http://dx.doi.org/10.1016/s0140-6736(09)61459-8


http://dx.doi.org/10.1097/jpt.0000000000000187

http://dx.doi.org/10.1080/09602011.2013.824383

http://dx.doi.org/10.1177/0954411917702931


http://dx.doi.org/10.1056/NEJMra1512592

80. Girard P, Cucherat M, Guez D, les participants de la Table Ronde No 2 de Giens XIX. Simulation des essais cliniques dans le


92. Haute Autorité de Santé. Avis de la CNEDIMTS du 2 février 2021. COREVALVE EVOLUT R avec système de pose EnVeo PRO, bioprothèses valvulaires aortiques implantées par voie artérielle transcutanée. Saint-Denis La Plaine: HAS. https://has-sante.fr/upload/docs/evamed/CNEDIMTS-6421_COREVALVE_EVOLUT_R_02_1%C3%A8vrier_2021_(6421)_avis.pdf


### Abbreviations and acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANSM</td>
<td>Agence Nationale de Sécurité du Médicament et des Produits de Santé (French National Agency for Medicines and Health Products)</td>
</tr>
<tr>
<td>CNAAMTS</td>
<td>Caisse nationale de l’assurance maladie des travailleurs salariés (National health insurance fund for salaried workers)</td>
</tr>
<tr>
<td>CNEDIMTS</td>
<td>Commission nationale d’évaluation des dispositifs médicaux et des technologies de santé (Medical Device and Health Technology Evaluation Committee)</td>
</tr>
<tr>
<td>CSS</td>
<td>Code de la sécurité sociale (French Social Security Code)</td>
</tr>
<tr>
<td>DGOS</td>
<td>Direction générale de l’offre de soins (Directorate General of Healthcare Provision)</td>
</tr>
<tr>
<td>DGS</td>
<td>Direction générale de la santé (Ministry of Health)</td>
</tr>
<tr>
<td>MD</td>
<td>Medical Device</td>
</tr>
<tr>
<td>CMD</td>
<td>Connected Medical Device</td>
</tr>
<tr>
<td>DSS</td>
<td>Direction de la sécurité sociale (French social security division)</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised controlled trial</td>
</tr>
<tr>
<td>EMA</td>
<td>European Medicines Agency</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>FEDEPSAD</td>
<td>Fédération des prestataires de santé à domicile (Federation of home healthcare providers)</td>
</tr>
<tr>
<td>HAS</td>
<td>Haute Autorité de santé (French National Authority for Health)</td>
</tr>
<tr>
<td>HTA</td>
<td>Health Technology Assessment</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>INAHTA</td>
<td>International Network of Agencies for Health Technology Assessment</td>
</tr>
<tr>
<td>INCa</td>
<td>Institut national du cancer</td>
</tr>
<tr>
<td>ITT</td>
<td>Intention to treat</td>
</tr>
<tr>
<td>LPPR</td>
<td>Liste des produits et prestations remboursables (List of products and services qualifying for reimbursement)</td>
</tr>
<tr>
<td>MOST</td>
<td>Multiphase Optimisation Strategy</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Care Excellence</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>PP</td>
<td>Per protocol</td>
</tr>
<tr>
<td>PREMs</td>
<td>Patient Reported Experience Measures</td>
</tr>
<tr>
<td>PROMs</td>
<td>Patient Reported Outcome Measures</td>
</tr>
<tr>
<td>SED</td>
<td>Service évaluation des dispositifs (HAS medical device assessment department)</td>
</tr>
<tr>
<td>SMART</td>
<td>Sequential Multiple Assignment Randomised Trials</td>
</tr>
<tr>
<td>SNDS</td>
<td>Système national des données de santé (French National health data system)</td>
</tr>
<tr>
<td>YHEC</td>
<td>York Health Economic Consortium</td>
</tr>
</tbody>
</table>
Glossary

**Big data**: corresponds to very large databases (with in the region of 100 million lines and columns, or more).

**External validity of a study**: the possibility of generalising from the results of a study in order to apply them to the general population.

**Internal validity of a study**: indicates that the results obtained are correct for the population studied; the methodology and the existence of bias impact the internal validity.

**Types of bias**

– Selection bias occurs when the two groups in the trial are not comparable. A difference may then appear between the two groups independent of any treatment effect. Randomisation aims to avoid selection bias by creating, on average, two comparable groups.

– Confounding bias is the bias that results from failing to take confounding factors into consideration.

– Follow-up (or performance or execution) bias occurs when the two groups are not followed up in the same way during the trial and the care received is different. The initial comparability is then lost and a difference may appear, independent of any treatment effect.

– Attrition (or exclusion) bias occurs when “drop-outs” from the study (patients lost to follow-up and missing data) are different in the two treatment groups.

– Assessment (or measurement or detection) bias occurs when the endpoint is not measured in the same way in both groups. Double-blind trials limit the risk of assessment bias.

– Bias related to intention-to-treat analysis occurs when patients are not analysed in their initial randomisation group. Secondary exclusions are likely to bias the result, mainly by damaging the initial comparability between groups, especially if exclusions may be related to the effect of treatment.

**Alpha and beta risks**: the risks of concluding wrongly as a result of random fluctuations, which may lead the observer to an incorrect conclusion. The alpha risk is the risk of concluding there is a difference where none exists; the beta risk is the risk of wrongly concluding there is no difference where a difference exists.

**Confidence interval (CI)**: generally set at 95%, this is the range of values with a 95% chance of containing the true value for the parameter estimated; this interval allows the uncertainty of the estimate to be seen.

**Credible interval**: in the Bayesian approach, this interval, deduced from the posterior distribution, indicates the confidence one may have in the value for the parameter concerned, i.e. the 95% probability that its true value falls within the boundaries of the interval.

**Intention-to-treat analysis (ITT)**: involves analysing patients in their randomisation group, whatever treatment they received.
**Per-protocol analysis (PP):** consists in analysing patients according to the treatment they received rather than their randomisation group.

**Confounding factor:** factor associated with the treatment which may also influence the result; it could weaken or strengthen an association between exposure and the results observed.

**Non-random censoring:** missing data caused by a patient lost to follow-up (deviation from the protocol not due to chance, but related to the treatment administered).

**Propensity score:** propensity to receive a treatment based on the patient’s characteristics.