



HAUTE AUTORITÉ DE SANTÉ

METHODOLOGY GUIDE

**Methodological Choices for the Clinical
Development of Medical Devices**

Assessment Report

4 October 2013

The scientific evidence for this assessment will be available for download at
www.has-sante.fr

Haute Autorité de Santé
Documentation and Public Information Department
2, avenue du Stade de France - F 93218 Saint-Denis La Plaine CEDEX
Tel.: +33 (0)1 55 93 70 00 - Fax: +33 (0)1 55 93 74 00

Foreword

In view of the shortcomings frequently observed in assessments of the efficacy of non-pharmacological treatments, the National Committee for the Assessment of Medical Devices and Health Technologies (CNEDiMTS) wished to identify a set of methods and conditions that will allow high-quality clinical assessment, particularly when conventional randomised controlled trials cannot be performed.

The methods described in the following document do however have their limitations and should be reserved for unique situations where it is considered impossible to conduct a conventional randomised controlled trial. Use of these alternative methods should be scientifically backed up and justified. Randomised controlled trials remain the standard for any clinical study that compares the efficacy of multiple treatments or treatment strategies.

The CNEDiMTS has produced this document for manufacturers, research organisations and project developers. It aims to provide an up-to-date overview of comparative methods that can be used to evaluate the potential clinical benefit of a new medical device or health technology, and to describe possible research designs. More specifically, the objectives of this document are as follows:

- to present all methods potentially available when randomisation and/or blinding are impossible to implement
- to identify the limitations of these methods
- to indicate, wherever possible, specific situations in which these methods can be used

Contents

Contributors.....	6
List of abbreviations	7
1. Introduction	8
2. Objective	8
3. Caveat	8
4. Background	9
4.1. <i>Special features of assessing medical devices and health technologies</i>	9
4.2. <i>Phases of clinical development</i>	10
5. Methods used	11
5.1. <i>Bibliographic data</i>	11
5.1.1. <i>Literature search</i>	11
5.1.2. <i>Search results</i>	11
6. Methodological problems.....	16
6.1. <i>Timing the assessment.....</i>	16
6.2. <i>Eligible population and recruitment</i>	16
6.3. <i>Acceptability</i>	16
6.4. <i>Randomisation.....</i>	17
6.5. <i>Blinding.....</i>	18
6.6. <i>Choice of control or comparator group</i>	19
6.7. <i>Factors related to operator experience.....</i>	21
6.7.1. <i>Learning curve</i>	21
6.7.2. <i>Volume of activity.....</i>	21
6.8. <i>Adjustment factors.....</i>	22
6.9. <i>Type of analysis.....</i>	22
6.9.1. <i>Intention-to-treat analysis</i>	22
6.9.2. <i>Per-protocol analysis</i>	22
7. Methods used to overcome the problems identified	23
7.1. <i>Achieving blinding or compensating for the lack of blinding</i>	23
7.2. <i>Other types of randomised controlled trials</i>	25
7.2.1. <i>Experimental designs</i>	25
7.2.1.1. <i>Zelen’s design or randomised consent design trial</i>	25
7.2.1.2. <i>Expertise-based randomised controlled trials.....</i>	28
7.2.1.3. <i>Tracker trial designs</i>	33
7.2.1.4. <i>Cluster randomised trials</i>	34
7.2.1.5. <i>Crossover trials.....</i>	37
7.2.2. <i>Interim analyses.....</i>	39
7.2.2.1. <i>Sequential trials</i>	39
7.2.2.2. <i>Adaptive randomisation trials</i>	44
7.2.2.3. <i>Bayesian methods</i>	45
7.3. <i>Comparative non-randomised observational studies.....</i>	48
7.3.1. <i>Type of study</i>	48
7.3.1.1. <i>Comprehensive cohort studies or trials based on patient preference</i>	48
7.3.1.2. <i>Prospective comparative observational studies</i>	49
7.3.2. <i>Analysis strategies.....</i>	51
7.3.2.1. <i>Propensity score.....</i>	51
7.3.2.2. <i>Instrumental variables</i>	53

8. In practice55
Appendix 1: Glossary 62
Appendix 2: Summary tables..... 64

Contributors

This document was produced by Isabelle Fournel (Project Manager, Registrar and Lecturer, Dijon, isabelle.fournel@chu-dijon.fr) and Michel Vaneau (Project Manager, Medical Devices Assessment Department, tel.: 01 55 93 37 56, email: m.vaneau@has-sante.fr).

The research and document management were undertaken by Aurélien Dancoisne (Researcher, Documentation and Public Information Department, tel.: 01 55 93 73 39, email: a.dancoisne@has-sante.fr) and Laurence Frigère (Assistant Researcher, Documentation and Public Information Department, tel.: 01 53 93 73 32, email: l.frigere@has.sante.fr).

Meetings were organised and secretarial support provided by Sandrine Bouvet and Fadela Chebili (tel.: 01 55 93 37 43, email: s.bouvet@has-sante.fr and f.chebili@has-sante.fr).

Senior Managers

Hubert Galmiche (Assistant Head of Medical Devices Assessment Department, tel.: 01 55 93 37 48, email: h.galmiche@has-sante.fr).

Catherine Denis (Head of Medical Devices Assessment Department, tel.: 01 55 93 37 40, email: c.denis@has-sante.fr).

Frédérique Pagès (Head of Documentation and Public Information Department, tel.: 01 55 93 73 23, email: f.pages@has-sante.fr).

Working Group

The working group consisted of the following professionals:

- Alain Bernard
- Michel Cucherat
- Isabelle Fournel
- Hubert Galmiche
- Pascal Giraux
- Bernard Guillot
- Jacques Machecourt
- Patrick Maison

Proofreading Group

The proofreading group consisted of the following professionals:

- Michèle Morin-Surocca
- Patrice Nony
- Françoise Roudot-Thoraval
- Sophie Stamenkovic

List of abbreviations

CNEDiMETS	National Committee for the Assessment of Medical Devices and Health Technologies
MD	Medical Device
HAS	Haute Autorité de Santé
FDA	Food and Drug Administration
CI	Confidence interval
ITT	Intention to treat
LPPR	List of products and services qualifying for reimbursement
PP	Per protocol
SED	Medical Devices Assessment Department

1. Introduction

Any efficacy assessment for a new treatment should be based on high-quality clinical trials. Only methodologically sound clinical trials allow judgements to be made about any causal links between the studied treatment and the observed effect and formal conclusions to be reached about the effect of treatment.

Since a clinical assessment has the aim of comparing the efficacy of two treatments, a well-conducted double-blind randomised controlled trial is the type of study that offers the highest level of evidence, enabling medical decisions to be backed up (1) and the efficacy of these treatments to be evaluated (2).

Randomisation, blinding and intention-to-treat analysis¹ are the main criteria with regard to methodological quality.

In most cases, a randomised controlled trial can be performed. However, in some very specific situations that require justification, conducting randomised controlled trials on medical devices (MDs) is known to be difficult for reasons relating to:

- the inability to randomise
- the choice of comparator
- the difficulties of implementing a double-blind procedure
- non-acceptance by the patient or clinician
- subjective endpoints in certain situations.

Other types of trials and observational studies should be reserved for situations where a conventional randomised controlled trial is not possible; the decision to use such a trial should be reasoned and justified.

2. Objective

This document focuses on aspects of the clinical efficacy assessment for a new medical device or a new health technology from development onwards, following feasibility studies. It aims to identify the methods and conditions that allow a high-quality clinical assessment of an MD to be made. It is intended for manufacturers, research organisations and project developers.

More specifically, the objectives of this document are as follows :

- to present all methods potentially available when randomisation and/or blinding are impossible to implement
- to identify the limitations of these methods
- to indicate, wherever possible, specific situations in which these methods can be used

3. Caveat

Whenever a methodology other than a randomised controlled trial is chosen, it must be scientifically justified.

¹ For studies aiming to demonstrate the superiority of a treatment

The clinical assessment methods presented in this document are the ones that should be considered when the respective efficacies of two treatments, or of one treatment and a placebo, are to be compared. These methods do not, however, guarantee the clinical relevance of the endpoints used for this purpose. In addition, these methods can only be applied to data of sufficient quality. Therefore, the studies conducted must be subject to a set of measures intended to ensure that this is achieved (a precise protocol entered in a register and/or published before the start of the trial, the selection and training of investigators, an appropriate data collection method, monitoring, quality control, etc.).

4. Background

4.1. *Special features of assessing medical devices and health technologies*

Non-pharmacological treatments may be therapeutic or diagnostic in nature or compensate for a disability, and consist of surgical techniques, interventional therapies (such as angioplasty) and the use of medical devices (for example, vascular stents, implantable cardiac pacemakers, ultrasound, electrical stimulation and orthotics).

The assessment of these treatments is considered to be important, but it is often insufficient (3). The lack of high-quality data is a significant obstacle for assessment agencies and decision makers.

In the literature, the life cycle of a technology has been described in five stages: future (not yet developed), emerging (just before introduction to clinical practice), new (in the adoption phase), accepted (in general use) and obsolete (should no longer be used) (3).

Others have defined emerging technologies as those not yet adopted by the healthcare system, and new technologies as those in the adoption phase which have only been available for clinical use for a short while (4).

Technologies need to be assessed before their widespread distribution, namely in the emerging phase (3). This requires information to be available about the clinical efficacy and safety of the technology. Indeed, proof of clinical benefit is essential from an ethical point of view.

The ideal assessment process for a health technology consists of three main phases (3): a detection phase, an assessment phase and a monitoring phase.

The first detection or “horizon scanning” phase relies on a surveillance system that can identify new and emerging medical technologies.

The second phase consists of the in-depth assessment (safety and efficacy) of emerging technologies.

Finally, the last phase involves surveillance and regular re-assessment of the use of a technology in practice.

4.2. *Phases of clinical development*

The clinical assessment of a new treatment's efficacy, which is the subject of this document, takes place after the pre-clinical 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 pre-clinical phase. They allow the technique to be refined and appropriate efficacy endpoints to be determined. Depending on the context, one or more studies may be necessary to answer different questions. In most instances, the most appropriate type of methodology is a prospective non-comparative study.

The choice of an efficacy threshold is a crucial issue. Indeed, before planning studies that include a large number of patients, the project developer or manufacturer must verify that their new MD is promising in terms of efficacy. The first step therefore involves choosing an efficacy endpoint from data in the literature or through expert opinion, and determining the probabilities of efficacy and inefficacy. These probabilities will allow rules to be created for stopping or continuing the research, while including the fewest number of patients possible.

Feasibility studies are useful for (5):

- Selecting patients who will benefit from the new medical device

This step involves clarifying the clinical forms of the disease and defining patients likely to benefit from the MD.

- Refining the technique

For implantable MDs, the implantation technique must be further defined by describing the different surgical procedures, technical facilities and personnel required. This can only be done in the context of a clinical trial. Although the implantation technique may continue to be improved after this step, this should not result in delays setting up a clinical trial to demonstrate the clinical benefit.

- Measuring clinical efficacy

The feasibility study is a prerequisite for hypothesis generation and testing and for calculating the number of participants needed in a comparative study. At this stage of development, determining the primary clinical endpoint will thus allow the clinical efficacy of the MD to be measured. The endpoint must be clinically relevant, such as the reduction in mortality, or a clinically measurable complication. The use of an intermediate endpoint must have been justified and validated in previous studies.

- Complications and risks

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

These stages are indispensable and provide essential information for subsequent demonstrations of efficacy through randomised controlled trials.

5. Methods used

In accordance with the regulations, experts who have contributed to HAS publications have completed a public declaration of interest, the aim of which is to inform HAS of any conflicts of interest between any member of the group and a manufacturer. These declarations of interest have been published on the HAS website.

5.1. Bibliographic data

5.1.1. Literature search

A systematic literature search was conducted for the period January 2000 to July 2012 by searching medical bibliographic databases:

- Medline (National Library of Medicine, USA)
- The Cochrane Library (Wiley Interscience, USA)
- BDSP (French Public Health Database)
- National registers
- Websites that publish guidelines, technological assessment reports or economic assessment reports
- Websites of learned societies competent in the field studied
- Specialist sources, particularly epidemiology and economics sources

The search was limited to publications in English and French. Monitoring continued until 27 June 2013.

Technological assessments, guidelines, consensus conferences, meta-analyses, systematic reviews, randomised controlled trials and other controlled trials, comparative studies and cohort studies were sought.

The search strategy and a list of sources used are detailed in the following section. The search was completed by the bibliographies of experts consulted and by manufacturers' data.

5.1.2. Search results

The search strategy for the Medline bibliographic database was constructed by using either terms from the thesaurus (Medline MeSH descriptors) or free-text terms (from the title or abstract) for each subject. These were combined with terms describing the types of study.

Table 1 presents the Medline search strategy, indicating its successive steps and giving the results by:

- **Number of unique references identified: 1172**
- **Number of references analysed: 206**
- **Number of references retained: 96**

The results of Medline alerts are not included in this table.

Table 1: Medline database search strategy

Type of study / subject Terms used	Period searched	Number of references
Choices of methodology for the clinical development of medical devices – Broad search		
Guidelines	01/2005 – 08/2010	128
Step 1 (design[ti] OR designing[ti] OR method[ti] OR methods[ti] OR model[ti] OR modeling[ti] OR modelling[ti] OR statement[ti] OR reporting[ti] OR report[ti] OR process[ti] OR processing[ti] OR nonpharmacologic[ti] OR blinding[ti] OR blind[ti] OR checklist[ti] OR elaboration[ti])		
AND		
Step 2 ("Epidemiologic Research Design"[Mesh] OR "treatment outcome"[Mesh] OR "Publishing/standards"[Mesh] OR "Bias (Epidemiology)"[mesh] OR "Research Design/standards"[Mesh] OR delphi technique[mesh])		
AND		
Step 3 (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		
Meta-analyses and systematic reviews	01/2005 – 08/2010	256
Steps 1 to 2		
AND		
Step 4 (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)/ti OR Meta-Analysis/pt OR Cochrane Database Syst Rev/so		
Randomised controlled trials	01/2005 – 08/2010	219
Steps 1 to 2		
AND		
Step 5 random*/ti OR (Random Allocation OR Double-Blind Method OR Single-Blind Method OR Cross-Over Studies)/de OR Randomized Controlled Trial/pt		
Controlled trials	01/2005 – 08/2010	33
Steps 1 to 2		
AND		
Step 6 random*/ti,ab 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 Multicenter Study)/pt		
Cohort studies	01/2005 – 08/2010	18
Steps 1 to 2		
AND		
Step 7 (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		
Other studies	01/2005 – 08/2010	30
Steps 1 to 2		
WITHOUT		
Step 8 Without previous steps		

Choices of methodology for the clinical development of medical devices - Initial search

Guidelines		01/2005 – 08/2010	6
Step 1	(Bayes Theorem[majr] OR Research Design[majr] OR Models, Statistical[majr] OR Bayesian[ti] OR Bayes[ti] OR study design[ti] OR methodological[ti] OR statistic*[ti] OR strateg*[ti] OR Premarket[tiab]) AND (Clinical Trials as Topic[mesh] OR Comparative Effectiveness Research[mesh] OR Meta-Analysis as Topic[mesh] OR Practice Guidelines as Topic[mesh] OR Device Approval[mesh] OR Product Surveillance, Postmarketing[mesh] OR Reproducibility of Results[mesh] OR premarket[tiab] OR sample size[mesh]) AND (Equipment Safety[mesh] OR Equipment and Supplies[mesh] OR Technology assessment, biomedical [mesh] 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		
Meta-analyses and systematic reviews		01/2005 – 08/2010	67
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)/ti OR Meta-Analysis/pt OR Cochrane Database Syst Rev/so		
Randomised controlled trials		01/2005 – 08/2010	51
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		
Controlled trials		01/2005 – 08/2010	10
Step 1			
AND			
Step 5	random*/ti,ab 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 Multicenter Study)/pt		
Other studies		01/2005 – 08/2010	141
Step 1			
WITHOUT			
Step 8	Without previous steps		

Table legend

Mesh: Descriptor; *: wildcard; de: descriptor, ti: title; ab: abstract; pt: publication type; so: journal title

• Websites consulted

Last consultation: June 2013.

French-language information:

- Agence d'Évaluation des Technologies et des Modes d'Intervention en Santé [Agency for the Assessment of Health Technologies and Interventions], Canada
- Agence Nationale de Sécurité Sanitaire (Anses) [French Agency for Food, Environmental and Occupational Health and Safety], France

- Agence Nationale de Sécurité du Médicaments et des produits de santé (ANSM) [National Medicines and Health Products Safety Agency], France
- Association Française de Normalisation [French Standardisation Association] - AFNOR
- Bibliothèque Médicale AF Lemanissier [AF Lemanissier Medical Library], France
- Catalogue et Index des Sites Médicaux Francophones [Catalogue and Index of French-Language Medical Websites], France
- Centre fédéral d'expertise des soins de santé [Federal Centre for Healthcare Expertise], Belgium
- Direction de la recherche, des études, de l'évaluation et des statistiques [Directorate for Research, Surveys, Assessment and Statistics], France
- Evaluation des technologies de santé pour l'aide à la décision (Fédération hospitalière de France) – ETSAD [Evaluation of Health Technologies for Decision-Makers (French Hospitals Federation)], France
- Expertise collective de l'INSERM [INSERM Collective Expertise], France
- Haute Autorité de Santé, France
- Institut national de prévention et d'éducation pour la santé (INPES) [National Prevention and Health Education Institute], France
- Institut de recherche et documentation en économie de la santé (IRDES) [Institute for Research and Information in Health Economics], France
- Institut de la statistique et des études économiques (INSEE) [National Institute of Statistics and Economic Studies], France
- Institut de veille sanitaire (InVS) [French Institute for Public Health Surveillance], France
- La Documentation française bookshop, France
- Portail de la statistique publique française [French Official Statistics Portal], France
- Société Française de Médecine Générale [French General Medical Society], France
- Unions Régionales des Caisses d'Assurance Maladie [Regional Association of Health Insurance Funds], France

English-language information:

- Adelaide Health Technology Assessment, Australia
- Agency for Healthcare Research and Quality, USA
- Alberta Heritage Foundation for Medical Research, Canada
- Alberta Medical Association, Canada
- American College of Physicians, USA
- American Academy of Orthopaedic Surgeons
- American Heart Association, USA
- Australian Therapeutic Goods Administration
- Blue Cross Blue Shield Association, USA
- BMJ Clinical Evidence, UK
- Canadian Agency for Drugs and Technologies in Health, Canada
- Canadian Task Force on Preventive Health Care, Canada
- Centers for Disease Control and Prevention Infection Control Guidelines, USA
- Centre for Clinical Effectiveness, Australia
- Centre for Reviews and Dissemination, UK
- CMA Infobase, Canada
- CONSORT Group
- Guidelines and Protocols Advisory Committee, Canada
- Guidelines International Network
- Institute for Clinical Systems Improvement, USA
- Minnesota Department of Health – Health Technology Advisory Committee, USA
- National Coordinating Centre for Health Technology Assessment, UK
- National Guidelines Clearinghouse, USA
- National Health Services Scotland, UK
- National Institute for Health and Clinical Excellence, UK
- National Institute for Health Research Horizon Scanning Centre, UK

- National Institutes of Health, USA
- National Library of Guidelines Specialist Library, UK
- New Zealand Guidelines Group, New Zealand
- New Zealand Health Technology Assessment, New Zealand
- Ontario Medical Advisory Secretariat, Canada
- Regional Assessment Panel, UK
- Scottish Intercollegiate Guidelines Network, UK
- Singapore Ministry of Health, Singapore
- U.S. Food and Drug Administration
- U.S. Preventive Services Task Force, USA
- Veterans Affairs Technology Assessment Program, USA

- **Monitoring**

Medline was monitored up until 27 June 2013 using alerts based on the equations in Table 1. This brought 213 additional references to light.

The contents pages of the following publications were examined throughout the project: British Medical Journal (BMJ), Journal of the American Medical Association (JAMA), The Lancet, The New England Journal of Medicine; the medical and paramedical daily press and Agence Presse Médicale (APM).

6. Methodological problems

Some methodological principles that are intrinsic to randomised trials of pharmacological treatments may be more difficult to apply when assessing medical devices and health technologies.

Problems related to conducting double-blind randomised controlled trials of medical devices are detailed below.

6.1. *Timing the assessment*

Choosing the most appropriate time in its life cycle to clinically assess an MD is one issue to consider. MDs usually undergo changes after they have been launched, which aim to improve them. Therefore, a study carried out too early may not reflect the true performance of the medical device if it does not sufficiently take into account the period needed to learn the technique. In this case, an unfavourable assessment may reflect a poorly mastered technique rather than a genuinely ineffective technique (6).

On the other hand, an assessment conducted too late is responsible for MDs or health technologies being used without any proof of efficacy. An assessment should take place before they are widely distributed. In fact, once an MD or technology is widely distributed, it is difficult to get doctors to adhere to a study protocol (3), because a technique already used is often empirically considered to be effective.

Over time, professionals will change the situations in which the MD is used. These developments may invalidate the initial assessment (3).

6.2. *Eligible population and recruitment*

The small size of the eligible population is also a particular feature of studies of medical devices. Indeed, the target population may be far less sizeable for medical devices than for most drug treatments (3), possibly only involving a few hundred patients in some cases (7). In this situation, a conventional parallel-group trial may be more difficult to implement due to its complexity and cost.

The selection of the population studied is important (8). 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 delineate the population most likely to benefit from the new treatment.

6.3. *Acceptability*

The acceptability of the study to patients plays an important role when assessing treatments.

Obtaining patient consent is a prerequisite for conducting a clinical trial. When patients are informed before giving their consent, they should be provided with clear, documented and reliable information. If patients will not consent, the feasibility of the study is called into question. Where there are reasons to believe that the risk/benefit ratio differs between the treatments, both patients and surgeons may prefer a specific intervention and refuse to take

part in the trial. This is particularly true of new surgical methods that may be used in emergency situations or paediatrics.

Some patients prefer to choose their treatment, and refuse to be randomised (6). These issues can disrupt patient recruitment and make randomisation difficult. Whether or not use of the technique is widespread may also be a source of difficulty when convincing patients to take part in a clinical trial.

Questions of acceptability may additionally be raised by surgeons, if they are absolutely convinced that the technique they normally use is the best strategy (9).

6.4. Randomisation

The randomisation techniques considered to be adequate include the use of random number tables and computer generation of the randomisation group (10). It is essential that a centralised randomisation procedure is used. The use of envelopes, including opaque and sealed envelopes, does not guarantee that the treatment received will be unpredictable in open-label trials. For example, a doctor may open an envelope and, depending on the treatment offered, refuse to enrol his or her patient. These techniques 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 secret allocation in less than a quarter of cases, according to a literature review examining the publication of clinical trials in journals with a high impact factor (11).

Several factors could explain these shortcomings. On the one hand, a randomised trial is more expensive than a case series (12). On the other hand, randomisation may be judged to be impossible from the outset, most often for practical reasons (e.g. preference for a new potentially effective treatment) or more rarely for ethical reasons (3).

Among the ethical reasons cited is the comparison of an invasive treatment with a non-invasive treatment, which could lead to patients refusing to take part. Although it is often quoted in the literature, this argument is highly questionable. In fact, disseminating treatments without proof of their efficacy or without any risk/benefit assessment is also an unethical position.

Furthermore, in the majority of cases such trials are possible. The literature actually provides many examples of assessments of invasive versus non-invasive treatments, such as stem cell transplants in Parkinson's disease (13) or coronary angioplasty versus medical treatment (14).

6.5. Blinding

Blinding is an important element in clinical trials, because it can reduce classification or measurement bias related to the doctor's or patient's subjectivity. A crucial element of blinding is that it must be impossible to distinguish between the treatments compared. The patient, doctor administering the treatment, endpoint committee and/or statistician may be blinded.

A literature review published in 2003 showed that barely a quarter of patients, 6% of doctors and 2/3 of assessors were blinded in non-pharmacological trials (11). This is particularly harmful given that the doctor's influence is also more marked than in pharmacological trials (11). Open-label trials overestimate the therapeutic effect by 14% in comparison with double-blind trials (15).

Blinding is impossible more often in non-pharmacological studies (16), for ethical or practical reasons (3). In these situations, in order to evaluate the efficacy of a non-pharmacological treatment as objectively as possible, alternatives have been developed.

Boutron et al. have summarised the different blinding methods used in non-pharmacological trials published in journals with a high impact factor (17). Blinding may be complete, partial, or only apply to the assessment of endpoints.

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 intervention is simulated in the control group with no contact between the operator (most often a surgeon) and the team responsible for the patient's follow-up. Despite the practical difficulties with therapeutic procedures, some trials have been conducted with complete blinding, involving an identical placebo, a sham procedure, and no communication with the surgeons.

In some cases, it is impossible to blind these key individuals (Table 2).

The absence of blinding may lead to biases such as selection bias, execution bias, attrition bias or detection bias (see Appendix 1), calling the internal validity of a study into question (6).

Table 2: Examples of cases where a blind study is impossible, from (18)

Blinding patients	Blinding the operator	Blinding the follow-up team	Blinding the endpoint assessment
<ul style="list-style-type: none"> - Appearance or perception of device - Scar that reveals the type of intervention (e.g. surgery vs. non-invasive treatment, laparotomy vs. laparoscopy, etc.) 	<ul style="list-style-type: none"> - Surgical intervention - Appearance or manipulation of device 	<ul style="list-style-type: none"> - Specific adverse effects characteristic of one of the treatments administered - Dye or characteristic marks left by the device - Radiological appearance suggestive of one of the treatments received - Scar that reveals the type of intervention 	<ul style="list-style-type: none"> - Practical organisation: an excessive number of visits or consultations limiting acceptability to the patient - If consultation is carried out only by one external assessor: the surgeon is reduced to the role of a mere technician, which is an acceptability problem for the surgeon and also for the patient

6.6. *Choice of control or comparator group*

The choice of the control or comparator group is crucial in non-pharmacological trials. This problem does not apply to studies where the treatment evaluated is added to the standard treatment², which is used alone in the control group.

According to the 11th directive of the Council for International Organizations of Medical Sciences regarding biomedical research in human beings, the use of a placebo (or an inactive treatment) may be considered ethically 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 (19)

From an ethical point of view, it is difficult to offer patients an invasive sham procedure (20). In fact, the more invasive the procedure, the harder it is to justify exposing patients in the control group to risks that may be substantial without any expected benefit (21). An important counter-argument to this point of view is that ethical considerations also apply to treatments received by future patients, in that the widespread use of an unassessed treatment is not ethical. This suggests that it may be important to make participants aware of the overall benefit of a study.

In particular, it has been suggested that it is not ethical to administer a placebo in place of a standard treatment with demonstrated efficacy (19), or indeed an “invasive placebo” that does not help to strengthen the demonstration. Thus, in an assessment of stem cell grafts for Parkinson’s disease, the control group underwent a sham procedure under general anaesthetic with incision and abrasion of the external cortical bone in the cranium (22). The decision to use this placebo (or sham) surgery was widely controversial for several reasons. Firstly, it was of no benefit to patients whereas it did entail risks (relating to anaesthesia) (23), and secondly, alternatives did exist (24) such as performing deep brain stimulation (25).

A literature review has summarised the main surgical placebos used in non-pharmacological trials (17). For surgery and technological interventions, 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 is simulated by making an incision similar to that made in the treated group, or by injecting a placebo.

In practice, so-called placebo surgery is virtually impossible and limited to cases where there is no suitable comparator and where it involves little risk (6). It is also important to standardise preoperative 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.

Boutron et al. have also reported the different placebos possible when using medical devices: placebo prostheses, hidden medical devices, identical but inactive medical devices, active devices made ineffective, or use of similar equipment (17,26) (Table 3: Examples of placebos).

² Known as “add on” studies

Table 3: Examples of placebos used in surgery and for medical devices, from (17) and (18)

		Placebo
Surgical techniques	Intraoperative	<ul style="list-style-type: none"> - Patients under general anaesthetic - Patients masked with a drape - Skin incisions to obtain a similar scar to the study procedure (e.g. coronary artery bypass surgery versus skin incisions alone in patients with angina pectoris, arthroscopic surgery versus skin incisions alone in patients with osteoarthritis) - Sham procedure (e.g. intracerebral injection of stem cells versus skin incisions and abrasion of the external cortical bone of the cranium in patients with Parkinson's disease)
	Postoperative	<ul style="list-style-type: none"> - Scars concealed - Use of the same dressings, possibly impregnated with blood, on real or sham wounds to create an identical appearance in a study comparing mini-cholecystectomy with laparoscopic cholecystectomy, or similar dressings when comparing resection of the colon by laparotomy with laparoscopic colectomy - Standardisation of postoperative care / associated treatments
Medical devices	Placebo MD	<ul style="list-style-type: none"> - Placebo not strictly identical to the treatment received - Similar prosthesis which does not provide the therapeutic effect (e.g. without producing heat) - Use of different MDs concealed (e.g. transfusion bag for apheresis control, dark glasses for the participant, etc.)
	Identical appearance	<ul style="list-style-type: none"> - Use of an identical but inactive device (same duration, frequency, precautions, etc.) - Use of an identical, active machine: using a barrier to block treatment, changing the position of the source so there is no exposure to the treatment

6.7. Factors related to operator experience

6.7.1. Learning curve

A particular feature of health technologies using medical devices is that the operator's experience has an impact on the results of the technique (3). Different levels of experience may lead to different levels of performance when carrying out interventions. A lack of experience may influence the result of the study, penalising the new treatment tested (performance bias) (27).

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 skill are variability factors (6) and they should be taken into account, for example with a breaking-in period (28).

An assessment performed too early risks reflecting complications related to learning the new procedure. During surgical interventions, the impact of learning may be assessed through indicators such as the duration of the procedure or the volume of blood lost.

The variability of the treatment effect according to the operator's level of expertise should also be assessed. The importance of operator experience has been emphasised in trials comparing angioplasty and thrombolysis for the treatment of myocardial infarction (29,30).

Therefore, an assessment of a new technology versus a control risks being unbalanced in favour of the control treatment, because of the operator's experience (6). The study should therefore incorporate the effect of learning, for example by recording training and experience (6). From a pragmatic point of view, this learning phase must be taken into account in the trial so that any benefit provided by the device or health technology can be evaluated accurately.

6.7.2. Volume of activity

Volume of activity must also be taken into account when evaluating a new technology. Almost 70% of studies found a significant association between favourable clinical results and the doctor's volume of activity (31). This association is even greater for complex, rarely performed and high-risk surgical procedures (20).

6.8. Adjustment factors

It is sometimes difficult to separate the effects of non-pharmacological treatments from those of other aspects of the healthcare system. 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 (3). Thus, surgical interventions are not only dependent on the surgeon (preference, experience, etc.) but also on the other members of the team (anaesthetist, nurse), on pre- and postoperative care (6) and on the healthcare organisation.

As practices may vary in different centres, it is essential that these are defined very precisely in the protocol in advance, in order to avoid biased results relating to the type of practice. This standardisation must be extremely detailed, meticulously describing the procedure and the equipment required, which must be included in the specifications (8). However, in cases where these practices are not reproducible, the aim should be rather to collect data on any differences between centres, so that the analysis can be adjusted for these factors.

6.9. Type of analysis

6.9.1. Intention-to-treat analysis

Intention-to-treat analysis involves analysing patients in their randomisation group, whatever treatment they received. It should be used wherever possible in superiority trials because it is the more conservative approach. It prevents the benefits of randomisation from being lost and circumvents deviations from the protocol which may be due not to chance but to the treatment administered.

6.9.2. Per-protocol analysis

Per-protocol analysis involves analysing patients according to the treatment they received rather than their randomisation group. This strategy maximises any difference in effect between the approaches compared, and should therefore be used wherever possible in non-inferiority or equivalence trials.

7. Methods used to overcome the problems identified

7.1. *Achieving blinding or compensating for the lack of blinding*

When it is impossible to blind healthcare professionals, a blind assessment of the endpoint should be planned. 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 treatment 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 must meet the assessor
- if the treatment is identifiable to 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 treatment has been received

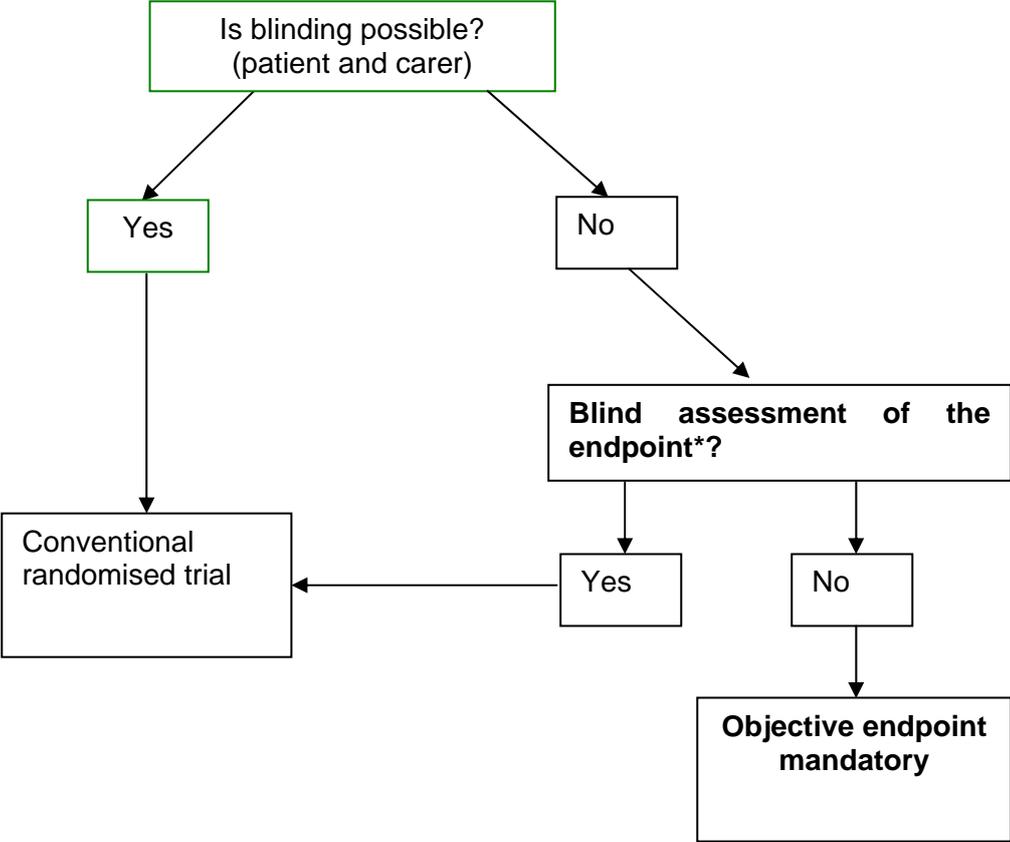
For surgical interventions, patient blinding can be achieved if the patient is under general anaesthetic or masked with a drape (17).

In some cases, an adjudication committee independent of the investigators is formed to check the endpoint. Blinding as to the study's hypotheses (or partial blinding) is also described as an alternative (17). 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 (e.g. the hypothesis that one treatment is superior), or of the randomisation outcome for patients under a modified Zelen's design. In these situations and for ethical purposes, patients must be advised that, for scientific reasons, 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 choose the most objective efficacy endpoint possible and for a blind assessment of this endpoint to be made (for example by an independent expert committee).

Figure 1 shows a decision tree for obtaining the best level of evidence as regards the endpoint.

Figure 1: Decision tree for limiting measurement bias in randomised trials



* or independent observer unaware of the treatment

7.2. Other types of randomised controlled trials

A trial with two parallel groups (or arms) should be considered before any other experimental design in any case where a comparative clinical assessment is needed. In such trials, the study treatment is compared with a control treatment by using two groups of patients formed by contemporaneous randomisation and followed up in parallel.

Because of the particular nature of medical devices, these conventional randomised controlled trials may be difficult to implement in practice.

Other types of experimental designs have been described, in particular for surgery, a field with similar difficulties to those encountered in medical device assessment.³

7.2.1. Experimental designs

7.2.1.1. Zelen's design or randomised consent design trial

- **Background**

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. In this way, 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 treatment rather than the standard treatment. In order to compensate for these inclusion issues, Zelen suggested only asking for patients' consent for the new treatment and not for the standard treatment (32).

- **Principle**

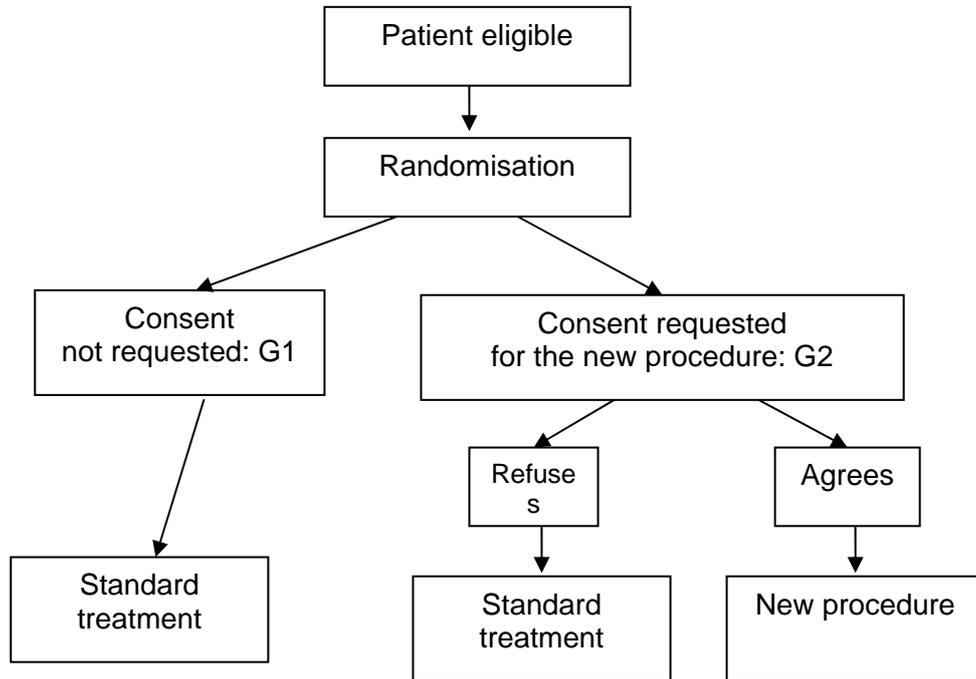
"Zelen's design" involves randomising patients without first obtaining their informed consent. Only patients randomised to the new treatment group must sign an informed consent form. Patients who refuse will be given the standard treatment (33).

Three options have been suggested by Zelen.

In the first (Figure 2), patients who are eligible are randomised to one of the two following groups: group G1 where consent is not requested, and group G2 where consent is requested (32). Patients in group G1 receive the standard treatment and are informed about their treatment. Patients in group G2 are asked to consent to the experimental treatment. If patients in group G2 give their consent, they will receive the new treatment. If they do not, they will receive the standard treatment (32). In the analysis, group G1 (standard treatment) must then be compared with group G2 (patients receiving either the new treatment or the standard treatment). The comparison must be made with all patients in group G2, whatever treatment they received. If only a small proportion of patients consent to receiving the experimental treatment, this design will be useless for evaluating the efficacy of the new treatment. However, a large proportion of new treatment refusals could also indicate that it is too early to use the experimental treatment in a clinical trial (32).

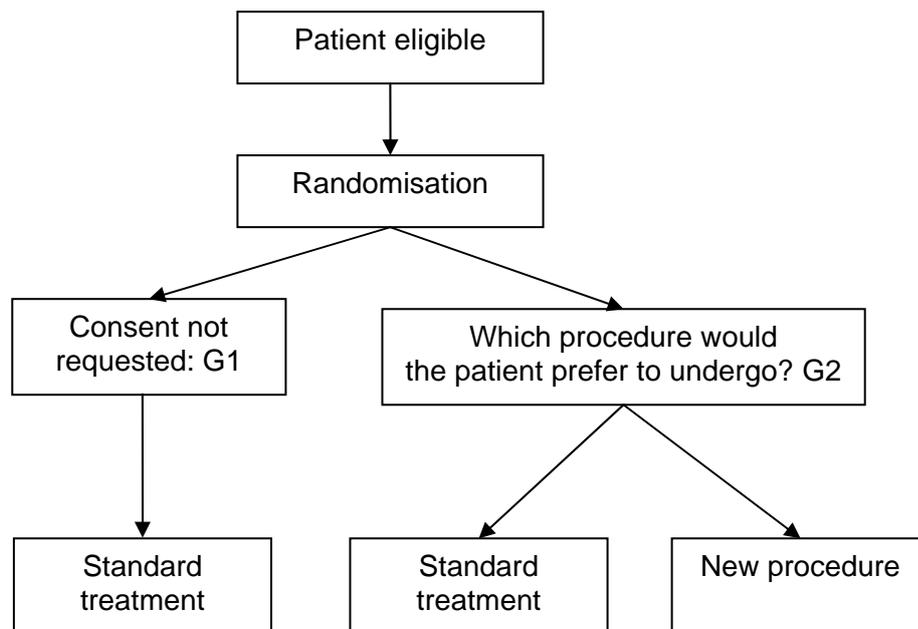
³ <http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf>

Figure 2: Pre-randomisation according to Zelen, case 1 (32)



In the second option (Figure 3), when patients are randomised to group G2, they are asked which treatment they would prefer to receive. This approach should be favoured in cases where treatment is particularly disabling (e.g. prostate cancer) (32).

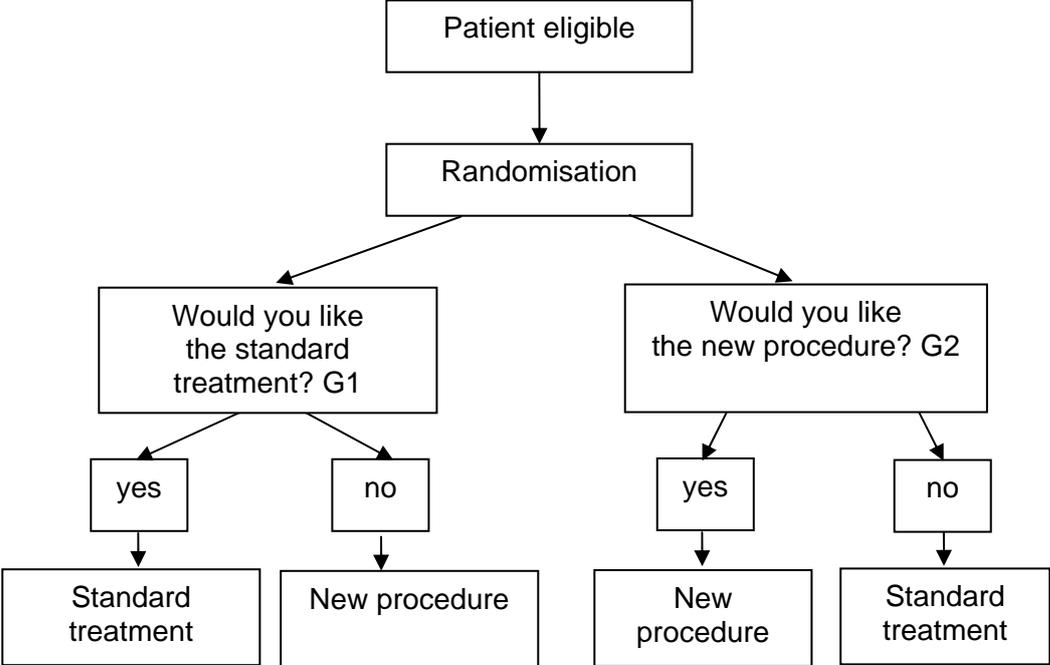
Figure 3: Pre-randomisation according to Zelen, case 2 (32)



Finally, the third option suggested by Zelen is also known as “double consent” (Figure 4). Patients are randomised to receive treatment A or treatment B. Those randomised to the

treatment A group are asked whether they would like to receive treatment A. If they refuse, they receive treatment B or another treatment. Those randomised to the treatment B group are asked whether they would like to receive treatment B. If they refuse, they receive treatment A or another treatment (34). This alternative is particularly useful when there is no established standard treatment.

Figure 4: Double consent according to Zelen, case 3 (34)



A comparison is made between the two groups, whatever the treatment ultimately received (34). This comparison looks at group G1 and at group G2 in their entirety, whatever the treatment received (32).

- **Advantages**

Patients know which treatment they will receive before they give their consent. This can be an advantage over conventional trials (apart from cases of double consent) where patients may withdraw their consent once they have been assigned a treatment. In addition, this type of study requires patients to make a decision only when they are to receive the experimental treatment (35). This can reduce stress in certain situations where a new technique is compared with the standard treatment. Furthermore, it is easier to recruit patients even when they have a strong preference for one treatment over another. This type of study does not affect patients' trust in their doctor (32), in that the doctor only offers the patient one treatment. From the patient's perspective, it does not appear as though treatment is being assigned randomly.

Overall, this study design allows more patients to be included than in a conventional trial (32).

- **Disadvantages**

A pre-randomised trial poses specific ethical problems, particularly in the case of the simple consent design, where a standard treatment is available. This raises the issue of whether patients ought to know how their treatment has been chosen (36). As there is no informed consent in group G1, this design cannot be used in placebo-controlled trials (37). 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 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 (37).

As the comparison is made between all patients randomised to group G1 and all patients randomised to group G2, the efficacy of the new treatment may be obscured if many patients in G2 choose the standard treatment (37). Thus, comparing group G1 with group G2 irrespective of the treatment received dilutes the measurable effect of the new treatment (32). There is therefore a risk that it will not be possible to draw any conclusions.

Finally, a double-blind design is impossible (32), and selection bias is the biggest disadvantage of this approach. In fact, selection bias is accentuated when patients with a poor prognosis are under-represented in the experimental group, because the refusal rate is likely to be proportional to the severity of the condition⁴.

Overall, this type of design more closely resembles an observational study.

- **Constraints**

Patients in the control arm must be informed that their data are being collected. It is also important to compare the characteristics of patients in group 1 with those of patients in group 2 receiving the standard treatment after refusing the new treatment, in order to identify any potential selection bias (32).

- **Applicability**

This experimental design is particularly useful when patients receiving the standard treatment do not require additional visits and when death is the only endpoint (37). It may be applicable in the fields of surgery, interventional techniques, or medical devices (38). Since no information is given to the group randomised to receive the standard treatment, this type of study is debatable from an ethical point of view. Very few studies in France use this design. In practice, it is used primarily for diseases where it is psychologically difficult for the patient to decide on treatment (oncology, palliative care, choosing treatment for one's child, etc.)

7.2.1.2. Expertise-based randomised controlled trials

- **Background**

Conventional randomised controlled trials may be difficult to implement in surgery due to issues of acceptability to the surgeon. Poor acceptability may hamper patient recruitment in a conventional clinical trial (9). Furthermore, surgeons may unconsciously perform the

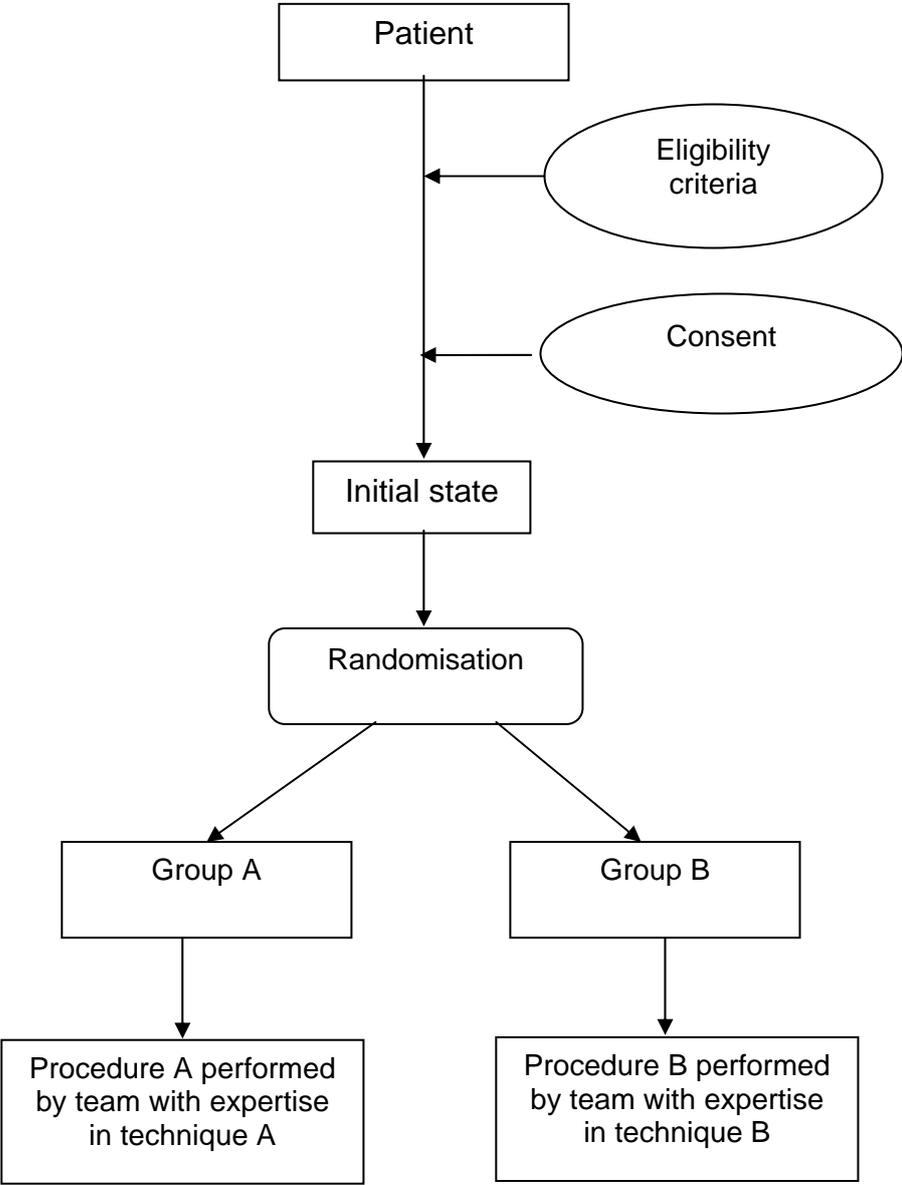
⁴ Under-representation will not occur in the control group because consent is not required in this group.

procedure more meticulously in patients randomised to receive the treatment they prefer, or they may follow up patients in a different manner to less experienced surgeons (9,39). Consequently, the study protocol must be precise and adhered to rigorously. These factors demonstrate the limitations of conventional randomised trials in fields where a procedure is performed. 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) (9). Each surgeon or team only carries out one procedure, at which they are proficient to the point of being experts. Care must be taken to ensure that proficiency in the procedure is comparable in each group (27). 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, from (27)



- **Advantages**

- *Improved internal validity*

Expertise-based randomised controlled trials have better internal validity (39) for the following reasons:

- *No differential expertise bias*

Expertise-based randomised controlled trials can reduce the differential expertise bias that exists in conventional trials (40).

- *Limitation of bias related to the absence of blinding*

This type of trial can also limit differential execution bias related to the absence of blinding. In fact, as each surgeon only performs the procedure that he or she specialises in, the risk of differences between patients as regards the procedure and related factors will be lower than in a conventional trial, where surgeons may follow up patients differently between the different groups (39). 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 obtain the most accurate assessment of a new treatment, patients must be given the treatment that they were randomised to receive. If this does not happen, the deviations from the protocol damage the internal validity of the clinical trial. Devereaux et al (39) have shown that in a conventional clinical trial of treatments for tibial shaft fractures (reaming vs. no reaming), there were differential deviations from the protocol between the two arms, with a higher proportion of patients randomised to the non-reamed procedure group (the more challenging technique). Such deviations from the protocol were more frequent in surgeons who rarely performed the procedure allocated to the patient than in the opposite case (41-43). 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 (39).

- *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 starting the clinical trial (39).

- *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 (9). In addition, there is improved ethical acceptability because surgeons only perform the procedure that they are used to carrying out, which they prefer, and in which they are "experts" (27,39). Overall, surgeons are less reluctant to take part in an expertise-based clinical trial than a conventional trial (27).

- ***Disadvantages***

As with conventional trials in this field, blinding is not possible (39) and there is a substantial risk of measurement (detection) bias.

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 (9). 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 (6,40).

- ***Constraints***

This type of design requires at least one expert in each randomised intervention at each centre (9). In addition, the initial pre-inclusion consultation must be undertaken by a neutral person who determines patients' eligibility (9).

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 relating to the surgeon's expertise (39). 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.

- ***Applicability***

Expertise-based trials are potentially useful in all situations where there is a recognised need to account for the presence of a learning curve. In fact, where using a new medical device requires experience to be acquired, a premature assessment 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
- when expertise in the technique substantially influences the outcome.

7.2.1.3. Tracker trial designs

- **Background**

Medical devices are often marketed before in-depth studies of their efficacy have been conducted. They are also more subject to substantial technological change (44). Tracker trials were developed so that these technological changes or improvements to the procedure could be taken into account during the trial (6).

- **Principle**

Changes during the trial are authorised, recorded and taken into account in the statistical analysis. Changes to the study protocol are also permitted (6).

- **Advantages**

Tracker trials can adapt to developments in technology and in clinical practice. They allow early assessment, thus respecting the principle of equipoise⁵ (where there is no argument to recommend one treatment over another), make maximum use of the different data available, and can rapidly evaluate new, innovative or potentially dangerous techniques (44).

- **Disadvantages**

Analysis is complex because the operator's experience and treatment developments must be taken into account. This type of trial calls for more sophisticated methods than conventional trials. Flexible budgets are required, taking into account the duration of the trial in particular (44).

From a practical point of view, they are difficult to organise (6,44), especially when the technology is evolving because most operators then have little experience (44).

There are few 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 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 (44).

- **Applicability**

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 for rejecting dangerous technologies, and thus could be used in the initial stages.

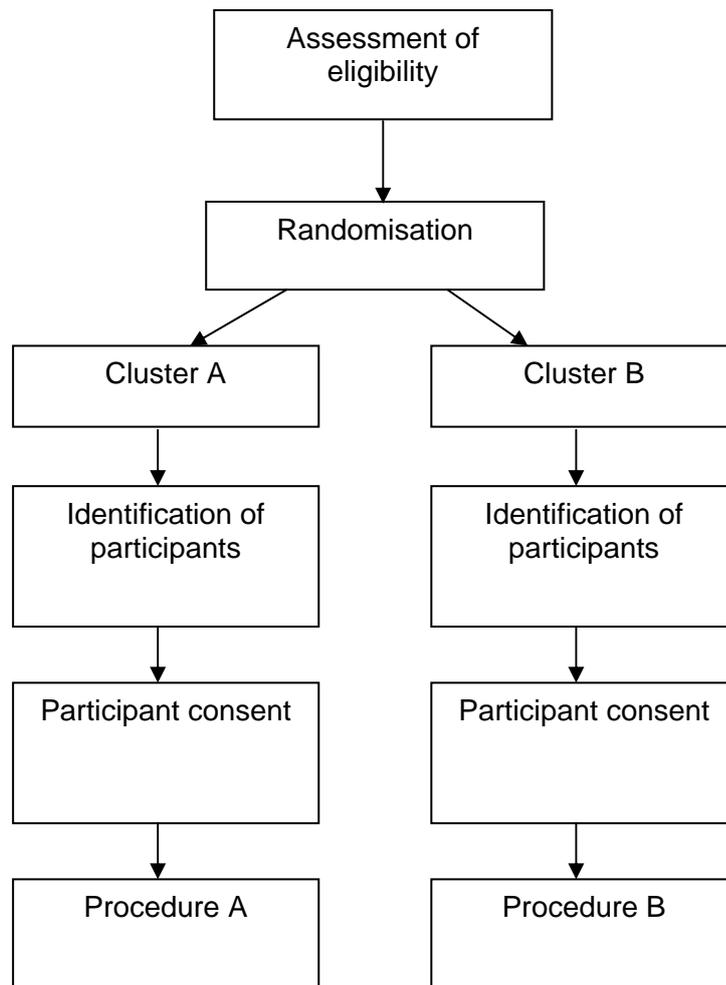
⁵ also known as "clinical equipoise"

7.2.1.4. Cluster randomised trials

- **Principle**

Cluster randomised trials involve randomising groups (clusters) of individuals (by centre, hospital or department) rather than randomising individuals directly (Figure 6). In practice, the technique a centre must use is selected randomly. For example, hospital A will systematically use technique 1, and hospital B will systematically use technique 2.

Figure 6: Diagram of cluster randomised trials, from (45)



Two types of cluster randomised trial design have been described: whole cluster randomised trials, and cluster randomised trials with active recruitment (46). 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

Randomising clusters rather than individuals helps to prevent any contamination between two techniques available at the same centre (46). For example, in the use of a postoperative antiseptic, the risk of administering antiseptic A instead of antiseptic 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 (46).

In addition, 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 (46).

Furthermore, there may be a difference in recruitment between clusters in trials with active recruitment, in terms of the number of patients included and their characteristics (46). An important factor for validity is that all patients are analysed. Finally, there is a risk of contamination between clusters in the case 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 (46). 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 (46).

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. Because patients are recruited after randomisation, participants are selected by someone who knows the treatment they will receive. Finally, patients consent not to randomisation but to participation in a predetermined group; this may result in selection bias (46). Differential recruitment may be considered equivalent to non-response bias, but the number of eligible participants is unknown and therefore difficult to estimate (46).

Cluster randomised trials often raise ethical issues because they require two levels of consent: health professionals give their consent, and then patients are informed that they are taking part in a study (47). In this way, patients have no choice in their treatment. In trials where a new treatment 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 leads to an additional problem with patient blinding (47).

Constraints

When calculating the sample size, account should be taken of the cluster effect, which increases the number of patients required (or the sample size) (46). The statistical analysis should also take this cluster effect into account (46).

In fact, this type of study has a two-level hierarchical data structure: cluster level and individual level. It is necessary to account for the fact that individuals within the same cluster tend to resemble each other more than individuals in different clusters.

For whole cluster randomised trials, steps should be taken to ensure that the person in charge of each cluster adheres to the study protocol before randomisation (46).

In order to limit selection bias in cluster randomised trials with active recruitment, Puffer et al. (48) 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 alternatives have been suggested, such as blinding recruiters to the allocation group or randomising clusters only once the first participant has been included (the index case concept). The latter option avoids empty clusters but cannot prevent differential recruitment (46).

- ***Applicability***

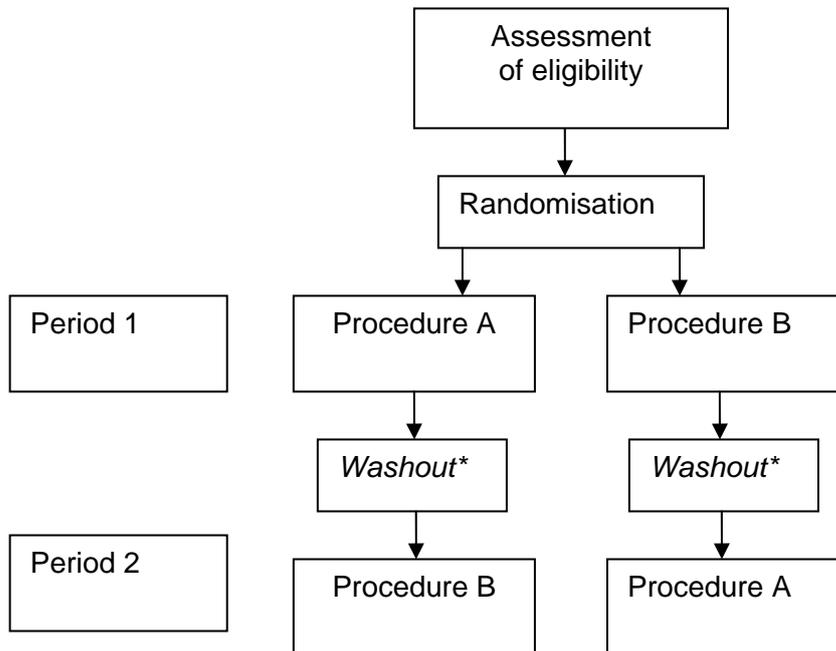
Cluster randomised trials are primarily used for organisation interventions, behavioural interventions and even health promotion programmes. Insofar as they are likely to improve feasibility or practical aspects, they could be considered for medical devices, interventional techniques and surgical techniques. They have already been used in this way to assess different hospital hygiene strategies.

7.2.1.5. 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). Each patient therefore receives both treatments (49). The different treatments are thus compared in the same patient during different periods (50).

Figure 7: The crossover principle



* washout = period without treatment or withdrawal phase

- **Advantages**

As participants are their own controls, with the same precision and with measurements correlated between treatment periods, a smaller number of patients is required (50), which is particularly useful in the case of rare diseases. In addition, patients may express their preference for one or other of the treatments (49). 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 recovery or death (49), or where there is a learning curve. If

measurements are not correlated between the treatment periods, the power of a crossover trial is two times weaker than a parallel group trial with an equal number of patients.

- **Constraints**

Crossover trials work on the assumption that the disease studied is stable, that there is no residual effect from the first treatment when the second treatment is administered, and that the endpoint measurement can be repeated. The washout period is the time without treatment or the withdrawal phase. This period is needed in order to reduce the risk of the carry-over effect (49).

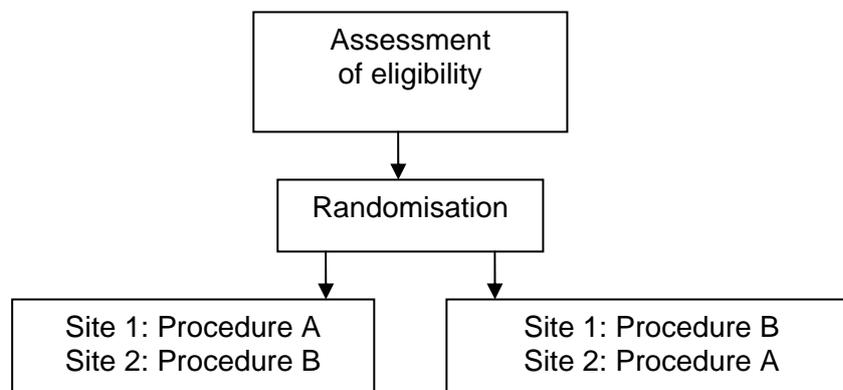
It is necessary to test for any interaction between the treatment effect and the treatment period (49) or any carry-over effect. If there is an interaction, only the data from the first period will be analysed.

Specific tests for matched data must be used to account for the fact that each participant is their own control. Finally, the treatment must have no irreversible effects.

- **Special cases of crossover**

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 presupposes that there is no possible interaction between the two application sites, such as systemic diffusion of the product administered.

Figure 8: Special case of a crossover trial where two treatments are applied simultaneously



- **Applicability**

Crossover trials are used to assess treatments for chronic diseases, in dermatology and in cosmetology. They could be used more widely in the assessment of some medical devices.

7.2.2. Interim analyses

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 groups. 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 planned before the start of the study).

Among these are sequential trials and adaptive trials.

The Bayesian methods described below are also a strategy for analysing data.

7.2.2.1. Sequential trials

- **Background**

Sequential trials are one technique that can be used for interim analysis.

Sequential trials were developed to avoid giving patients a less effective treatment for longer than is necessary (51).

- **Principle**

The different stages are planned independently of any data previously collected (52).

Various sequential analysis methods have been described. The traditional sequential method includes the analysis of pairs of patients (53), sequential group analyses every n participants (54-58), flexible methods (59,60) used to choose when to perform analyses (57,58), and continuous sequential analysis such as the sequential probability ratio test and the triangular test (57,58,61).

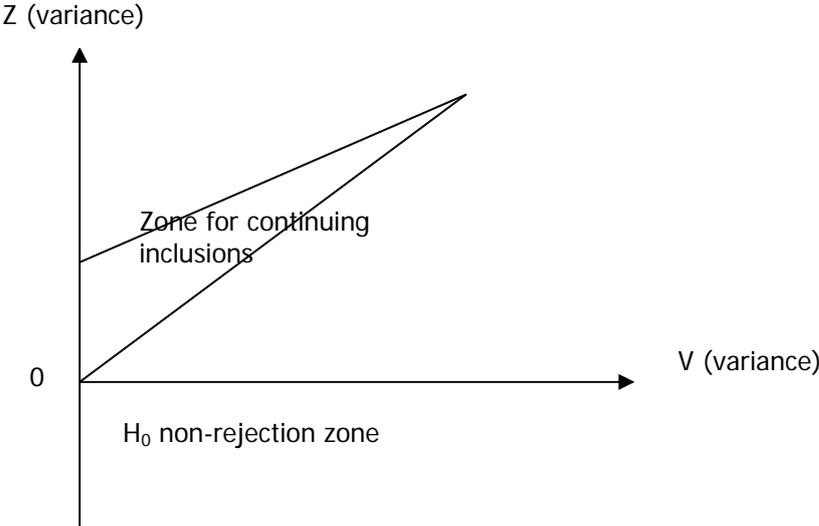
Sequential trials are randomised controlled trials where the study results from patients included are examined before randomising new patients, and a decision is made as to whether to continue the trial or not. The aim is to avoid ethical problems with continuing the study if the difference between treatments turns out to be more significant than was expected (62).

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 (51). In both these cases, new inclusions are stopped.

Figure 9 and Figure 10 show the graphical rules for deciding whether to stop or continue inclusions.

The specifics of these different methods are presented in Table 4.

Figure 9: Graphical decision-making rules for the triangular test (closed design), from (63,64)
Sequential analysis of clinical trials



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

Figure 10: Graphical decision-making rules for the sequential probability ratio test (open design)

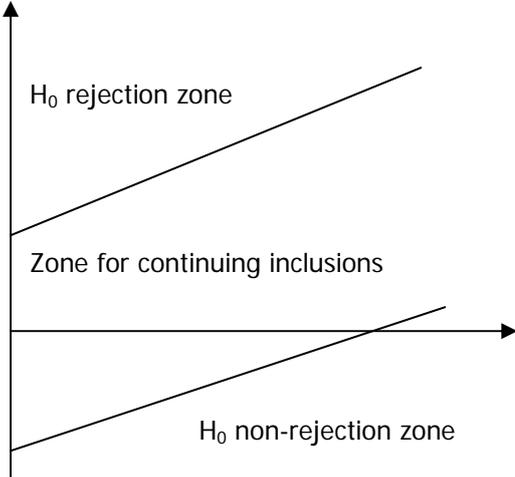


Table 4: Specific principles, advantages and disadvantages of different methods for sequential analysis

Name	Principle	Advantages	Disadvantages
Conventional sequential method			
Armitage 1975 (53)	<ul style="list-style-type: none"> - Matching: either the patient is his/her own control, or there are 2 separate subjects with prognostic factors taken into account or selected randomly - Analysis is performed every 2 patients for endpoints obtained rapidly after inclusion 		<ul style="list-style-type: none"> - Hard work to implement - Requires frequent analyses - Requires a single and rapidly obtained endpoint - Not suited to censored data - Open design: a conclusion may never be reached
Group sequential methods: analysis every n subjects			
Pocock 1977, Pocock 1982 (54,65)	<ul style="list-style-type: none"> - Same alpha risk for every analysis, determined by the number of interim analyses - Interim analyses performed after inclusion of a fixed number of patients 	<ul style="list-style-type: none"> - Simple - Allows earlier conclusion than O'Brien and Fleming rule (57,58,65) - Allows earlier conclusion than Peto rule (57,58,65) 	<ul style="list-style-type: none"> - Does not allow early conclusion in the case of equivalence between 2 treatments - Number of analyses fixed from the start - Results sometimes conflicting, depending on whether an interim analysis was done or not (57,58) - Requires the highest number of subjects (66)
O'Brien and Fleming 1979 (55)	Alpha risk increases with each interim analysis (57)	Allows earlier conclusion than Peto	Number of analyses fixed from the start
Peto (56)	Boundary $p < 0.001$ for each analysis	More flexible in terms of number of interim analyses (57)	
Succession method Falissard 1992 (58)	Requires results to be significant to 0.005 for a succession of interim analyses		

Flexible methods			
Lan (67)	Different alpha for each analysis, determined from past and ongoing interim analyses	Number of interim analyses not predefined	- Risk of increased sample size - Low probability of concluding during the trial (51)
Semi-sequential method (57)	Adapted succession method with free choice of date and number of interim analyses (Falissard, Lelouch, Resp)		
Continuous sequential analysis (59), SPRT and TT (61,63)			
Triangular test		Method of choice Closed design Earlier conclusion	
Sequential probability ratio test			Open design: risk of not concluding and of a long trial duration

- ***Advantages of sequential methods***

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) and 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 (62).

Furthermore, the mean number of patients needed in this type of trial is lower than in conventional trials, independent of the therapeutic effect or the power (51,68).

Benichou et al. more specifically compared the triangular test, sequential probability ratio test and group sequential analysis. They showed that these three methods have correct statistical properties, but that the triangular test allowed an earlier conclusion, whatever the size of the difference between the two groups compared (63,64,69). The triangular test is therefore preferable to the group sequential analysis, which does not allow an early conclusion if the two treatments compared are equivalent.

- ***Disadvantages of sequential methods***

The sequential methods have practical limitations. On the one hand, they allow two treatments to be compared. On the other hand, 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 (62). Stopping the trial early risks 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 (51). Finally, the number of participants is unknown at the start, and the confidence intervals are wider than those obtained with a conventional trial (51).

The external validity of sequential trials has also been called into question because of the fact that results can be obtained from a small sample, potentially with a selection bias, but this argument has been refuted by Van der Lee et al. because of randomisation (51).

The boundary approach presupposes that the time from inclusion to measuring the endpoint is short. If many 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 (51). This will lead not to a reduction of the sample size but to a reduction in the duration of the trial. The boundaries are defined from prior knowledge and cannot be changed as a result of data collected from the sequential analyses (51).

- ***Constraints***

The results at each stage must be confidential. An independent data monitoring committee must therefore be formed, which will decide on changes to the number of patients necessary and on whether inclusions should be continued or stopped, without giving detailed information on the analysis results to people involved in the trial (investigators, evaluators, patients) (51). There must be regular inclusions so as to comply with the necessary frequency of analyses, and good-quality follow-up to avoid delays in updating data. Procedures allowing patients to be monitored regularly with prompt submission of information are required. Long-term follow-up after the trial has been stopped is also essential for identifying adverse effects.

The use of specific methods to calculate confidence intervals (57,58) and correct bias in assessments is also required.

- ***Applicability***

Currently, these analysis techniques are of particular value for rare diseases and in paediatrics. More broadly, they could allow a treatment strategy that proves ineffective, including the loss of opportunity for patients exposed to it, to be stopped earlier,.

7.2.2.2. Adaptive randomisation trials

- ***Background***

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 1:1 to each arm, but 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 (62).

Adaptive trials should not be confused with trials where the allocation ratio is adapted to the preliminary results of the trial (70).

- ***Principle***

Like sequential trials, adaptive trials are based on the use of interim analyses. But in this type of trial, the study procedure may be changed on the basis of the results observed during interim analyses (51) without increasing the alpha risk (71). All information collected during the different stages is used, and not just that from the current stage (52). The allocation of a treatment to a patient depends on the results obtained in previous patients (71).

If changes are made, a new phase of the trial starts, and the analysis covers the data gathered in this new phase and no longer all of the data. The various amendments possible are a re-assessment of the number of subjects needed, amendments to endpoints, and the addition or deletion of interim analyses (51).

Adaptive trials may consist of two phases or several phases (72). 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 adaptive trials in two phases.

- **Advantages**

Adaptive trials are highly flexible and reduce the sample size. In addition, all the information collected is used in the analysis (51,52). On the other hand, it has been shown that this study design can handle a larger number of adequately treated patients in comparison with sequential trials and randomised trials (62,73).

- **Disadvantages**

Adaptive trials have more logistical constraints than other types of trials (52). Extreme vigilance is also needed 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 (62). As with sequential trials, a simple and unequivocal efficacy endpoint must be defined (62). Van der Lee raises the issue of how to define rules for stopping the trial or how to interpret the results if the primary endpoint has changed during the study (51).

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 (53,62).

Because of these disadvantages, adaptive trials are considered to be of little benefit in comparison with sequential trials (62,74,75).

- **Constraints of adaptive trials**

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.

An independent steering committee must therefore be formed, which will decide on changes to the number of patients necessary and on whether inclusions should be continued or stopped, without giving detailed information on the results of the analysis to people involved in the trial (investigators, assessors, patients) (51,76).

Finally, it is crucial for the credibility of the final conclusions that hypotheses and adaptations made as the trial is conducted are recorded prospectively (51).

- **Applicability**

Adaptive trials may be useful for operator-dependent techniques, in surgery, in studies of interventional or surgical techniques that require a learning or improvement phase, and in the assessment of medical devices.

7.2.2.3. Bayesian methods

Bayesian methods can be used as alternatives to conventional sequential methods (77).

Bayesian methods are also used to perform interim analyses. Thus, in a randomised controlled trial comparing paclitaxel-eluting stents and a control group in patients with severe emphysema, the planned interim analysis was based on Bayesian methods (38).

These approaches are also better at managing multiple endpoints (78).

- **Background**

In conventional randomised trials, data from previous studies are used when the trial is being designed. Subsequently, only information collected during the trial is used. By contrast, the Bayesian approach combines existing information with information from the ongoing trial (78).

To stay within an acceptable Bayesian approach, the existing information must:

- remain non-informative and not influence the result of the study;
- or use the results of other valid studies.

The use of Bayesian methods for medical devices is recommended in the recent FDA guidelines (79).

- **Principle**

The Bayesian approach allows clinical information available *a priori* to be used in the trial's analysis. This involves using existing information which may be supplied by the literature. Some publications argue that it is even possible to use expert opinions by modelling them. This point is debatable because it introduces a substantial arbitrary element to the result of the study.

The existing information and the trial results are viewed as coming from the same set of data.

In a Bayesian clinical trial, uncertainty about a parameter of interest is described in terms of probabilities, updated as information is collected during the trial. The probabilities in the prior distribution are based on data from previous trials (78) (80). With Bayes' theorem, posterior probabilities are estimated from the ongoing trial and are contingent on prior probabilities. This approach combines previous information with the data provided by the trial (78). These models are extremely dependant on existing information.

Bayesian estimation provides not confidence intervals but credibility intervals based on the posterior distribution (78). Unlike the so-called "frequentist" approach, there are no statistical tests but results with a 95% credible interval. These methods offer a certain amount of flexibility.

- **Advantages**

Bayesian methods can indicate the probability that the treatment is effective. In addition, they allow more information to be used to make a decision. In a similar way to a meta-analysis, the use of valid previous data can increase the information and precision that a trial offers. In some cases, this type of method can reduce the sample size, either because of the addition of existing information, or by using adaptive trials. Another advantage is the great flexibility of these techniques during a trial.

Thus, the Bayesian method allows all available information (both past and current) on a particular technology to be used. In addition, a Bayesian approach is particularly useful in situations where the number of subjects is small. Finally, the use of existing information can reduce the number of participants and thus also the duration and resources needed for a clinical trial (78,80).

- **Disadvantages**

The **major** disadvantage is the risk of including arbitrary information as existing information, thus substantially influencing the final result.

This would amount to including in a meta-analysis a large study, biased in favour of the treatment, fully compensating for the results of a small study with no bias.

- **Constraints**

Bayesian analysis involves generating hypotheses and simulations, namely relating to existing information (supplied by previous data or expert consensus), posterior information (which will be obtained from the trial data), and the mathematical model used to combine the two. The influence of covariates on events or missing data must also be considered, with sensitivity analyses planned (78).

This initial work is particularly important.

- **Applicability**

While the so-called “frequentist” approach of conventional trials remains the norm in biomedical literature, the publication of trials using Bayesian methods has been on the increase for several years (79).

7.3. Comparative non-randomised observational studies

7.3.1. Type of study

This type of study should be used only under circumstances where it is impossible to conduct a randomised controlled trial. The choice of an observational study should remain the exception. Observational studies offer lower-quality evidence than randomised trials.

7.3.1.1. Comprehensive cohort studies or trials based on patient preference

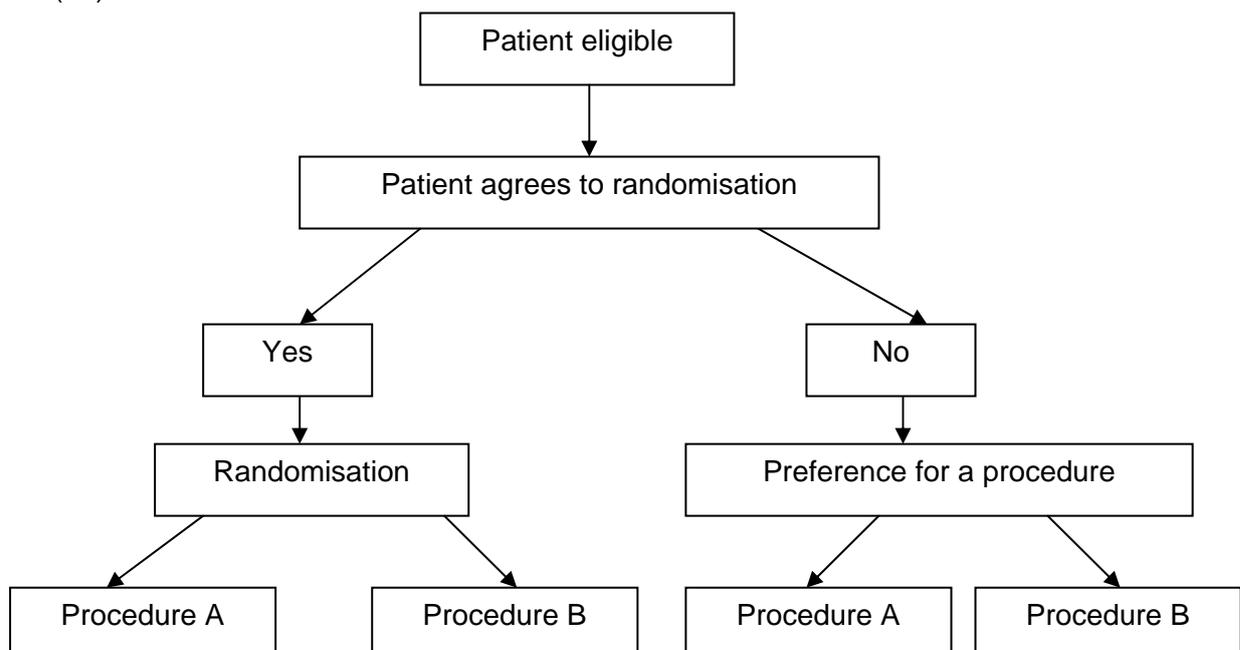
- Background

In trials comparing a surgical treatment and a non-surgical treatment, patients may view the non-surgical treatment group as a less effective intervention, and refuse to participate in the study (81). 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 or trials based on patient preference were developed to overcome these obstacles.

- Principle

All patients meeting the eligibility criteria for a trial are recruited, independently of whether they consent to randomisation (81). Patients who refuse to be randomised receive the treatment of their choice (Figure 11).

Figure 11: Principle of studies based on patient preference or comprehensive cohort studies, from (82)



- 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 (83).

- 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. The difference observed may be due not to the treatment but to other non-controlled variables (confounding factors).

- Constraints

The randomisation variable is introduced to the model as an adjustment covariate.

- Applicability

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.

7.3.1.2. Prospective comparative observational studies

- Background

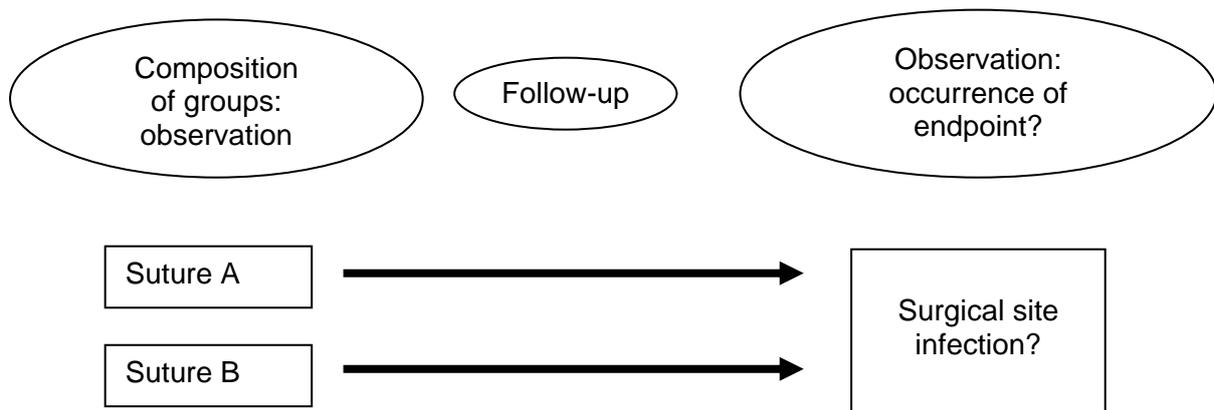
Observational studies may be used when randomised controlled trials cannot be considered, for example for ethical reasons (62). Their main limitation is non-comparability between the groups, which could explain any apparent treatment effect. It is crucial that confounding factors in observational studies are taken into account (84).

- Principle

Prospective comparative non-randomised studies involve comparing the incidence of an endpoint (for example infection following surgery) between a group receiving procedure A (an intervention or medical device, for example the application of antiseptic X) and a group not receiving procedure A (either receiving another procedure, or not receiving anything). The distribution of patients in each of these groups is not generated by the investigator, but is observed data (Figure 12). The principle is the same as in exposed/unexposed studies.

^

Figure 12: Example of a prospective observational study comparing the occurrence of postoperative infections depending on the type of suture used



- Advantages

A prospective design means that plans can be made to collect all the information required. It also allows precise information on the procedure to be collected (doses administered by imaging equipment, for example, or duration of the procedure). In addition, it is possible to measure the incidence of an event or relative risk and to take into account the time interval between the procedure and occurrence of the event.

- Disadvantages

Like all observational studies, prospective comparative studies are subject to confounding. There is no guarantee that the two groups will be comparable on other variables, especially variables that have not been measured. In addition, there may be a long follow-up period, increasing the risk of patients being lost to follow-up.

- Constraints

The control group must be comparable to the group receiving the study procedure on other variables, and must be followed up in an identical way to the study group. If the follow-up period is especially long, procedures must be put into place to limit the risk of patients being lost to follow-up. The observation period must be defined.

- Applicability

Currently, most of these studies are exposed/unexposed cohort studies.

When a randomised trial is impossible, they can be used to compare very different interventions such as surgical techniques, medical devices and other procedures.

7.3.2. Analysis strategies

7.3.2.1. Propensity score

- **Background**

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 one receiving treatment B. There is no guarantee that the groups are comparable. The propensity score is a statistical method that aims to reduce the confounding factors associated with observational studies (85).

- **Principle**

The propensity score is defined as the conditional probability of a patient receiving treatment A rather than treatment B given his baseline characteristics (propensity to receive a treatment according to characteristics). The objective is to balance the distribution of covariates (such as age, gender, comorbidities and severity of disease) between the groups, so as to neutralise confounding factors (86,87). Comparisons between the treated and control group are made within the same propensity score category (85). In this way, treated participants and controls with equal propensity scores can be considered as similar (84).

Once calculated, the propensity score can be used as a matching variable, a stratification variable or an adjustment variable. The methods of choice are stratification and matching, for example stratifying the analysis into quintiles based on the propensity score. In this way, within each of these strata, most bias related to measured confounding factors disappears. Matching can be more complex because the propensity score has a continuous scale (84).

Example

In a study aiming to compare the incidence of deep venous thrombosis depending on whether low-molecular-weight heparin (LMWH) was administered or not, the propensity score corresponded to the probability of a patient receiving LMWH depending on his or her baseline characteristics. Once the propensity score was determined, it was classified into quintiles with increasing values, with patients in the first quintile having the lowest probability of receiving LMWH and patients in the fifth quintile having the highest (88). The propensity score was used as a stratification variable. The authors also used the propensity score as a matching variable: each treated patient was matched to the non-treated patient with the closest matching score, as long as the difference was less than 0.1.

The propensity score method works best under three conditions (87) :

- when the event studied is rare
- when there are many 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.

In addition, it allows a number of unbalanced variables to be adjusted for simultaneously, thus reducing bias in the comparison between the two treatments (87). In this way, stratifying based on the propensity score reduces bias caused by differences between the characteristics observed by more than 90% (89).

- **Disadvantages**

Adjustments can only be made for recorded variables (87). There is therefore no guarantee that the groups will be comparable for unmeasured variables (84). However, sensitivity analyses can be performed to assess the robustness of the results (85,90).

Propensity score analysis is not appropriate in small samples (< 50). In these cases, some variables may be unbalanced (87). This is also the case if there is a severe imbalance between the groups in terms of variables on inclusion (87).

Finally, the propensity score cannot replace a randomised trial or eliminate all selection bias (87).

- **Constraints**

Propensity score analysis must be planned in advance. All relevant variables to be collected must be specified in the protocol. In addition, patient populations must be comparable between the two treatments. 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 confirmed that the distribution of the propensity score is comparable in both groups. Finally, missing data must be estimated using specific techniques (such as multiple imputation) so as not to exclude patients, which would make the validity of the results questionable. A sensitivity analysis is usually performed for unobserved variables to identify any hidden bias (87).

- **Applicability**

Propensity score analysis is very useful in observational studies when a randomised trial cannot be conducted because of medical practices, patient preference, the organisation of healthcare or economic constraints (88). The sample must be of a sufficient size. This analysis strategy could be used for medical devices, in surgery and for interventional techniques.

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 event due to the differential use of this treatment may be very small and, consequently, difficult to assess (84).

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 (96). In addition, a good instrumental variable estimator relies on large numbers of patients (96). Where numbers are small, the estimates are incorrect. Finally, interpretation is difficult, particularly when the effect of treatment differs.

- ***Constraints***

A good instrumental variable must satisfy three major assumptions (92):

- It must have a strong relationship with the treatment 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 treatment.

- ***Applicability***

Instrumental variables are frequently used in econometrics. However, the inherent constraints of these variables make it difficult to consider their use in medical device assessment.

8. In practice

Under some circumstances, a conventional randomised controlled trial cannot be conducted during the clinical development of a new medical device. Other experimental designs may be proposed to demonstrate the clinical benefit of the MD.

The choice may be guided by:

Characteristics of the medical device

- Substantially and rapidly evolving technology:
 - Adaptive randomisation trials
 - Bayesian methods
 - Possibly tracker trials
- A small target population:
 - Bayesian methods
 - Crossover trials, as long as the disease studied is stable and the endpoint outcome can be repeated
 - Sequential trials
- Potentially serious adverse events:
 - Sequential trials

Medical acceptability

- Experience in the technique influences the result or the techniques are very different:
 - Expertise-based randomised controlled trials
 - Under some circumstances, cluster randomised trials

Acceptability to the patient

- The comparator is an invasive technique or use of the technique is widespread:
 - Zelen's design or randomised consent design
 - Trials based on patient preference or comparative observational studies if randomisation is impossible

Figure 14 shows a decision tree used to guide the choice of type of study. The respective quality of evidence for different types of study is illustrated by Figure 15.

Prospective non-randomised trials are justified only when no other type of trial can be used. This decision must be justified by an unequivocal scientific argument.

Figure 14: Decision tree for type of study

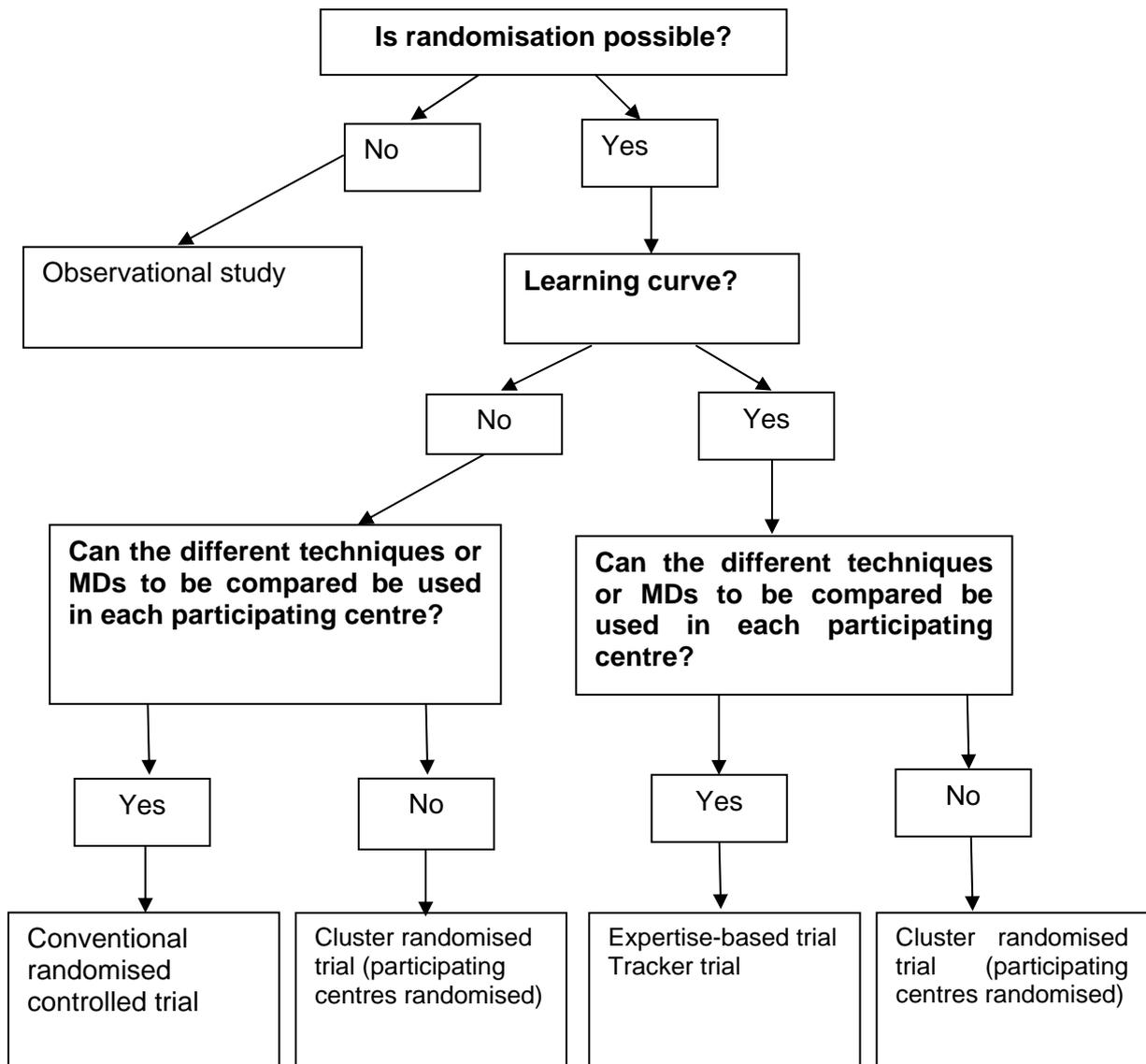
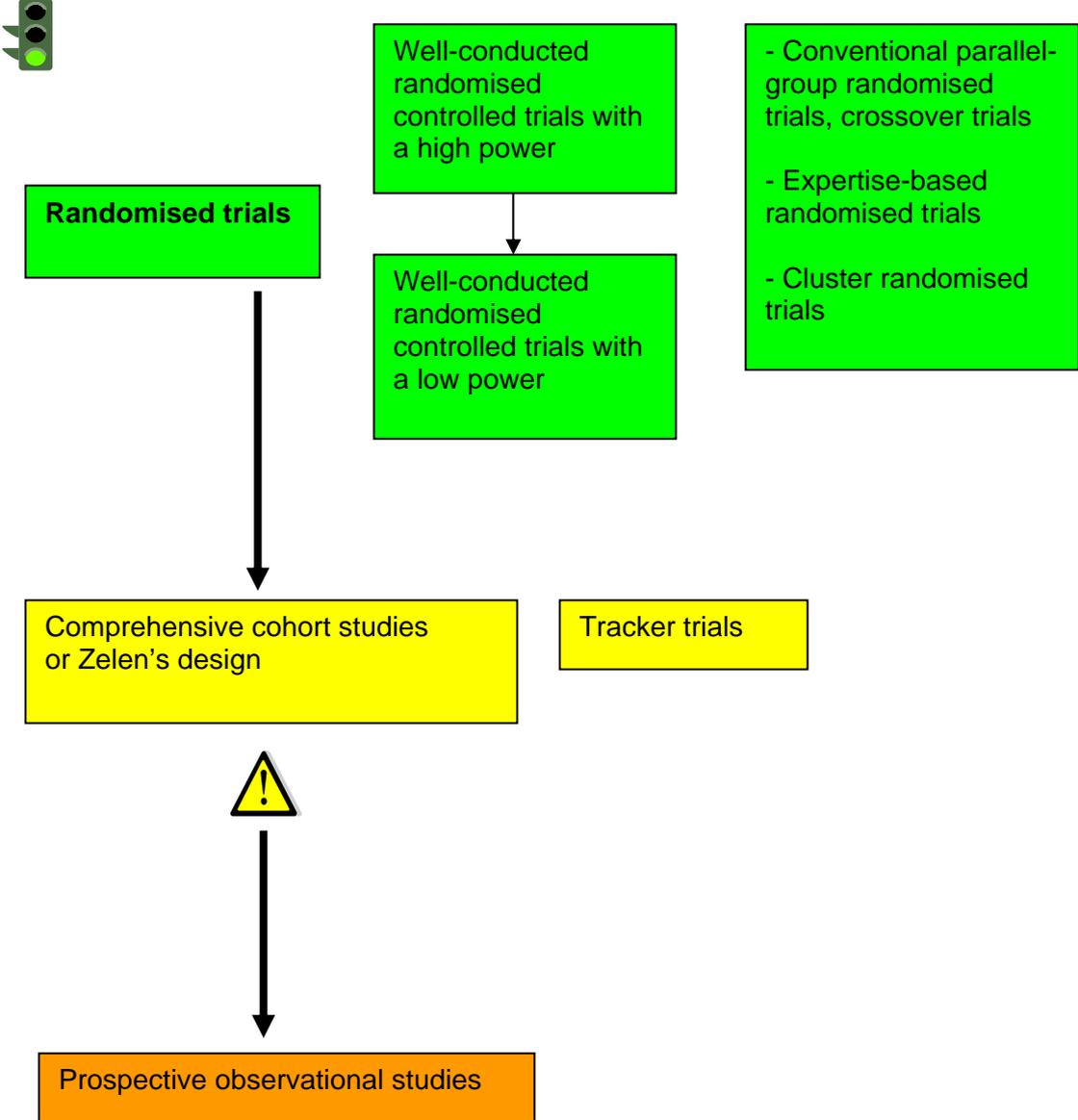


Figure 15: Suggested classification of study types by quality of evidence



Bibliography

1. Agence nationale d'accréditation et d'évaluation en santé. Guide d'analyse de la littérature et gradation des recommandations. Saint-Denis la Plaine: ANAES; 2000. http://www.has-sante.fr/portail/upload/docs/application/pdf/analite_rat.pdf
2. Boutron I, Moher D, Tugwell P, Giraudeau B, Poiraudou S, Nizard R, *et al.* A checklist to evaluate a report of a nonpharmacological trial (CLEAR NPT) was developed using consensus. *J Clin Epidemiol* 2005;58(12):1233-40.
3. Centre fédérale d'expertise des soins en santé, Vinck, I, Neyt, M, Thiry, N, Louagie, M, Ghinet, D, *et al.* Procédure d'évaluation des dispositifs médicaux émergents. Bruxelles: Centre fédérale d'expertise des soins en santé (KCE); 2006.
4. EuroScan, Simpson, S, Hiller, J, Gutierrez-Ibarluzea, I, Kearney, B, Norderhaug, I, *et al.* A toolkit for the identification and assessment of new and emerging health technologies. Birmingham: EuroScan; 2009.
5. Bernard A. Le développement clinique des nouveaux dispositifs médicaux. e-mémoires de l'Académie Nationale de Chirurgie 2012;11(4):027-31.
6. Ergina PL, Cook JA, Blazeby JM, Boutron I, Clavien PA, Reeves BC, *et al.* Challenges in evaluating surgical innovation. *Lancet* 2009;374(9695):1097-104.
7. Bernard A, Vicaut E. Quelles études pour évaluer des nouveaux dispositifs médicaux ? *J Chir* 2009;146(2):129-35.
8. Bernard A, Vicaut E. Dispositifs médicaux. Première partie : de quels types d'études avons-nous besoin ? *Médecine* 2008;4(10):464-7.
9. Bednarska E, Bryant D, Devereaux PJ. Orthopaedic surgeons prefer to participate in expertise-based randomized trials. *Clin Orthop Relat Res* 2008;466(7):1734-44.
10. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, *et al.* CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 2008;5(1).
11. Boutron I, Tubach F, Giraudeau B, Ravaud P. Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. *JAMA* 2003;290(8):1062-70.
12. Weil RJ. The future of surgical research. *PLoS Med* 2004;1(1):e13.
13. Freed CR, Greene PE, Breeze RE, Tsai WY, DuMouchel W, Kao R, *et al.* Transplantation of embryonic dopamine neurons for severe Parkinson's disease. *N Engl J Med* 2001;344(10):710-9.
14. Tamis-Holland JE, Palazzo A, Stebbins AL, Slater JN, Boland J, Ellis SG, *et al.* Benefits of direct angioplasty for women and men with acute myocardial infarction: results of the Global Use of Strategies to Open Occluded Arteries in Acute Coronary Syndromes Angioplasty (GUSTO II-B) Angioplasty Substudy. *Am Heart J* 2004;147(1):133-9.
15. Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. Systematic reviews in health care. *BMJ* 2001;323(7303):42-6.
16. Boutron I, Tubach F, Giraudeau B, Ravaud P. Blinding was judged more difficult to achieve and maintain in nonpharmacologic than pharmacologic trials. *J Clin Epidemiol* 2004;57(6):543-50.
17. Boutron I, Guittet L, Estellat C, Moher D, Hróbjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Med* 2007;4(2).
18. Kinnaert P. La chirurgie factuelle : aspects éthiques et méthodologiques. *Rev Med Brux* 2006;27:451-8.
19. Kinnaert P. Placebo et effet placebo (deuxième partie) : aspects éthiques. *Rev Med Brux* 2007;28:39-44.
20. Ravaud, P. Assessment des traitements non médicamenteux. 4^{ème} conférence d'épidémiologie clinique. 26-28 mai 2010. Paris: 2011.
21. Li H, Yue LQ. Statistical and regulatory issues in nonrandomized medical device clinical studies. *J Biopharm Stat* 2008;18(1):20-30.
22. Freeman TB, Vawter DE, Leaverton PE, Godbold JH, Hauser RA, Goetz CG, *et al.* Use of placebo surgery in controlled trials of a cellular-based therapy for Parkinson's disease. *N Engl J Med* 1999;341(13):988-92.
23. Macklin R. The ethical problems with sham surgery in clinical research. *N Engl J Med*

- 1999;341(13):992-6.
24. Macklin R. Placebo surgery in trials of therapy for Parkinson's disease. *N Engl J Med* 2000;342(5):355.
25. Master Z, McLeod M, Mendez I. Benefits, risks and ethical considerations in translation of stem cell research to clinical applications in Parkinson's disease. *J Med Ethics* 2007;33(3):169-73.
26. Boutron I, Estellat C, Guittet L, Dechartres A, Sackett DL, Hróbjartsson A, *et al.* Methods of blinding in reports of randomized controlled trials assessing pharmacologic treatments: a systematic review. *PLoS Med* 2006;3(10):e425.
27. Rudicel S, Esdaile J. The randomized clinical trial in orthopaedics: obligation or option? *J Bone Joint Surg Am* 1985;67(8):1284-93.
28. Campbell G. Statistics in the world of medical devices: the contrast with pharmaceuticals. *J Biopharm Stat* 2008;18(1):4-19.
29. Keeley EC, Boura JA, Grines CL. Primary angioplasty versus intravenous thrombolytic therapy for acute myocardial infarction: a quantitative review of 23 randomised trials. *Lancet* 2003;361(9351):13-20.
30. Stone GW, Grines CL, O'Neill WW. Primary coronary angioplasty versus thrombolysis. *N Engl J Med* 1997;337(16):1168-9.
31. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. *Ann Intern Med* 2002;137(6):511-20.
32. Zelen M. A new design for randomized clinical trials. *N Engl J Med* 1979;300(22):1242-5.
33. Blondel B, Goffinet F, Bréart. *Assessment des soins en obstétrique*. Paris: Masson; 2001.
34. Zelen M. Innovations in the design of clinical trials in breast cancer. *Breast Cancer Res Treat* 1983;3(2):137-42.
35. Zelen M. Strategy and alternate randomized designs in cancer clinical trials. *Cancer Treat Rep* 1982;66(5):1095-100.
36. Zelen M. Randomized consent designs for clinical trials: an update. *Stat Med* 1990;9(6):645-56.
37. Horwitz RI, Feinstein AR. Advantages and drawbacks of the Zelen design for randomized clinical trials. *J Clin Pharmacol* 1980;20(7):425-7.
38. Shah PL, Slebos DJ, Cardoso PFG, Cetti EJ, Sybrecht GW, Cooper JD. Design of the exhale airway stents for emphysema (EASE) trial: an endoscopic procedure for reducing hyperinflation. *BMC Pulm Med* 2011;11:1.
39. Devereaux PJ, Bhandari M, Clarke M, Montori VM, Cook DJ, Yusuf S, *et al.* Need for expertise based randomised controlled trials. *BMJ* 2005;330(7482):88-91.
40. Mastracci TM, Clase CM, Devereaux PJ, Cinà CS. Open versus endovascular repair of abdominal aortic aneurysm: a survey of Canadian vascular surgeons. *J Can Surg* 2008;51(2):142-8.
41. DeTurris SV, Cacchione RN, Mungara A, Pecoraro A, Ferzli GS. Laparoscopic herniorrhaphy: beyond the learning curve. *J Am Coll Surg* 2002;194(1):65-73.
42. Menon VS, Manson JMCK, Baxter JN. Laparoscopic fundoplication: learning curve and patient satisfaction. *Ann R Coll Surg Engl* 2003;85(1):10-3.
43. Lobato AC, Rodriguez-Lopez J, Diethrich EB. Learning curve for endovascular abdominal aortic aneurysm repair: assessment of a 277-patient single-center experience. *J Endovasc Ther* 2002;9(3):262-8.
44. Lilford RJ, Braunholtz DA, Greenhalgh R, Edwards SJL. Trials and fast changing technologies: the case for tracker studies. *BMJ* 2000;320(7226):43-6.
45. Hahn S, Puffer S, Torgerson DJ, Watson J. Methodological bias in cluster randomised trials. *BMC Med Res Methodol* 2005;5:10.
46. Giraudeau B, Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med* 2009;6(5).
47. Campbell MJ. Cluster randomized trials in general (family) practice research. *Stat Methods Med Res* 2000;9(2):81-94.
48. Puffer S, Torgerson DJ, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ* 2003;327(7418):785-9.
49. Elbourne DR, Altman DG, Higgins JPT, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. *Int J Epidemiol* 2002;31(1):140-9.
50. Hills M, Armitage P. The two-period cross-over clinical trial. *Br J Clin Pharmacol*

- 1979;8(1):7-20.
51. van der Lee JH, Wesseling J, Tanck MWT, Offringa M. Efficient ways exist to obtain the optimal sample size in clinical trials in rare diseases. *J Clin Epidemiol* 2008;61(4):324-30.
52. Wassmer G. Basic concepts of group sequential and adaptative group sequential test procedures. *Stat Pap* 2000;41:253-79.
53. Armitage P. *Sequential medical trials*. 2nd ed ed. New York: Wiley; 1975.
54. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 1977;64(2):191-9.
55. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979;35(3):549-56.
56. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, *et al*. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer* 1976;34(6):585-612.
57. Falissard B, Lellouch J. Some extensions to a new approach for interim analysis in clinical trials. *Stat Med* 1991;10(6):949-55.
58. Falissard B, Lellouch J. A new procedure for group sequential analysis in clinical trials. *Biometrics* 1992;48(2):373-88.
59. DeMets DL, Ware JH. Group sequential methods for clinical trials with a one-sided hypothesis. *Biometrika* 1980;67(3):651-60.
60. DeMets DL, Ware JH. Asymmetric group sequential boundaries for monitoring clinical trials. *Biometrika* 1982;69(3):661-3.
61. Whitehead J, Stratton I. Group sequential clinical trials with triangular continuation regions. *Biometrics* 1983;39(1):227-36.
62. Clayton DG. Ethically optimised designs. *Br J Clin Pharmacol* 1982;13(4):469-80.
63. Benichou J, Chastang C. A simulation study of three sequential methods for the comparison of two treatment groups when the response criterion is censored. *Stat Med* 1986;5(4):375-85.
64. Benichou J, Chastang CK. Mise en œuvre de méthodes séquentielles dans les essais thérapeutiques dont le critère de jugement est censuré. *Rev Epidemiol Santé Publ* 1986;34:196-208.
65. Pocock SJ. Interim analyses for randomized clinical trials: the group sequential approach. *Biometrics* 1982;38(1):153-62.
66. Pocock SJ, Hughes MD. Practical problems in interim analyses, with particular regard to estimation. *Control Clin Trials* 1989;10(4 Suppl):209S-21S.
67. Lan GKK, Simon R, Halperin M. Stochastically curtailed tests in long-term clinical trials. *Sequential Analysis* 1982;1(3):207-19.
68. Sebillé V, Bellissant E. Comparison of four sequential methods allowing for early stopping of comparative clinical trials. *Clin Sci* 2000;98(5):569-78.
69. Benichou J, Chastang C. Analyse séquentielle des essais thérapeutiques randomisés dont le critère de jugement est la survie. Utilisation du test triangulaire. *Thérapie* 1987;42(3):295-9.
70. Coad DS, Ivanova A. The use of the triangular test with response-adaptive treatment allocation. *Stat Med* 2005;24(10):1483-93.
71. Bauer P, Brannath W. The advantages and disadvantages of adaptive designs for clinical trials. *Drug Discov Today* 2004;9(8):351-7.
72. Weinstein MC. Allocation of subjects in medical experiments. *N Engl J Med* 1974;291(24):1278-85.
73. Day NE. A comparison of some sequential designs. *Biometrika* 1969;56:301-11.
74. Lellouch J, Schwartz D. L'essai thérapeutique: Ethique individuelle ou éthique collective. *Rev Inst Int Stat* 1971;39:127-36.
75. Simon R. Adaptive treatment assignment methods and clinical trials. *Biometrics* 1977;33(4):743-9.
76. Hung HMJ. Considerations in adapting clinical trial design. *J Formos Med Assoc* 2008;107(12 Suppl):S14-S18.
77. Freedman LS, Spiegelhalter DJ. Comparison of Bayesian with group sequential methods for monitoring clinical trials. *Control Clin Trials* 1989;10(4):357-67.
78. Center for Devices and Radiological Health, Center for Biologics Assessment and Research. Guidance for industry and FDA staff. Guidance for the use of bayesian statistics in medical device clinical trial. Silver Spring, Rockville: CDRH, CBER; 2006.

<http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf>

79. Campbell G. Bayesian statistics in medical devices: innovation sparked by the FDA. *J Biopharm Stat* 2011;21(5):871-87.
80. Pibouleau L, Chevret S. Bayesian statistical method was underused despite its advantages in the assessment of implantable medical devices. *J Clin Epidemiol* 2011;64(3):270-9.
81. Kearney RS, Achten J, Parsons NR, Costa ML. The comprehensive cohort model in a pilot trial in orthopaedic trauma. *BMC Med Res Methodol* 2011;11:39.
82. Olschewski M, Scheurlen H. Comprehensive Cohort Study: an alternative to randomized consent design in a breast preservation trial. *Methods Inform Med* 1985;24(3):131-4.
83. Torgerson DJ, Sibbald B. Understanding controlled trials. What is a patient preference trial? *BMJ* 1998;316(7128):360.
84. Klungel OH, Martens EP, Psaty BM, Grobbee DE, Sullivan SD, Stricker BH, *et al.* Methods to assess intended effects of drug treatment in observational studies are reviewed. *J Clin Epidemiol* 2004;57(12):1223-31.
85. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc* 1984;79(387):516-24.
86. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41-55.
87. Yue LQ. Statistical and regulatory issues with the application of propensity score analysis to nonrandomized medical device clinical studies. *J Biopharm Stat* 2007;17(1):1-13.
88. Labarère J, Bosson JL, Francois P, Fine MJ. L'analyse par score de propension. Exemple

d'application à une étude observationnelle sur la prophylaxie de la maladie thromboembolique veineuse. *Rev Med Interne* 2008;29(3):255-8.

89. Stukel TA, Fisher ES, Wennberg DE, Alter DA, Gottlieb DJ, Vermeulen MJ. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA* 2007;297(3):278-85.
90. Rosenbaum PR. Discussing hidden bias in observational studies. *Ann Intern Med* 1991;115(11):901-5.
91. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000;29(4):722-9.
92. Chen Y, Briesacher BA. Use of instrumental variable in prescription drug research with observational data: a systematic review. *J Clin Epidemiol* 2011;64(6):687-700.
93. Schmoor C, Gall C, Stampf S, Graf E. Correction of confounding bias in non-randomized studies by appropriate weighting. *Biom J* 2011;53(2):369-87.
94. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. *Pharmacoepidemiol Drug Saf* 2010;19(6):537-54.
95. Rassen JA, Brookhart MA, Glynn RJ, Mittleman MA, Schneeweiss S. Instrumental variables I: instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships. *J Clin Epidemiol* 2009;62(12):1226-32.
96. Baser O. Too much ado about instrumental variable approach: is the cure worse than the disease? *Value Health* 2009;12(8):1201-9.

Appendix 1: Glossary

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:

The results obtained are unbiased for the population studied; methodology and the existence of bias determine 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 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 “dropouts” 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 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.

Credibility interval (or Bayesian confidence 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, namely 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. In order to fully preserve the huge benefit of randomisation, all randomised participants should be included in the analysis, exclusively in the group to which they were allocated. These two conditions define an “intention-to-treat” analysis, which is widely recommended as the preferred analysis strategy. Intention-to-treat analysis corresponds to analysing the groups exactly as randomised.

Strict intention-to-treat analysis is often hard to achieve for two main reasons: missing outcome measures for some participants and non-adherence to the trial protocol.

Per-protocol analysis (PP): Involves analysing patients according to the treatment they received rather than their randomisation group. Per-protocol analysis is a comparison of treatment groups that includes only those patients who completed the treatment originally allocated.

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.

Appendix 2: Summary tables

Table 5: Analysis strategies in observational studies

Type of study	Principle	Advantages	Disadvantages
Propensity score	Use of a score summarising the characteristics of the covariates observed	Can reduce confounding	Only takes observed variables into account
Instrumental variable	Use of a variable strongly associated with the treatment indication, but not related to risk factors for the occurrence of an event and not affecting the endpoint	<ul style="list-style-type: none"> - Takes measured and unmeasured confounding factors into account - Can estimate the efficacy of a treatment even in the presence of unmeasured risk factors 	<ul style="list-style-type: none"> - Limited use in practice - Limited possibility of generalisation - Instrumental variable difficult to find - Incorrect estimation if study population is small

Table 6: Specific analysis designs for randomised trials

Type of analysis	Principle	Advantages	Disadvantages
Sequential trial	Interim analysis allows a trial to be stopped early. The trial results for patients included are analysed before new patients are randomised.	<ul style="list-style-type: none"> - does not inflate alpha risk - power conserved - reduces the sample size - avoids ethical problems with continuing the trial if the difference between treatments is greater than expected 	<ul style="list-style-type: none"> - only 2 treatments can be compared - problems evaluating secondary endpoints or adverse effects, with a risk of insufficient power - boundary set <i>a priori</i> and inflexible - overestimation of treatment effect - trials with small sample size may not be representative
Adaptive trial	Interim analysis allows a trial to be stopped early. The trial's protocol may be changed on the basis of results observed in the interim analysis.	<ul style="list-style-type: none"> - does not inflate alpha risk - preserves the trial's power - reduces the sample size - maximises the total number of patients treated correctly by changing the trial's procedure based on the results of interim analyses: more ethical because more patients correctly treated - highly flexible 	<ul style="list-style-type: none"> - practical difficulties - questionable internal validity
Bayesian methods	Use of existing data integrated with data collected during the trial.	<ul style="list-style-type: none"> - More flexible - Reduces the sample size - Shorter duration - Uses all available prior data 	<ul style="list-style-type: none"> - The result may depend far more on the existing data considered than on the data obtained - No statistical tests - No confidence interval (but an equivalent: credibility interval)
Maximum bias analysis	When the failure of the tested treatment results in the other evaluated procedure being used with a favourable outcome, the patient will be considered as a failure in his randomisation group.	<ul style="list-style-type: none"> - Avoids wrongly concluding success when this is actually due to the salvage therapy 	Loss of contrast: reduces differences observed between the groups

Table 7: Types of non-conventional randomised trial

Type of trial	Principle	Advantages	Disadvantages
Expertise-based randomised controlled trial	Patients randomised to a clinician who will perform the procedure in which s/he is an "expert".	<ul style="list-style-type: none"> - Reduced risk of differential deviation from the protocol - Limited differential follow-up bias - Improved internal validity - Reduced risk of differential bias related to experience - Better acceptability - Better feasibility (shorter time to starting the trial) - Fewer ethical problems 	<ul style="list-style-type: none"> - If there is a difference, is this related to the techniques or the surgeon using the techniques? - Longer time to inclusion - Waiting list
Zelen's design	Randomisation before consent.	<ul style="list-style-type: none"> - Fewer withdrawals of consent in the control group - Improved feasibility of recruitment - No damage to doctor-patient relationship 	<ul style="list-style-type: none"> - Selection bias - Ethical problems - Blinding impossible - Diluted treatment effect - Not usable in placebo-controlled trials
Tracker trial	Changes during the trial are authorised, recorded and taken into account in the statistical analysis. Changes to the study's protocol are also authorised.	<ul style="list-style-type: none"> - Early assessment - Respects the principle of equipoise - Maximises data collected 	<ul style="list-style-type: none"> - Very difficult in practice in terms of analysis and logistics - Problems with different degrees of operator experience
Cluster randomised trials	Groups of individuals randomised rather than direct randomisation of individuals.	<ul style="list-style-type: none"> - Prevents contamination bias - Simpler from a logistics point of view 	<ul style="list-style-type: none"> - Comparability of groups not systematically guaranteed - Differential recruitment bias - Increases the sample size - Ethical problems with consent
Crossover trials	Each subject receives the different treatments alternately after a washout period. The order of administration is chosen randomly.	<ul style="list-style-type: none"> - Allows increased power with the same number of subjects - Better precision 	<ul style="list-style-type: none"> - Requires a stable disease - Treatments must not have any permanent effect

Table 8: Procedures for increasing patient recruitment

Type of trial	Principle	Advantages	Disadvantages
Zelen's design	Randomisation before consent.	<ul style="list-style-type: none"> - Fewer withdrawals of consent in the control group - Improved feasibility of recruitment - No damage to doctor-patient relationship 	<ul style="list-style-type: none"> - Selection bias - Ethical problems - Blinding impossible - Diluted treatment effect - Not usable in placebo-controlled trials
Comprehensive cohort study	<p>All eligible patients are invited to take part in the study.</p> <p>If they agree to randomisation, they are randomised to receive one of the two procedures.</p> <p>If they refuse to be randomised, they will be asked which procedure they would prefer.</p>	<ul style="list-style-type: none"> - Can increase recruitment of subjects - Information on treatment acceptability can be collected - Can limit the lack of representativeness observed in clinical trials (highly selective population which may be far from that seen in routine practice) 	<p>The arm corresponding to patient preference is equivalent to an observational study, and is thus subject to confounding, so the comparability of the groups is no longer guaranteed.</p>



All HAS publications can be downloaded from
www.has-sante.fr

n° ISBN : 978-2-11-139070-6 (Édition originale : n° ISBN : 978-2-11-138103-2)