Specific methodological features of clinical evaluation of a connected medical device (CMD)

Preparatory report of the guide to specific features of clinical evaluation in view of its application for reimbursement

"CNEDiMTS approval date:" 29 January 2019
Contents

Introduction ........................................................................................................................................... 6

1. Scope of medical devices covered ................................................................................................. 7
   1.1 Products covered ......................................................................................................................... 7
   1.2 Products not covered .................................................................................................................. 8

2. Method ............................................................................................................................................... 9
   2.1 Documentary search and article selection criteria ...................................................................... 9
   2.2 Search results ............................................................................................................................. 10

3. Evaluation – Data analysis .............................................................................................................. 12
   3.1 Data provided by health technology assessment agencies in other countries ......................... 12
   3.2 Data obtained from applications submitted to the CNEDIMTS ............................................. 14
   3.3 Data from clinical trial databases ............................................................................................. 19
   3.4 Data from the literature .............................................................................................................. 21

4. General conclusion and opinion of the working group ................................................................. 42

5. Appendices .................................................................................................................................. 44
   Annexe 1. Work method .................................................................................................................. 45
   Annexe 2. Documentary search ....................................................................................................... 46
   Annexe 3. Questionnaire sent to stakeholders ............................................................................. 49
   Annexe 4. Review grid ..................................................................................................................... 57
   Annexe 5. List of guide review contributions ................................................................................... 59
   Annexe 6. List of tables .................................................................................................................... 60
   Annexe 7. List of figures .................................................................................................................... 61

References ............................................................................................................................................. 62
Participants – Working Group ............................................................................................................... 64
Fact sheet .............................................................................................................................................. 65
### Abbreviations and acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANSM</td>
<td>Agence nationale de sécurité du médicament et des produits de santé (French National Agency for Medicines and Health Products Safety)</td>
</tr>
<tr>
<td>CNAMTS</td>
<td>Caisse nationale de l’assurance maladie des travailleurs salariés (National health insurance fund for salaried workers)</td>
</tr>
<tr>
<td>CNEDiMTS</td>
<td>Commission nationale d’évaluation des dispositifs médicaux et des technologies de santé (Medical Device and Health Technology Evaluation Committee)</td>
</tr>
<tr>
<td>CSS</td>
<td>Code de la sécurité sociale (French social security code)</td>
</tr>
<tr>
<td>DGOS</td>
<td>Direction générale de l’offre de soins (Directorate General of Health Care Provision)</td>
</tr>
<tr>
<td>DGS</td>
<td>Direction générale de la santé (Directorate-General for Health)</td>
</tr>
<tr>
<td>MD</td>
<td>Medical Device</td>
</tr>
<tr>
<td>CMD</td>
<td>Connected Medical Device</td>
</tr>
<tr>
<td>DSS</td>
<td>Direction de la sécurité sociale (French social security division)</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised controlled trial</td>
</tr>
<tr>
<td>EMA</td>
<td>European Medicines Agency</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>FEDEPSAD</td>
<td>Fédération des prestataires de santé à domicile (Federation of home healthcare providers)</td>
</tr>
<tr>
<td>HAS</td>
<td>Haute Autorité de santé (French National Authority for Health)</td>
</tr>
<tr>
<td>HTA</td>
<td>Health Technology Assessment</td>
</tr>
<tr>
<td>INAHTA</td>
<td>International Network of Agencies for Health Technology Assessment</td>
</tr>
<tr>
<td>INCa</td>
<td>Institut National du Cancer (French National Cancer Institute)</td>
</tr>
<tr>
<td>LPPR</td>
<td>Liste des produits et prestations remboursables (List of products and services qualifying for reimbursement)</td>
</tr>
<tr>
<td>MOST</td>
<td>Multiphase Optimisation Strategy</td>
</tr>
<tr>
<td>MSA</td>
<td>Mutualité sociale agricole (Agricultural Social Mutual Fund)</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Care Excellence</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>PREMs</td>
<td>Patient Reported Experience Measures</td>
</tr>
<tr>
<td>PROMs</td>
<td>Patient Reported Outcome Measures</td>
</tr>
<tr>
<td>RSI</td>
<td>Régime social des indépendants (Health and retirement scheme for independent workers)</td>
</tr>
<tr>
<td>SED</td>
<td>Service évaluation des dispositifs (HAS medical device assessment department)</td>
</tr>
<tr>
<td>SIDIV</td>
<td>Syndicat de l'industrie du diagnostic in vitro (In vitro diagnostic industry association)</td>
</tr>
<tr>
<td>SNITEM</td>
<td>Syndicat national de l'industrie des technologies médicales (National medical technology industry association)</td>
</tr>
<tr>
<td>UNPDM</td>
<td>Union nationale des prestataires de dispositifs médicaux (National union of medical device providers)</td>
</tr>
<tr>
<td>UPSADI</td>
<td>Union des prestataires de santé à domicile indépendants (Union of independent home healthcare providers)</td>
</tr>
<tr>
<td>SMART</td>
<td>Sequential Multiple Assignment Randomised Trials</td>
</tr>
<tr>
<td>SNADOM</td>
<td>Syndicat des associations d’assistance à domicile (Federation of home help associations)</td>
</tr>
<tr>
<td>YHEC</td>
<td>York Health Economic Consortium</td>
</tr>
</tbody>
</table>
Glossary

Algorithm
“Description of a finite and unambiguous series of steps or instructions for producing a result on the basis of input elements.” (1)
A learning algorithm is able to autonomously develop the parameters of the instructions of which it is made up over time, according to the results previously received. It is the opposite of conventional algorithms which do not have the ability to self-modify.

Supervised machine learning
“The algorithm learns input data qualified by humans and thus defines rules on the basis of examples which are as many validated cases.” (1)
Supervised learning takes place in two steps. During the first, the algorithm learns, on the basis of solved cases, to adjust its parameters in order to enhance its predictive performances (e.g. case classification). The second step consists of verifying that the resulting algorithm is generalisable by applying it to a set of new cases, the result for these cases having to be validated by an expert. We thus check that the algorithm has learned properly, if the learning was biased (over learning) or, conversely, if it requires more examples (under learning).

Unsupervised machine learning
“The algorithm learns from raw data and creates its own classification which is free to develop into any final state where a pattern or element is presented to it. Practice which requires instructors to teach the machine how to learn.” (1)
In the case of unsupervised learning, the data provided to the algorithm are unsolved cases (we do not know the conclusion). The algorithm relies on functions of similarity, differences etc. between the cases to pool them into groups. The expert then verifies later on that the resulting groups effectively classify the data (by comparing the algorithm result and its own decision on new data) in order to decide whether the algorithm is valid and can be used in real life.

Machine learning
“Branch of artificial intelligence, based on methods for learning and automatically capturing new knowledge using computers, used to have them perform tasks without having to be explicitly programmed.” (1)
The main methods are known as supervised or unsupervised. These methods can be combined together. When used in real-life situations, the algorithm can be temporarily fixed, then updated sequentially and regularly by persons in charge of its development, according to predetermined version management. The algorithm can also not be fixed and can update dynamically when used and when processing new data.

Big data
“Denotes the convergence between, on one hand, huge volumes of data beyond the processing ability of current digital technologies and, on the other, new technologies suitable for processing such data, or extracting such data by identifying unexpected data correlations.” (1)
The development of digital technologies has resulted in the generation of huge volumes of data in the context of growth of the Internet. The parallel development of new approaches to data processing and scaling-up of computer power has given rise to an international branch concerned with the mass processing of heterogeneous data sets known as “big data”.

Regression testing
“The testing required to determine that a change to a system component has not adversely affected functionality, reliability or performance and has not introduced additional defects.”

Introduction

The Medical Device and Health Technology Evaluation Committee (CNEDiMTS) is the HAS committee which evaluates, in particular, medical devices (MD) and other health products in view of their reimbursement by the French health insurance scheme. It plays an advisory role to health authorities recommending the reimbursement of the MDs or not (inclusion on the list of products and services qualifying for reimbursement – LPPR), helping to determine the conditions for their proper use and their role in the therapeutic, diagnostic or prevention strategy. If required, it sets out the conditions allowing optimised technology use in terms of user expertise and environment required².

As for any MD category, a connected medical device (CMD) eligible for the LPPR can be the subject of a request for inclusion under its brand name, submitted by the company responsible for its operation. Such requests must be accompanied by supporting data allowing the assessment of the benefit of the products and its public health benefit, justifying its the argumentation for actual clinical benefit (ACB) and clinical added value (CAV). Indeed, the CNEDiMTS evaluation criteria are regulatory criteria³ which apply regardless of the type of medical device.

However, CMDs can have specific features related to their mode of action, their impact on patients, carers, healthcare professionals or organisation of care. Their clinical development must take these specific features into account.

The aim of this project undertaken in 2018 is to identify any specific methodological features in the clinical evaluation of CMDs, so as to enable companies submitting an application to the CNEDiMTS to include them in their clinical development strategy and prepare for their reimbursement by French health insurance scheme. This project is in line with the HAS work programme⁴.

A “Guide to the specific features of clinical evaluation of a connected medical device (CMD) in view of its application for reimbursement” has been produced. This guide is available to view and download on the HAS website (www.has-sante.fr).

This supplementary report to the guide details the method used to develop the guide.

³ Article R. 165-11 of the French social security code [viewed on 21/12/2018].
⁴ https://www.has-sante.fr/portail/upload/docs/application/pdf/2018-02/programme_de_travail_has_vd.pdf [viewed on 21/12/2018].
1. Scope of medical devices covered

1.1 Products covered

The connected medical devices area is vast, however, in light of the CNEDiMTS’ missions, the medical devices (MDs) it evaluates only represent a small part of connected medical devices (CMDs).

The scope of this work fully underpins the scope of the MDs evaluated by the CNEDiMTS, that is to say those with CE marking and candidates for individual funding from the French health insurance scheme (in addition to existing procedures or healthcare packages).

CMDs evaluated by the CNEDiMTS meet the following four criteria.

1. They are intended for use for medical purposes\(^5\), their end-use implying they are CE-marked.
2. They are for individual use (implanted or used by the patient themselves).
3. They have a telecommunication function.
4. The company has submitted an application for reimbursement by national solidarity.

It is the fourth condition that triggers evaluation of a CMD by the CNEDiMTS: the company takes the initiative to register a new technology on the list of products and services qualifying for reimbursement (LPPR) (article R. 165-7 of the CSS).

To illustrate this scope, some examples of CMDs likely to qualify for individual reimbursement by the French health insurance scheme are given below; MDs:

- used for medical telemonitoring purposes (e.g. an implantable cardiac defibrillator and remote medical monitoring);
- prompting action from the patient for self-treatment or self-monitoring purposes (e.g. nerve stimulator to treat pain connected to a smartphone application allowing the patient to manage their treatment themselves);

- producing or receiving information in view of treatment optimisation (e.g., an insulin pump combined with a sensor for the continuous measurement of interstitial glucose using the patient’s electronic diary to optimise their treatment).

1.2 Products not covered

Numerous products used in health, although they are connected, are not covered in this work:

- applications and connected objects that are not CE-marked medical devices (e.g. software or apps used to increase physical activity by calculating the number of steps per day);
- medical devices which are not for individual use and which are not subject to individual funding (e.g. connected balances, thermometers or blood pressure monitors used in hospitals for more than one patient);
- medical devices used exclusively by a healthcare professional or between healthcare professionals (e.g., professional decision aid tools, prescription or dispensing aid software, teleconsultation software, diagnostic or therapeutic decision aid medical imaging devices, etc.);
- software for general uses, even when used in a healthcare environment (e.g. administrative management software).
2. Method

The method used is based on:

- a questionnaire-based survey on methods and requirements used in other countries;
- a review of the applications submitted to the CNEDiMTS in respect of this type of technology;
- a review of clinical trial databases;
- a critical review of the scientific literature data;
- the opinion of experts included in a working group.

The working group is made up of HAS Departments and experts with no major conflicts of interest (as per the HAS “Guidance on declarations of interest and management of conflicts of interest”) chosen in view of their experience in medical device assessment and their eHealth or methodological expertise.

Prior to the working group, for the framing of this project and review of the resulting guide, stakeholders were contacted, including:

- representatives of manufacturers and distributors;
- “Fédération des Spécialités Médicales” and learned societies specialised in eHealth in the medical device sector;
- service providers and distributors of equipment for home use, specialised in eHealth in the medical device sector;
- patient associations concerned by ongoing trials or with assessment platforms in place;
- focus groups or experts specialised in eHealth.

These studies were reviewed and approved by the CNEDiMTS and subsequently submitted to the HAS College for information purposes. The Committee may propose supplementary amendments to the studies prepared by the expert group.

The details of this method are described in Appendix 1.

2.1 Documentary search and article selection criteria

2.1.1 Documentary search

The documentary search was restricted to publications in English and French. It covered the period from January 2013 to October 2018. Monitoring was conducted until December 2018 on the websites consulted.

Health technology assessment reports, guidelines, consensus conferences, guidance documents, meta-analyses, systematic reviews, and clinical trials were searched. The search was conducted by querying multiple bibliographic databases, including Medline and various websites such as the Cochrane Library, websites publishing guidelines, websites publishing health technology assessment reports, and websites of learned societies with expertise in the field studied.

The search strategy and the list of queried sources are detailed in Appendix 2.

This automated search was supplemented by a manual search conducted based on the references cited in the selected publications. The experts' bibliography and industry data were also taken into consideration as sources of data. If the studies provided by these sources met the selection criteria defined for the documentary search, they were included in the bibliography.
2.1.2 Selection criteria

In order to be selected, the data were required to meet the following criteria.

► Study type
Guidelines, guidance documents or health technology assessment reports, meta-analyses, systematic reviews, clinical trials.

► Primary objective
Evaluate connected medical devices. The guidelines, systematic reviews, guidance documents or health technology assessment reports were required to identify clinical trial methodologies for connected medical devices. The clinical trials or meta-analyses were required to use a specific clinical evaluation methodology other than that usually described in evaluation applications for the various categories of medical devices submitted to the CNEDiMTS.

► Topic addressed or devices used
Connected medical devices within the scope of this project (section 1.1).

► Study population
Patients using a connected medical device.

► Endpoints
All endpoints were taken into account, provided that they made it possible to assess the individual benefit and that they were specific to connected medical devices.

► Exclusion criteria
General articles, narrative articles such as editorials or opinion pieces, in vitro, biomechanical, histological studies, conducted on animals, studies evaluating CMDs with no specific clinical features (in terms of design, endpoints, duration, sample size, etc.), duplicate studies.

► Selection period
From 2013 to 2018.

2.2 Search results

In total, 1,163 bibliographic references were identified from the systematic search run on Medline and 22 from websites. Of these references, 75 were reviewed and 10 were retained.

The selection process is illustrated in the figure below. These references are broken down as follows:
- 3 methodological guidance documents;
- 1 literature review analysing the ClinicalTrials.gov databases;
- 6 supplementary data sources addressing the issue of the clinical evaluation of digital technologies and proposing methodological adaptations.
Article selection flow chart

1,185 references obtained from the initial documentary search (1,163 from the Medline search and 22 from websites)

Review of the full-text version of 75 references

Addition of 5 references

10 references retained:
- 3 methodological guidance documents;
- 1 literature review analysing the ClinicalTrials.gov clinical trial databases;
- 6 supplementary data sources addressing the issue of the clinical evaluation of digital technologies and proposing methodological adaptations.
3. Evaluation – Data analysis

The clinical data, obtained from the literature or provided by stakeholders, were selected according to the criteria set out in section 2.1.2, and concern the specific methodological features of clinical evaluation of connected medical devices.

3.1 Data provided by health technology assessment agencies in other countries

A survey was conducted in the INAHTA network – International Network of Agencies for Health Technology Assessment (49 assessment agencies – 30 countries).

The data provided by other international health technology assessment agencies were reviewed. The items covered by the survey were as follows:

- scope of assessment and types of CMDs covered;
- number of CMDs assessed;
- identification of any specific features during the clinical evaluation of CMDs, compared to the evaluation of other medical devices;
- existence of guidelines in relation to CMDs (recommendations of a specific design or guidance document on evaluation methods);
- expected or planned adaptation of CMD evaluation methods.

The responses from the 12 agencies located in 8 countries that responded are detailed in Table 1.

Table 1. Data provided by INAHTA network agencies

<table>
<thead>
<tr>
<th>Agencies</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>JAZMP</strong> (Slovenia)</td>
<td>No agency is responsible for the assessment of medical devices. No feedback in terms of clinical evaluation of CMDs.</td>
</tr>
<tr>
<td>Javna agencija Republike Slovenije za zdravila in medicinske pripomočke</td>
<td></td>
</tr>
<tr>
<td><strong>RedETS</strong> (Spain)</td>
<td>In Spain, medical devices are assessed by the HTA agency network (RedETS) that coordinates 8 regional HTA agencies (AQUAS, SESCS, AETS, OSTEBA, AVALIA-T, IACS, AETS, UETS-M). The scope of their assessment includes all types of MD (used by the patient or among healthcare professionals). Three documents published in Spanish only were provided. They cover “Information and communication technologies in mental health (telepsychiatry)”, “Teleneurology vs. face-to-face consultation. Assessment of its effectiveness and of its cost-effectiveness”, “Telediagnosis design, assessment and implementation.” No specific features in terms of clinical evaluation are identified and no methodological adaptations are envisaged.</td>
</tr>
<tr>
<td>Red Española de Agencias de Evaluación de Tecnologías Sanitarias y Prestaciones del Sistema Nacional de Salud</td>
<td></td>
</tr>
<tr>
<td>SESCS (Spain)</td>
<td>No data relating to the clinical evaluation of CMD are available. No specific features in terms of clinical evaluation are identified, clinical evaluation is similar regardless of the medical device. No guidelines are available.</td>
</tr>
<tr>
<td>Servicio Evaluación del Servicio Canario de la Salud (HTA unit of the Regional Government of the Canary Islands)</td>
<td></td>
</tr>
<tr>
<td><strong>AETS</strong> (Spain)</td>
<td>Some MDs have been assessed, such as ventricular assist devices or implantable cardiac defibrillators. No specific features in terms of clinical evaluation are identified and no methodological adaptations are envisaged. No guidelines are available.</td>
</tr>
<tr>
<td>Agencia de Evaluación de Tecnologías Sanitarias</td>
<td></td>
</tr>
<tr>
<td><strong>AQuAS</strong> (Spain)</td>
<td>Connected MDs may be assessed at the level of regional HTA agencies or at a national level. There is no specific agency assessing connected MDs. The scope of the assessment includes all types of MD (used by the patient or among healthcare professionals). No CMD assessments are provided. No guidelines are available.</td>
</tr>
<tr>
<td>Agència de Qualitat i Avaluació Sanitàries de Catalunya</td>
<td></td>
</tr>
<tr>
<td><strong>ASSR</strong> (Italy)</td>
<td>The G-BA (Federal Joint Committee) is responsible for assessing new technologies. The IQWIG is commissioned by the G-BA to assess their benefits.</td>
</tr>
<tr>
<td>Agenzia Sanitaria e Sociale Regionale dell’Emilia-Romagna (Regional Agency for Health and Social Care),</td>
<td></td>
</tr>
<tr>
<td><strong>IQWIG</strong> (Germany)</td>
<td></td>
</tr>
<tr>
<td>The G-BA (Federal Joint Committee) is responsible for assessing new technologies. The IQWIG is commissioned by the G-BA to assess their benefits.</td>
<td></td>
</tr>
</tbody>
</table>
Specific methodological features of clinical evaluation of a connected medical device

<table>
<thead>
<tr>
<th>Agencies</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institute for Quality and Efficiency in Health Care</td>
<td>The G-BA assesses new technologies in specific indications — the broad and non-specific use of devices used in a telehealth context does not fall within the remit of the G-BA and IQWIG. A preliminary report demonstrating a lack of benefit of tele-monitoring is cited. It relates to telemonitoring of active implantable cardiac devices in patients with ventricular arrhythmia and heart failure. No specific features in terms of clinical evaluation are identified, standard methodology applies regardless of the medical device. There is no need for a specific methodological adjustment. No guidelines are available.</td>
</tr>
<tr>
<td>SV (Austria) Hauptverband der österreichischen Sozialversicherungsträger</td>
<td>No agency is specifically responsible for the assessment of CMDs. No feedback in terms of clinical evaluation of CMDs. No guidelines are available.</td>
</tr>
<tr>
<td>GOeG (Austria) Gesundheit Österreich GmbH</td>
<td>There is no specific document on CMDs. Assessment methods are identical to those for other MDs. No adaptation of assessment methods is envisaged.</td>
</tr>
<tr>
<td>LBI-HTA (Austria) Ludwig Boltzmann Institute for Health Technology Assessment</td>
<td>LBI-HTA has assessed technologies within a telecardiology and tele-dermatology context. An assessment of telemedicine is envisaged by 2020. There would appear to be a need to issue guidelines on the clinical evaluation of CMDs. Nevertheless, the assessment methods in respect of apps/services in Austria are the same as those for other MDs (PICO), with a focus on organisational data. No guidelines are available. No adaptation of assessment methods is envisaged.</td>
</tr>
<tr>
<td>DEFACTUM (Denmark) Social &amp; Health Services and Labour Market</td>
<td>No specific agency is responsible for CMDs. DEFACTUM assesses MDs, working together with various professional organisations. The scope of the assessment includes all types of MD (used by the patient or among healthcare professionals). The specific features associated with CMD assessment stem from the fact that the assessment covers a chain of technologies which are depending on an organisation to be set up (and not a single connected device). The implementation of connected health technology often varies, further complicating the assessment. The context and the organisation set up are important factors to be included in the assessment. As such, the components in the chain can influence each other differently depending on the patient group. No guidelines are available. No adaptation of assessment methods is envisaged.</td>
</tr>
<tr>
<td>NCPHA (Bulgaria) National Center of Public Health and Analyses</td>
<td>No specific agency is responsible for CMDs. The specific features associated with CMD assessment concern their indirect impact on patient behaviour. No guidelines are available. No adaptation of assessment methods is envisaged.</td>
</tr>
<tr>
<td>SFOPH (Switzerland) Swiss Federal Office of Public Health</td>
<td>In Switzerland, CMDs for personal use for medical purposes and also for use by healthcare professionals are assessed, with a view to a health insurance reimbursement decision. The assessments are conducted by the SFOPH and subsequently approved by permanent evaluation committees that make recommendations to the decision-making body (Ministry of Health). To date, two devices have been assessed: one for monitoring cardiac arrhythmia (telemedicine) and one for diabetic self-monitoring of glucose levels. The same methodology is applied for assessments of other technologies and there are no plans to issue recommendations of guidelines.</td>
</tr>
</tbody>
</table>

Conclusion
Based on the responses from the other assessment agencies contacted, it is not possible to identify any specific methodological features for the clinical evaluation of connected medical devices. This survey does not indicate any existing specific guidelines or prospects of envisaged methodological adaptation. CMD assessment methods would seem to be identical to those for other medical devices. The survey responses highlight the complexity of the assessment of a CMD, due to its organisational impact and its impact on patients. These responses also indicate that CMD assessment is required to account for a set of components which are continually interacting and interdependent of an organisation to be set up, each element of the system being capable of influencing others differently depending on the patient group.
3.2 Data obtained from applications submitted to the CNEDiMTS

A number of categories of connected medical devices have previously been assessed by the CNEDiMTS, following the submission of applications by companies. These CMD categories are detailed in Table 2. Note that, to date, no applications have cited CMDs involving AI-based data processing.

Table 2. CMD categories assessed by the CNEDiMTS, following the submission of applications by companies

<table>
<thead>
<tr>
<th>Field of application</th>
<th>CMD categories</th>
<th>General indications/Comments</th>
</tr>
</thead>
</table>
| Locomotor system     | Power propulsion assist devices for manual wheelchair | **Indications:**
|                      |                | Patients with the cognitive abilities allowing them to operate the power propulsion assist device and who, despite being capable of propelling a manual wheelchair themselves, require an electrical power propulsion assist device intermittently or definitively for medical reasons.
|                      |                | **Specificities:**
|                      |                | - operated using a “Pushtracker” wristband attached to the user’s wrist that communicates with the motor unit via Bluetooth when the power assist device is activated/deactivated;
|                      |                | - all data can be viewed via an app (average number of pushes per day, number of km, etc.). |
| Audiology            | Bone-anchored devices (bone-anchored hearing aids) | **Indications:**
|                      |                | - cases of conductive hearing loss or combined hearing loss unsuitable for middle ear surgery and for whom conventional air or bone conduction hearing aids are ineffective or unsuitable;
|                      |                | - cases of at least severe unilateral sensineural hearing loss.
|                      |                | **Mode of action:**
|                      |                | - the external speech processor converts the acoustic pressure into a force of variable intensity;
|                      |                | - this force is retransmitted from the transcutaneous abutment to the bone-anchored implant;
|                      |                | - the implant generates an elastic deformation of the bone cortex according to the frequency;
|                      |                | - this vibration is reverberated on the temporal bone where the labyrinth capsule containing the inner ear is situated.
|                      |                | Data logging feature. |
| Middle ear implant system | **Indications:**
|                      |                | Cases of unilateral or bilateral conductive or combined hearing loss, in children and adults, for whom the following have failed or are unsuitable:
|                      |                | - middle ear surgery;
|                      |                | - and conventional air or bone conduction hearing aids;
|                      |                | - and bone-anchored hearing aids.
|                      |                | **Mode of action:**
|                      |                | The principle is based on the direct conduction of sounds to a vibratory structure of the middle ear via an electromagnetic transducer, the FMT.
|                      |                | The audio processor captures the sounds via a microphone, processes the signal, and sends the information transcutaneously to the implanted component. The VORP internal receiver transfers the electrical signal via the conductive link, to the output transducer, the FMT. The FMT converts the electrical signal into mechanical vibrations transmitted to the middle ear. These vibrations are then transmitted to the inner ear. |
| Brainstem implants   | **Indications:**
|                      |                | Severe to profound bilateral sensineural hearing loss (perception hearing loss), following failed or ineffective conventional hearing aid use.
|                      |                | The indications are restricted to circumstances in which the cause of the hearing loss does not allow cochlear implantation:
|                      |                | - either due to tumour excision (bilateral vestibular schwannoma) in the context of neurofibromatosis type 2;
### Field of application | CMD categories | General indications/Comments
--- | --- | ---
Cochlear implants | Indications: Severe to profound bilateral neurosensory hearing loss (perception hearing loss), following failed or ineffective conventional hearing aid use. Bilateral cochlear implantation in children is also indicated for cases of profound bilateral sensineural hearing loss (perception hearing loss), following failed or ineffective conventional hearing aid use. Mode of action: The sound processor picks up sounds via a multidirectional microphone. This signal is processed, digitised, encoded, and transmitted at the antenna to the internal part of the implant by inductive coupling. This modulated signal passes through the electrodes of the cochlear or brainstem implant, stimulating the fibres of the auditory nerve or cochlear nucleus, respectively. The action potentials generated are interpreted as sounds by the brain.

Cardiology /cardiac rhythm study | Remote monitoring system (of implantable cardiac defibrillator, pacemaker or of heart failure patients) Implantable cardioverter defibrillator, Implantable cardiac pacemaker | Indications: For remote monitoring systems: of pacemakers/ cardioverter defibrillators in the indications below. For single-chamber implantable cardiac pacemakers: - rate-responsive single-chamber atrial pacemaker: sinus node dysfunction with chronotropic incompetence, if atrioventricular conduction is normal in the absence of progressive heart disease; - rate-responsive single-chamber ventricular pacemaker: second- or third-degree atrioventricular block with chronotropic incompetence; if the auricle is not suitable for predominant pacing or not detectable. For dual-chamber implantable cardiac pacemakers: Second- or third-degree atrioventricular block requiring continuous or intermittent ventricular pacing (if atrioventricular conduction is preserved): - for patients with chronotropic incompetence, where it is possible to maintain physical activity, and if the auricle is suitable for predominant pacing; - sinus node dysfunction with chronotropic incompetence associated either with atrioventricular conduction abnormalities, or with atrial rhythm disorder. For implantable cardioverter defibrillators: - ventricular arrhythmia inducing haemodynamic instability (sudden cardiac arrest survivor, poor VT tolerance) and life expectancy > 1 year with good functional status; - patients with NYHA II or III symptomatic heart failure, left ventricular ejection fraction (LVEF) ≤ 35%, despite optimal pharmacological treatment ≥ 3 months and with a life expectancy > 1 year with good functional status, of ischaemic origin and > 40 days after the acute myocardial infarction phase or of non-ischaemic origin; - genetic disease involving a high risk of sudden cardiac arrest by VT/ventricular fibrillation without any other known effective treatment. For implantable cardiac resynchronisation systems (“triple-chamber” pacemakers or defibrillators): - NYHA II, III and IV class ambulatory chronic symptomatic heart failure patients, receiving optimal medical treatment, with ejection fraction ≤ 35%, in sinus rhythm: • with QRS duration > 150 ms, • with QRS duration between 120 and 150 ms and with left branch block;
### Field of application | CMD categories | General indications/Comments
--- | --- | ---
| | | - permanent atrial fibrillation patients with chronic heart failure, having a QRS ≥ 120 ms and LVEF ≤ 35%, classified as NYHA III and IV ambulatory despite optimal medical treatment, provided that it is possible to obtain approximately 100% biventricular pacing.

**Portable external cardiac defibrillator system**

**Indications:**
- After explantation of an implantable defibrillation system due to infection from the housing or electrodes, until reinplantation (recovery from infection).
- Pending heart transplantation. The indication has to be reviewed every 3 months (assessment of benefit/risk ratio and of compliance).
- After myocardial revascularisation if the left ventricular ejection fraction (LVEF) is less than 30%, until LVEF review and discussion of the indication of an implantable automatic defibrillator at month 3.
- During acute myocardial infarction if the LVEF is less than 30% after the first 48 hours, until LVEF review and discussion of the indication of a permanent implantable automatic defibrillator after the first month.

**Insertable cardiac monitors**

**Indications:**
- Aetiological diagnosis of syncope.
- Aetiological diagnosis of cerebral ischaemic attack.

**Continuous interstitial glucose monitoring (CGM) system coupled with an insulin pump**

**Indications:**
- type 1 diabetes patients (adults and children) with poor prior glycaemic control (HbA1c level ≥ 8%) despite well-conducted intensive insulin therapy by means of continuous subcutaneous insulin infusion (external pump) for more than 6 months and glycaemic self-monitoring multiple times a day (≥ 4/d);
- type 1 diabetes patients (adults and children) having presented with severe hypoglycaemic episodes resulting in emergency medical treatment, in the previous 12 months, despite well-conducted intensive insulin therapy by means of continuous subcutaneous insulin infusion (external pump) for more than 6 months and glycaemic self-monitoring multiple times a day (≥ 4/d).

These systems are reserved for patients who have received therapeutic education, as well as specific training on use of the continuous glucose monitoring system.

**Mode of action:**
The CGM + pump system enables insulin administration (pump function) and continuous interstitial glucose monitoring (glucose monitor function).

Coupling both functions enables insulin infusion to be discontinued temporarily in the event of established or predicted hypoglycaemia, without patient intervention.

The pre-hypo stop function estimates the predicted risk of exceeding (in the next 30 minutes) a predetermined estimated blood glucose threshold.

**Blood glucose meters offering blood glucose data download capability via Bluetooth**

**Indications:**
- for type 1 diabetes patients: at least 4 self-monitored blood glucose readings/day;
- for women with gestational diabetes: at least 4 self-monitored blood glucose readings/day;
- for some type 2 diabetics defined below:
  - insulin-treated patients;
  - patients treated by oral glucose lowering drugs (sulfonylureas or glinides, alone or in combination with other antidiabetic drugs) in order to detect or confirm hypoglycaemia and adapt the dosage of these medicinal products if required. However, cover is restricted to 200 strips per year for these patients.
  - patients for whom it is sought to improve their glycaemic control, where the target is not achieved, as an educational tool providing an indication of the effect of exercise, diet and medication. However, cover is restricted to 200 strips per year for non-insulin-dependent patients.

**Mode of action:**
Blood glucose monitoring from a capillary sample, using an algorithm integrated in the meter.

**Flash glucose self-monitoring system**

**Indications:**
Interstitial glucose monitoring in the treatment of patients with type 1 or type 2 diabetes (adults and children of at least 4 years of age) treated with intensified insulin therapy (via external pump or ≥ 3 injections per day) and carrying out glycaemic self-monitoring multiple times a day (≥ 3/day).
<table>
<thead>
<tr>
<th>Field of application</th>
<th>CMD categories</th>
<th>General indications/Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>This system is reserved for patients who have received therapeutic education, as well as specific training on use of the flash interstitial glucose self-monitoring system. <strong>Mode of action:</strong> Interstitial glucose monitoring. The meter can be used to show trends in terms of blood glucose fluctuations using algorithms.</td>
</tr>
<tr>
<td></td>
<td>Software for basal-bolus insulin regimen and technical service coupled with medical telemonitoring</td>
<td><strong>Indications:</strong> Adult type 1 diabetic patient (diagnosed more than 1 year previously), poorly controlled (HbA1c ≥ 8%) with basal-bolus insulin regimen administered by means of multiple injections or pump (for at least 6 months). This system is reserved for patients who have received specific training on its use. <strong>Mode of action:</strong> This system consists of software coupled with medical telemonitoring and a technical service teaching software use. The purpose of the software is to assist the patient in the daily calculation of long-acting and short-acting insulin doses, based on targets predefined by the prescribing physician. It is available via an app on a mobile device (smartphone or tablet) for the patient, and via a Web portal for healthcare professionals involved in their treatment (prescribing physician and nurse). The short-acting insulin dose calculator recommends a short-acting insulin dose based on the data logged by the patient (blood glucose levels, blood glucose events, carbohydrate consumption and exercise). These data are processed based on algorithmic parameters initially defined by the healthcare professional via the remote processing configuration function.</td>
</tr>
<tr>
<td></td>
<td>Implantable intraperitoneal insulin administration pump system</td>
<td><strong>Indications:</strong> Adult type 1 diabetes patients poorly controlled with subcutaneously administered insulin (including via a pump) and presenting with severe, frequent or unexplained hyperglycaemic and/or hypoglycaemic episodes. <strong>Mode of action:</strong> Implantable pump connected to a catheter and a portable component enabling the physician and patient to communicate with the pump via radiofrequency. The portable component is used to program administration rates, activate bolus doses, and stores the various programming data in memory.</td>
</tr>
<tr>
<td>Neurology</td>
<td>Rechargeable or non-rechargeable spinal cord stimulation systems</td>
<td><strong>Indications:</strong> 1- Chronic neuropathic pain, after failed alternative treatments subsequent to: - chronic radicular pain syndrome persisting for at least 1 year post-surgery; - chronic truncal pain syndrome (due to diabetes, zona, injury or surgery) persisting for at least 1 year; - type I or II complex regional pain syndrome persisting for at least 6 months. 2- Pain of ischaemic origin, after failed alternative treatments subsequent to Buerger’s disease. <strong>Mode of action of algorithm:</strong> - charge or battery indicator; - stimulation programs; - for one range: detection and adaptation of upright/seated position and adaptation of therapy delivery.</td>
</tr>
<tr>
<td></td>
<td>Implantable dorsal root ganglion systems</td>
<td><strong>Indications:</strong> Chronic neuropathic pain, after failed alternative treatments subsequent to type I or II complex regional pain syndrome persisting for at least 6 months. <strong>Mode of action of algorithm:</strong> - charge or battery indicator - stimulation programs</td>
</tr>
<tr>
<td></td>
<td>Implantable vagus nerve stimulation systems</td>
<td><strong>Indications:</strong> Child or adult with established incapacitating and drug-resistant epilepsy (seizure recorded on electroencephalogram – EEG) for whom intracranial surgical treatment is not indicated. Drug-resistant epilepsy is defined as the persistence of seizures after 2 years of suitable treatment, i.e. prior use of sequential monotherapy regimen of at least 2 anti-epileptic medicines and at least one combination of 2 anti-epileptic medicines for a sufficient period of time in order to assess efficacy. <strong>Mode of action of algorithm:</strong> - charge or battery indicator; - stimulation programs; - for an MD: automatic seizure detection algorithm based on heart rate monitoring.</td>
</tr>
</tbody>
</table>
The scope of CMDs falling within the remit of CNEDIMTS assessment is thus very varied, with the data transmission and reception capabilities having different end purposes. The evaluation of these various CMD categories for which a request for inclusion on the list of product and services qualifying for reimbursement has been submitted did not highlight any specific methodological features in the design of the clinical trials used.

The content of applications involving data transmission for medical telemonitoring purposes, for which a CNEDIMTS assessment had been carried out, was specifically reviewed.

- **Main therapeutic fields covered**: cardiology and diabetes

The CMDs consisted of apps installed on smartphones or communication devices enabling telemonitoring via data collection, transmission and analysis.

- **Study types**

The clinical study designs did not differ from those used for the assessment of other technologies. The clinical studies analysed were essentially randomised controlled trials or RCT meta-analyses.

The other types of clinical studies listed were as follows:

- non-randomised comparative prospective study;
- non-comparative observational study;
- non-comparative retrospective study.

- **Endpoints**

The primary endpoint was clinical (validated clinical endpoints) or biological (validated surrogate endpoint, in this instance, it consisted of the HbA1c level).

Other endpoints were assessed:

- quality of life;
- time spent by physicians on patient consultations (conventional or telephone consultation) and time spent by patients on travel or attending hospital;
- number of consultations;
- patient satisfaction;
- number of hospital admissions;
- compliance.
Conclusion
Based on the applications submitted in respect of connected medical devices for which a clinical evaluation has been carried out by the CNEDiMTS, no specific methodological features in the design of the clinical trials conducted have been identified.

3.3 Data from clinical trial databases

The search was run in two stages successively using two different key-words: i) algorithm; ii) mhealth.

► CLINICAL TRIALS database – algorithm

The search was run using the key-word *algorithm* and identified 482 studies.

**Health technologies**
The studies identified involve a wide range of health technologies. Indeed, while some algorithms, such as learning algorithms, require the use of computer systems, others are decision trees displayed on screen or on print-outs.

**Study types**
482 studies were recorded, of which 304 are interventional and 178 are observational. 258 are described as completed.

Of the interventional studies (n = 304), 192 are comparative. The study design is randomised between two or more arms for 176 of these (58%); the 16 other studies compare outcomes (before/after) based on a single patient group. For 31 interventional studies, a cross-over design was used. Sequential randomisation designs (n = 5) are described.

The main conditions or fields covered by these studies are:
- cardiovascular diseases (24%);
- diabetes (16%);
- respiratory medicine (8%);
- oncology (7%).

For diabetes, the studies essentially (96%) relate to the use of algorithms for therapeutic purposes (e.g. insulin administration regulation or selection of potential responders). Conversely, in oncology, the algorithms are mostly (58%) for diagnostic purposes (e.g. image analysis algorithms or diagnostic algorithms based on non-invasive test results).

**Endpoints**
For the comparative studies, the endpoints are essentially clinical or biological, whether systems involved learning algorithms or not.

The non-comparative studies are mostly aimed at validating algorithms; measurements of method specificity and sensitivity are described as endpoints.

► CLINICAL TRIALS database – mhealth

The search run using the key-word *mhealth* (which was associated with the term *mobile health*) identified 415 studies.
**Health technologies**

The health technologies in question are almost exclusively apps installed on smartphones (mobile apps) with or without associated sensors. The descriptive elements of the interventions used are sometimes not featured in the database or else they are not provided in detail. The devices may be used on a one-off basis: capturing a questionnaire or a photo. However, most often, they enable active or passive iterative use by the subject.

**Study types**

415 studies were recorded, including 380 interventional and 35 observational. 151 are described as completed.

Of the interventional studies (n = 380), the study design is randomised between two or more arms in the case of 300 of these studies. For 15 of these studies, a cross-over design was used. It is observed that many designs propose that the control arm also wear the device without having access to the results of the assessed function (hidden measurements). Sequential randomisation designs (n = 9) are described.

Of the other interventional studies (n = 80), 12 are comparative and 68 are not.

The observational studies (n = 35) are essentially studies only involving a single patient group, that using the device.

The main conditions or fields covered by these studies are:

- cardiovascular diseases (13%);
- psychological disorders (13%), with half being associated with addictions, the other being associated with psychiatry, or with anxiety or depression in equal measure;
- diabetes (11%);
- sexually transmissible diseases (10%);
- nutrition (9%);
- oncology (9%).

**Endpoints**

The specific endpoints associated with mobile health include those assessing the apps per se:

- app use frequency;
- rating of app use acceptability or of app quality (e.g. with the Mobile App Rating Scale (MARS), specifically designed for mobile apps);
- time spent on the app;
- app use compliance;
- app user satisfaction.

The endpoints are measured by the system directly, by sensors embedded in the smartphone (e.g. the accelerometer measures the number of times getting up or the number of steps) or associated sensors (e.g. activity sensors or blood glucose or pressure sensors).

The other endpoints may be clinical or biological (impact on HbA1c), or behavioural (number of cigarettes consumed, number of post-exposure preventive treatments taken).

Many trials favour patient self-reporting and use self-questionnaires to rate the level of anxiety, stress, or depression, and also quality of life.

Other endpoints are studied, and concern patient follow-up organisation:

- number of consultations or hospital admissions (emergency or not);
- interval prior to consultation.

**Sample size**

The sample size indicated is, depending on the case, the number of subjects needed, the number of subjects included at the latest update or the number of subjects included at the end of the study. Note that some studies have substantial sample sizes, which is rarely the case with MDs in general: of all completed and ongoing studies, 10% include more than 1000 patients. Many feasibility studies on small samples are referenced.
**Study duration**
The envisaged duration of the studies identified is extremely variable – ranging from one day to several months or years.

**Conclusion**
The designs of the clinical trials identified in the ClinicalTrials.gov clinical trial database do not differ from those usually described in assessment applications for the various medical device categories submitted to the CNEDiMTS. There are no common studies in the two reviews carried out.
For the comparative studies including algorithms, the endpoints assessing the impact of each method (algorithm arm or comparator arm) are clinical or biological. For algorithm validation, the sensitivity and specificity are primarily studied.
For mobile health-related trials, only a few app-specific endpoints are described (e.g. the MARS score, usability, time or frequency of use). As regards the sample size and follow-up period, given the different objectives of these trials, it is not possible to identify a specific feature in respect of the use of connected devices.
A majority of interventional studies are controlled and randomised (58% for studies involving algorithms and 79% for mobile health-related studies).

### 3.4 Data from the literature

No guidelines, health technology assessment report or meta-analysis relating to the specific methodological features of the clinical evaluation of connected medical devices were identified.

The search identified 1 literature review (2) and 7 guidance documents (3-9).

Three United States Food and Drug Administration (FDA) guidance documents were identified:
- the “Mobile medical applications” guidance (5). This guidance document does not provide details of any specific assessment methods in respect of clinical trials. It especially contains definitions and regulatory requirements relating to these products;
- the draft “Multiple function device products” guidance (3). This draft guidance document defines these products and sets out the principles and content of applications to be submitted to the FDA. It does not provide details of any specific assessment methods in respect of clinical trials;
- the “Software as a Medical Device (SAMD)” guidance (4). This guidance document was produced by the International Medical Device Regulators Forum (IMDRF) The “clinical evaluation” section does not describe any specific assessment methods in respect of clinical trials.

The guidance document published on the New Zealand Ministry of Health website concerns mobile “health apps” (7). This guidance document does not describe any specific methodological features for conducting clinical trials. Nevertheless, examples of tools enabling users to rate a mobile app are mentioned, such as the “Mobile App Rating Scale” (MARS) and “App Chronic Disease Checklist” (ACDC) allowing a score to be assigned based on functionality, aesthetics, information quality, ease of use, etc. Websites listing health apps, classified according to themes, are also cited. This guidance document specifies that an assessment framework for “clinical apps” is underway (details of the projected schedule are not provided).

Furthermore, 6 publications addressing the issue of the clinical evaluation of digital technologies and proposing methodological adaptations were reviewed (10-15). Details of these supplementary data are provided in a specific section.
Finally, based on the specific search run on “in silico” trials in the medical device sector, no guidelines, health technology assessment reports, meta-analyses or systematic literature reviews were identified. Research and areas of study, particularly in the United States and in Europe, were detailed with a view to identifying the development prospects of these trials.

3.4.1 Methodological guidance documents

Three methodological guidance documents were selected:

- the “digital health technologies” methodological guidance, published by the National Institute for Health and Care Excellence (NICE) in 2018 (8);
- the “digital applications” methodological guidance, published by the York Health Economic Consortium (YHEC) in 2016 (6). This guidance document is based on a report identifying the different types of studies concerned by digital health evaluation (16). Details of examples of trials including mobile apps are provided (17);
- the “digital health interventions” methodological guidance, published by the World Health Organization (WHO) in 2016 (9).

3.4.1.1. National Institute for Health and Care Excellence (NICE) guidance, United Kingdom, 2018

This guidance document covers digital health technologies (DHT), such as mobile apps, wearable devices, and software algorithms.

It was developed between June and December 2018, without conducting a systematic literature review or referral to a multidisciplinary working group.

The objective is to identify the evidence standards required in terms of effectiveness and economic impact, in order to assist developers in setting up their clinical development plan.

This guidance document is based on a risk level-based approach according to digital technology functionality. The evidence standards required in terms of effectiveness stem from a tiered functional classification (1, 2, 3A, 3B). Details of this classification are provided in Table 3.

Table 3. Functional classification according to intended use of technology

<table>
<thead>
<tr>
<th>Functional classification according to intended use</th>
<th>Functional classification</th>
<th>Description</th>
<th>Includes (for example)</th>
<th>Excludes (for example)</th>
</tr>
</thead>
</table>
| Tier 1  
No direct user benefits                      | System service            | Improves system efficiency with no measurable individual patient outcomes. | Electronic prescribing systems, electronic health record platforms. | Systems that provide advice, treatment or diagnoses. |
| Tier 2  
Helps users to understand healthy living and illnesses  
No measurable user outcomes                   | Inform                    | Provides general information on specific conditions or about healthy living. | Apps providing advice for healthy lifestyles (such as recipes). | Tools that collect data or provide treatment for a condition. Apps that allow communication (among users or with third party). |
|                                                   | Simple monitoring         | Allows users to record health parameters to create health diaries. This information is not shared with or sent to others. | Electronic health diaries (symptoms, health status, etc.). | DHTs that share information (among users or with third party), tools that provide treatment for a condition. |
## Functional classification according to intended use

<table>
<thead>
<tr>
<th>Functional classification</th>
<th>Functional classification</th>
<th>Description</th>
<th>Includes (for example)</th>
<th>Excludes (for example)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicate</td>
<td></td>
<td>Allows communication between users and professionals, carers, etc. Advice is provided by a professional, but not by the DHT.</td>
<td>Instant messaging apps. Video conference-style consultation software. Platforms for communication with carers.</td>
<td>DHTs that provide clinical content themselves (such as cognitive behavioural programmes for depression).</td>
</tr>
<tr>
<td>Preventative behaviour change</td>
<td></td>
<td>Designed to change lifestyle, behaviour (smoking, eating, alcohol, etc.).</td>
<td>Smoking cessation DHTs and those used as part of weight loss programmes.</td>
<td>DHTs that describe themselves as a treatment for a diagnosed condition, apps that provide general healthy lifestyle advice.</td>
</tr>
<tr>
<td>Self-manage</td>
<td></td>
<td>Aims to help people with a diagnosed condition to manage their health. May include symptom tracking function that connects with a healthcare professional.</td>
<td>DHTs that allow users to record, and optionally to send, data to a healthcare professional to improve management of their condition.</td>
<td>DHTs that describe themselves as a treatment for a diagnosed condition, apps that automatically monitor and report data to a healthcare professional or third-party organisation.</td>
</tr>
<tr>
<td>Treat</td>
<td></td>
<td>Provides treatment for a diagnosed condition (such as CBT for anxiety).</td>
<td>DHTs for treating mental health. Apps that advise on treatments.</td>
<td>Apps that provide general health advice for patients, DHTs intended to improve the care pathway.</td>
</tr>
<tr>
<td>Active monitoring</td>
<td></td>
<td>Automatically records health information and transmits the data to a professional, carer or third-party organisation, without any input from the user.</td>
<td>Medical devices such as implants, sensors in which the data are automatically transmitted.</td>
<td>DHTs that allow a user to choose if and when to send recorded data.</td>
</tr>
<tr>
<td>Calculate</td>
<td></td>
<td>Tools that perform clinical calculations that are likely to affect clinical care decisions.</td>
<td>DHTs for use by professionals or users to calculate parameters (software).</td>
<td>DHTs that diagnose or provide treatment for a condition.</td>
</tr>
<tr>
<td>Diagnose</td>
<td></td>
<td>Uses data to diagnose a condition, or to guide a diagnostic decision.</td>
<td>DHTs that diagnose specified clinical conditions using clinical data.</td>
<td>DHTs that offer general lists of signs and symptoms for healthcare conditions.</td>
</tr>
</tbody>
</table>

The health technologies included are extremely diverse; not all fall under medical device status in view of their intended use. For classification tier 3B corresponding to potentially CE-marked health technologies, the evidence standards should be considered as complementary to the requirements for regulatory approval under the CE marking framework.

For each tier, some flexibility in terms of the evidence standard is proposed (minimum evidence standard and best practice standard). The evidence standards are cumulative to address the various issues raised by health technology use.

For example, support services and health technologies with only data transmission capabilities (with no user parameter measurement) are classified as tier 1 and 2, respectively. This means that these health technologies require evidence of reliability, reproducibility, safety, content accuracy, and feasibility of implementation of the health technology in the healthcare system and of its acceptability by users, including compliance with existing standards and guidelines.

As regards self-monitoring devices and medical telemonitoring systems classified in tier 3B, these devices are required to meet the requirements for the previous tiers (1 to 3A) and also provide evidence of their effectiveness with the aid of high-quality experimental or quasi-experimental studies, with the RCT representing the best practice standard.

Clinical trial methodologies are described for functional classification tiers 3A and 3B. Details of the evidence standards for functional classification tiers 3A and 3B are provided in Table 4.
### Table 4. Evidence standards for functional classification tiers 3A and 3B

<table>
<thead>
<tr>
<th>Evidence category</th>
<th>Minimum evidence standard</th>
<th>Best practice standard</th>
</tr>
</thead>
</table>
| **Demonstrating effectiveness** | High quality observational or quasi-experimental studies demonstrating relevant outcomes. These studies should present comparative data. Comparisons could include:  
• relevant outcomes in a control group;  
• use of historical controls;  
• routinely collected data. Relevant outcomes may include:  
• behavioural or condition-related user outcomes such as reduction in smoking or improvement in condition management;  
• evidence of positive behaviour change;  
• user satisfaction. | High quality intervention study (quasi-experimental or experimental design) which incorporates a comparison group, showing improvements in relevant outcomes, such as:  
• patient-reported outcomes (preferably using validated tools) including symptom severity or quality of life;  
• other clinical measures of disease severity or disability;  
• physiological measures;  
• user satisfaction and engagement;  
• health and social care resource use. The comparator should be a care option that is reflective of standard care. |

| Use of appropriate behaviour change techniques | Be able to show that the techniques used in the DHT are:  
• consistent with recommended practice (NICE or relevant professional organisations);  
• appropriate for the target population. | Published qualitative or quantitative evidence showing that the techniques used in the DHT are:  
• based on published and recognised effective behaviour change techniques;  
• aligned with recommended practice;  
• appropriate for the target population. |

### Functional classification tier 3B

| Demonstrating effectiveness | High quality intervention study (experimental or quasi-experimental design) showing improvements in relevant outcomes, such as:  
• diagnostic accuracy;  
• patient-reported outcomes (PRO) including quality of life;  
• other clinical measures of disease severity or disability;  
• physiological measures;  
• user satisfaction and engagement. The comparator should be a care option that is reflective of the current care pathway. | High quality randomised controlled studies done in a setting relevant to the use of the DHT, comparing the DHT with a relevant comparator and demonstrating consistent benefit including a validated clinical outcome. Alternatively, a well-conducted meta-analysis of randomised controlled studies if there are enough available studies on the DHT. |

A commentary referring to the publication of the NICE guidance document has been published (18). It is particularly noted that the specific features of these health technologies relate to their very rapid development and their ability to provide real-time data. These characteristics have entailed the development of methodologies, such as Multiphase Optimisation Strategy (MOST) and Sequential Multiple Assignment Randomised Trials (SMART).

This commentary notes that the studies required must conform to the same clinical standards as other technologies, including transparency of methods, *a-priori* analysis plans, and full publication of all results. The author notes that, beyond effectiveness, there are other issues to be considered, such as ethical and safety aspects.

#### 3.4.1.2. York Health Economic Consortium (YHEC) guide and associated report, United Kingdom, 2016

##### 3.4.1.2.1. York Health Economic Consortium (YHEC) guide

The guide prepared in 2016 by the York Health Economic Consortium (6) aims to support app developers to understand and create the evidence required to demonstrate the health and care benefits of their apps.
The York Health Economic Consortium provides health economics consultancy to the National Health Service (NHS). Decisions on adopting innovative technologies into the NHS are based on a health technology approach (HTA).

The scope of this guide covers digital applications, which may include a mobile app, a web-based application or in certain cases a digital service.

The guide is structured around six criteria (detailed below) that must be supported with evidence. The developer must compare the app intervention to current practice and consider the system as a whole and not the app separately from these components. For each criterion, specific questions and answers to these research questions are set out, aiming to help app developers produce the evidence required.

The guide states that all criteria are important. However, for some apps, a developer may be able to demonstrate that specific criteria do not apply to their app. Hence not all developers will be required to provide evidence on all criteria, rather the evidence required will be proportionate to the potential impact of the app on each criterion and the risk of such impact occurring.

A summary of the types of studies proposed in the guide for each criterion is detailed below.

► **Safety**

The app intervention should be able to demonstrate that it reduces, or has no impact on, the risks to patients and carers from harm, compared to current practice.

This guide states that clinical studies alone are unlikely to detect all potential harms to patients and users. Therefore, app developers are encouraged to identify potential issues as early as possible, via regular clinician and user feedback.

Safety studies can take several forms. **Cohort studies and case series** usually have a longer duration than a randomised controlled trial (RCT) and ideally have a comparator arm. They are often pragmatic studies and set in routine clinical practice. The function of the app intervention may determine the type of study required. Nevertheless, for safety studies, there is no single gold standard.

If the app intervention is distributed and marketed, app developers are encouraged to incorporate safety outcomes data collection, particularly via **register studies** that can be used for the systematic collection of data on safety.

► **Effectiveness**

The app intervention should be able to demonstrate it improves the health of users, or has no impact on this compared to current practice. If there are no clinical benefits, then the app must be able to demonstrate benefit in some of the remaining categories.

Study designs in this domain need to focus on generating evidence on the impact of the app intervention on clinical and well-being outcomes. The organisation implemented for the use of the health technology is a key factor influencing effectiveness. It is thus useful to determine whether the app is used in accordance with recommended practice. The guide specifies that the drop-out rate with the use of apps in practice is higher than the rate for most other medical interventions. Thus, how the app is used in practice and the impact of that on health outcomes should be considered.

Clinical studies should identify the optimal mix of each component of the intervention. The proposed methodologies relate to **Multiphase Optimisation Strategy (MOST)** or Sequential Multiple Assignment Randomised Trials (SMART). The guide specifies that once the app intervention is optimised, further trials, typically RCTs, are still required to identify the relative clinical benefits. An RCT is the gold standard trial design for evaluating relative effectiveness, but other designs including cohort studies and case series may provide sufficient evidence to inform an evaluation, provided that potential biases can be adequately addressed.
If the app is distributed and marketed, app developers are encouraged to conduct “real world” studies which collect evidence of outcomes from using the app in a natural environment, particularly by means of register studies or cohort studies. Indeed, RCTs and other quantitative studies are conducted in a sample population. “Real world” studies particularly help identify whether the results from this sample apply to the whole population of interest.

For apps which are designed for use by patients, it is important that app developers measure outcomes which reflect the impact of the app on their health and well-being. Using validated specific or generic questionnaires can demonstrate the impact of using the app on generic health outcomes (using, for example, the EQ-5D questionnaire) or disease-specific outcomes such as Patient Reported Outcome Measures (PROMs). Hence, app developers should ideally provide evidence using such tools (likely to be incorporated into evaluations using RCT, cohort or case series designs).

► Patient/user-centredness

The app intervention should be able to demonstrate that patients/users find it useful, that it improves outcomes relevant to them and enhances their health experience, compared to current practice.

The guide specifies that evidence may relate to the usability of apps, or “usability testing”. Patient involvement is influential in developing the app pre-launch. Such involvement may take the form of observations, interviews, focus groups, questionnaires and more structured user testing. Capturing the patient experience when using the app by means of Patient Reported Experience Measures (PREMs) is a key item of evidence to be collected in clinical studies.

The benefits on patient well-being are harder to measure, although there are some validated scales for measuring aspects of well-being (scales not specified in the guide). App developers may choose to specify what the anticipated benefits are, even if there is no empirical data to support these. The study methodologies may include a variety of study designs, both qualitative and quantitative.

► Timeliness of care

The app intervention should be able to demonstrate it reduces, or has no impact on, unnecessary delays for patients in the health and social care system and may increase the likelihood that each patient receives prompt attention, compared to current practice.

Evidence could come from economic studies, and also qualitative and quantitative studies.

► Efficiency

The app intervention should be able to demonstrate it can reduce, or not change, the need for, and hence cost of, resources required to deliver high-quality patient-centred care, compared to current practice. If it costs more, this must be justified by benefits across other criteria.

Evidence could come from economic studies alongside qualitative and quantitative studies (RCTs).

► Equity

The app intervention should be able to demonstrate there are minimal barriers to its use arising from a patient’s disability, age, race, ethnicity, gender, sexual orientation, and religion or belief.

The evidence may come from intervention studies (for example, from details of characteristics of recruited users), usability studies, data collected through the app itself, and from user feedback elicited from questionnaires, interviews and focus groups. There is no specific type of study.
3.4.1.2.2. York Health Economic Consortium (YHEC) report

The methodological guidance document published by the York Health Economic Consortium (YHEC) is based on a report identifying the different types of studies concerned by digital health evaluation (16).

This report was completed in 2016 by the York Health Economic Consortium (YHEC) on behalf of the National Institute for Health and Care Excellence (NICE). Details of the method used to draft the report are not provided (systematic literature review, search criteria, work group set-up, etc.).

The scope relates to digital health, which includes mobile health (mHealth), health information technology (IT), wearable devices, telehealth and telemedicine, and personalised medicine.

The aim of this report was to provide information on the strengths and limitations of various research designs, at different stages of product development, and the limitations in terms of interpretation of results.

The authors specify that developers will likely build a cumulative evidence base, gained at each stage of development, before widespread adoption of the technology. As a consequence of the probable need for several study designs over the lifetime of the product to accumulate evidence, study types are not mapped to specific research questions. The report provides a range of study types available to generate the evidence required to answer each research question. Moreover, as each product is very different, no single study design is recommended. The importance of evaluating the intervention package and not the product in isolation from its components is highlighted.

The studies listed include quantitative studies (experimental and observational), qualitative studies, mixed-method studies (both quantitative and qualitative) as well as engineering studies.

Figure 1 shows the various types of studies proposed.

Figure 1. Study designs mentioned in the YHEC report, 2016 (16)

For the evaluation of the benefit/risk of a new health technology, **well-designed RCTs are preferred**. Qualitative studies can be used to explore patient behaviours, experiences, attitudes, and interactions, with a view to enhancing understanding of the health technology with respect to users in a specific context. Both methodologies may be combined; this is known as mixed methods research. It is defined as integrating quantitative and qualitative data collection and analysis in a single study.

Engineering study designs are used to address optimisation problems. Such designs seek to identify the optimal set of parameters, which maximise the probability that multiple objectives can be met, subject to constraints, which can include behavioural aspects and cost of the product. These designs have been used widely since the 1970s and are increasingly being adopted in healthcare, particularly with behavioural interventions. These types of studies are carried out before an RCT is conducted.

The engineering study designs detailed in this report are as follows.

► **Ecological momentary assessment (EMA) studies**

These studies enable participants to self-report on changes in symptoms e.g. pain, moods and behaviour, in near real time in their daily lives, with electronic devices. These studies can use the longitudinal data to correlate changes in mood/behaviour to changes in interventions over time. Initially, these studies were adopted in the field of clinical psychopharmacology. They can thus examine relationships across variables over time and potentially develop algorithms to indicate when there is a step change in behaviour or mood.

The authors point out that qualitative user feedback (e.g. questionnaires, focus groups) is important especially during the early stages of intervention development to optimise the interventions before an RCT is conducted.

Identified strengths:
- Reliability of the data collected.
- Qualitative data can be incorporated particularly during the early stages of the intervention development so that deficiencies in the systems can be identified and corrected before controlled trials are conducted.
- Enables large quantities of data to be collected.
- Study data can be used by clinicians and patients to monitor treatment progress and inform when changes in medication or an intervention are required.
- Study design including data collection and analysis.

Identified weaknesses:
- Some patients or some clinicians may not be open to this form of intervention.
- Finding a suitable control can be difficult.
- Methodology is still developing.

► **Small studies and N-1 studies**

These studies involve a very small number of participants (under 10) and focus on observing changes in individual behaviour or event rates over time. Hence, the design framework essentially consists of studying a single person or small group of persons over time with repeated measurement of the outcome.

Identified strengths:
- Patients and clinicians may recognise ineffective interventions quickly, thereby reducing costs and potential exposure to adverse effects.
They may help engage patients in their own care and thereby improve outcomes.
- They enable research to be conducted as part of clinical practice.

**Identified weaknesses:**
- The benefits to the patient and clinician must be clear from the outset.
- Ethics boards may not accept the design as a valid form of exploratory research.
- Statistical analysis procedures need further development.
- The interventions must be carefully selected and have no lasting effect.
- Low external validity.

**Interrupted times series**
A time series is a sequence of data points made over a continuous time interval, using successive measurements across that interval and at the same unit of time between each consecutive measurement. Examples include a daily chart of blood pressure or pain scores. These examples relate to an individual's experience but a time chart can also measure population-level events such as the incidence rates of new diseases or mortality or uptake of new public health interventions.

A time series of repeated observations of a particular event collected over time is divided into two segments. The first segment comprises rates of the event before the intervention and the second segment is the rates after the intervention. A statistical analysis called “segmented regression” is used to measure the changes in the rate of the outcomes and trend in the post-intervention period compared with the pre-intervention period. These studies may also include a comparator arm where no intervention occurred.

**Identified strengths:**
- The use of trend analysis over time can avoid the mistake of attributing a continuing reduction benefit to the intervention. The change may be wrongly related to the intervention when in fact the change was due to other factors.
- One can conduct stratified analyses to evaluate the differential impact of an intervention on subpopulations (e.g. by age, sex, race).
- This type of study provides clear information on the shape and direction of change in the post-intervention period compared with the pre-intervention period.

**Identified weaknesses:**
- Requires statistical knowledge to undertake the regression analysis.
- The causality assumption is only valid if no other variable or process changes at the same time.
- A minimum of 8 observations is required to provide adequate power for the regression analysis.
- Interpretation is enhanced if there is a direct control but this adds to complexity.

**Multiphase Optimisation Strategy (MOST)**
This strategy is aimed at optimising and evaluating multicomponent behavioural interventions. MOST includes a randomised study for intervention evaluation, but includes other phases of research before the RCT. These earlier phases are aimed at intervention optimisation using selected criteria and to achieve a specific goal.

The MOST approach seeks to use highly efficient experimental designs to optimise the various components using feedback informed by the behaviour of participants. A component is any part of an intervention that can reasonably be separated out for study. These are conducted before an RCT is undertaken.
MOST comprises of 3 phases:

- a preparation phase used to create the theoretical model to identify which intervention components to examine and what the optimisation criterion is;
- an optimisation phase in which the investigator empirically tests options to identify which individual components make up the intervention that meets the optimisation criterion. This phase uses a randomised factorial trial. The outcome from this phase is identification of the components that make up the optimised intervention, together with an understanding of its likely effectiveness. At this point, a decision is required on whether there is sufficient promise to warrant undertaking an RCT. If so, the evaluation phase starts. If not, the investigator may return to phase 1 and pilot new components. Hence the process can be iterative;
- an evaluation phase consisting of conducting an RCT, to compare the optimised intervention to a suitable control. If the RCT indicates that the optimised intervention is not effective, then the investigator may return to the first phase.

**Identified strengths:**

- It enables testing of an optimised intervention efficiently. Conventional RCTs test whether an intervention works or not and these results may inform a revision to the optimisation of the intervention, requiring a further randomised trial.
- It enables the effects of the individual components and their combinations to be tested, unlike a conventional RCT which just measures the efficacy of the intervention only as a whole.
- Design strategies can be flexible with few limitations, in theory, on the number of components being tested, ability to use short-term surrogate endpoints to capture behavioural change and enabling a range of design methods including factorial experiments in the evaluation.
- Traditional measures of statistical significance can be applied to the results so the analyses do not lose any rigour despite the flexibility.

**Identified weaknesses:**

- Implementing the design can be challenging to ensure each participant receives the appropriate components, at the appropriate times, and outcomes are recorded and analysed. Thus, careful planning, good data collection and well-trained staff are essential.
- Decisions may require trade-offs between outcomes.
- Funding bodies may be unfamiliar with the framework and unwilling to depart from the gold standard 'RCT' approach, particularly given the uncertainties at the start on what the intervention will look like.

**Sequential Multiple Assignment Randomised Trials (SMART)**

These trials are used to develop adaptive interventions to manage complex health disorders such as cancers, addictions and depression. Managing these diseases often entails sequencing, and repeating, different potential approaches depending on participants' responses. An adaptive intervention is a sequence of decision rules that specify whether, how, when and based on which measures, to alter the intensity, type or delivery of treatments at pre-defined decision stages in the management care plan.

The aim of a SMART is to develop the optimal adaptive intervention, but clinical evaluation is still required. SMARTs are multistage randomised trial designs. Each participant in a SMART moves through multiple stages of treatment and at each stage is randomised to the next therapy/intervention.

The authors point out that SMART is a form of a factorial experimental design and hence standard data analysis methods can be applied.
The strengths and weaknesses of SMART trials are similar to those for MOST (MOST trials encompass SMART). The authors specify that these study designs are the only study designs which can reliably tailor interventions to individuals because of their unique randomisation features. The design does not require large sample sizes.

### Real world studies

Results from clinical trials may not always be a useful aid for decision-making – particularly if these do not measure the value of the intervention when used in a practical, real life setting. There is no standard design for this type of study. The focus has been on collecting ‘real life’ data to demonstrate effectiveness in a naturalistic environment which may differ from the controlled environment of an RCT.

The most common sources of data are:
- databases: user data are fully anonymised and aggregated in order to conduct research. Clinicians can also use the data, which patients record in the connected device to inform decision-making about treatment options, supporting a personalised treatment approach;
- patient and population surveys;
- observational data from cohort and case studies;
- registries: these involve analysing all patients treated at a particular centre, using pre-defined fields to capture clinical outcomes and adverse events.

**Identified strengths:**
- These trials should contain a more representative patient population than an experimental study.
- Can assess clinically relevant endpoints, rather than the often more short-term surrogate endpoints used in clinical trials.
- Can provide longer-term information on safety, resource use and costs.

**Identified weaknesses:**
- The cost of setting up a registry or database are high.
- Incentives to enter data may be weak so missing or poor-quality data can be problematic.
- Definitions for events or the threshold to include events in say a patient record may differ across users.
- Vital information needed to interpret results such as severity of condition may not be captured.
- Biases may creep into the data.

### A/B testing

The authors specify that most apps, particularly those that are web-based will have been subject to formal or informal A/B testing. A/B testing offers a randomised experiment with 2 variants, A and B, which are normally called the control and the variation. User responses to each version are compared to inform on their relative effectiveness. With web applications, the variables are often web pages which essentially are the same except for variations in individual elements like layouts, images or colours. In such online settings, the goal of A/B testing is to identify changes to web pages that increase or maximise an outcome of interest (e.g. click-through rate for a banner advertisement).

The current web page is associated with the null hypothesis.

These tests have randomisation in common with RCTs but their objectives are very different.
Specific methodological features of clinical evaluation of a connected medical device

Identified strengths:
- This test can be performed continuously.
- The results can be actioned immediately enabling the site to be refreshed regularly.
- The planning and conduct of such tests is reasonably straightforward and not that expensive.
- Existing website users provide real-time randomised focus groups at a fraction of the cost of doing direct recruitment.

Identified weaknesses:
- Only one aspect can be changed at a time, so sequencing testing over time can be complex.
- Inefficient data collection, as none of the information from a previous test can be reused to draw conclusions about other variables in future.

Further study designs are detailed, particularly innovative study designs.

► Instrumental variables estimation

Instrumental variables estimation is a technique used to control for confounding variables and measurement error. Studies that are not randomised, such as observational studies, may be subject to biases from confounding variables, which limit the ability to interpret outcomes. Most confounding variables are observable, and may be controlled in regression models to minimise their impact on the results. However, some are unobservable. Instrumental variables estimation identifies the variables or instruments that have two properties: 1) they affect the treatment variable, i.e. the intervention, but 2) have no direct effect on the outcome variable. Once the instrumental variable has been identified, it can be entered into the regression models (along with the other confounding variables) to determine the effect caused by the intervention.

► Use of Big Data in healthcare

Big data is defined in this report as a huge volume of data that exceed the processing capacity of conventional database systems and that may be characterised by 6 “V” features: 1) Volume, 2) Velocity, 3) Variety, 4) Value, 5) Variability and 6) Veracity.

In terms of healthcare, Big Data refers to large, complex electronic datasets that are difficult to manipulate with existing methods and software. It may include genomic, clinical, and behavioural data, as well as publication and reference data, administrative and business data.

Sophisticated analysis techniques may be applied to this integrated data to provide individual patient profiles alongside cost and clinical efficacy markers, enabling the most clinically and cost-effective interventions to be identified and tailored to individual patients (the techniques are not detailed in the report).

The identified strengths are as follows:
- improved health outcomes due to more accurate and precise diagnostics (tailored to individuals);
- cost reductions through earlier identification of disease;
- predicting and managing public health, such as obesity and diabetes, more effectively;
- reducing unplanned and emergency admission through better healthcare management.

These types of data may be garnered through patient’s own mobile devices, allowing, for instance, automated monitoring of subjective scores (mood, pain), as well as clinical markers (weight, blood pressure, glucose levels).

The authors point out the potential of integrated Big Data in healthcare systems, for collecting data via apps or other digital devices, despite there still being a number of barriers, such as budget and infrastructure, as well as issues surrounding data confidentiality to overcome.
An example of a mobile app evaluation strategy is detailed in Table 5.

### Table 5. Example of mobile app evaluation strategy

<table>
<thead>
<tr>
<th>Research methods</th>
<th>Purpose</th>
<th>Example measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Formative research</strong></td>
<td>Focus groups; Online surveys.</td>
<td>To inform the development of the intervention content and process.</td>
</tr>
<tr>
<td><strong>Pretesting</strong></td>
<td>Online surveys; Focus groups; Individual interviews.</td>
<td>• To determine acceptability of proposed intervention to target audience; • To improve and refine intervention on basis of feedback.</td>
</tr>
<tr>
<td><strong>Pilot study</strong></td>
<td>Small and nonrandomised.</td>
<td>• To test content and regimen of intervention; • To test processes.</td>
</tr>
<tr>
<td><strong>Randomised control trial</strong></td>
<td>Pragmatic community-based randomised control trial.</td>
<td>To test the effect of the intervention in comparison with a control group.</td>
</tr>
<tr>
<td><strong>Qualitative research</strong></td>
<td>Semi-structured interviews.</td>
<td>• To improve the intervention further; • To determine implementation issues and methods.</td>
</tr>
<tr>
<td><strong>Evaluation of implementation impact</strong></td>
<td>Phone / online surveys; Semi-structured interviews.</td>
<td>To determine the effect of the intervention once scaled up.</td>
</tr>
</tbody>
</table>


**3.4.1.3. World Health Organization (WHO) guide, 2016**

The WHO guide entitled “Monitoring and evaluating digital health interventions” (9) is aimed at guiding the reader through the pathway from the stage of development of the digital health intervention, ranging from technical requirements to feasibility testing up to the stage of clinical evaluation and measurement of their impact. The scope covers mHealth and eHealth interventions. Details of the method used to draft the report are not provided (systematic literature review, search criteria, working group set-up, etc.).

The rapid development of these health technologies is highlighted.

In terms of clinical evaluation, the various stages are indicated according to the stage of maturity of the digital health intervention:

- **feasibility**: assess whether the digital health system works as intended in a given context;
- **usability**: assess whether the digital health system is used as intended;
- **efficacy/effectiveness**: assess whether the digital health intervention achieves the intended results in a research or non-research (controlled or uncontrolled) setting;
• implementation research: assess the digital health intervention in a given context, including medical practices.

The types of clinical studies listed include quantitative studies (experimental and observational), qualitative studies, and mixed-method studies (both quantitative and qualitative). They are featured in figure 2.

Figure 2. Study designs mentioned in the WHO guide, 2016 (9)


The authors point out that multicentre RCTs are the gold standard. To evaluate the digital health intervention under real world conditions (implementation phase), an observational or quasi-experimental study including qualitative and quantitative endpoints may be most appropriate.
Conclusion

Three recently published guidance documents were selected, the 2016 WHO guidance document and 2 guidance documents issued in the United Kingdom, the 2016 York Health Economic Consortium guide, and the 2018 NICE guide. Note: the NICE guidance document does not include NHS data.

They are intended to guide developers in identifying the evidence standards required to demonstrate the benefits provided by their health technologies.

The guidance document published by the York Health Economic Consortium (YHEC) relates to mobile eHealth apps and covers six criteria (safety, effectiveness, patient-centredness, timeliness of care, efficiency, equity). For each criterion, study methodologies are suggested.

The NICE guidance document relates to digital health technologies and has defined a tiered classification of evidence standards according to health technology functionality. For health technologies falling under the definition of a medical device, evidence of their effectiveness is required with RCTs deemed to be the best practice evidence.

The WHO guidance document relates to digital health interventions (mHealth or eHealth) and defines evidence standards according to the stage of maturity of the digital health intervention.

The scope of health technologies covered by these guidance documents is wide-ranging and does not align completely with the scope of devices covered by the CNEDIMTS project. Nevertheless, the study methodologies applicable to a connected medical device do not differ from those of other medical devices, with RCTs deemed to be the gold standard for assessing the effectiveness of digital health technologies.

The types of clinical studies listed in these guidance documents include quantitative studies (experimental and observational), qualitative studies, and mixed-method studies (both quantitative and qualitative). The YHEC report identified other studies, particularly engineering studies and so-called “innovative” study designs, such as instrumental variables estimation and the use of big data.

3.4.2 Literature review

The search identified 1 literature review analysing the methodologies of clinical trials involving mobile health apps (2).

The study by Pham et al. (2) conducted a review, on the ClinicalTrials.gov registry, of clinical trials involving mHealth apps between November 2014 and November 2015.

The search was conducted using the term “mobile application”.

The inclusion criteria were as follows:
- evaluation of mHealth apps;
- measurement of clinical outcomes;
- app deployed exclusively on a mobile phone as a native app and not as a Web-based app.

The exclusion criteria were as follows:
- interventions based solely on sending text messages “SMS-based interventions” or on phone calls;
- studies mixing mobile and non-mobile interventions in the same group;
- apps serving solely as an appointment reminder service;
- apps not requiring user input through data entry.

The search strategy identified 137 clinical trials of which 71 met the inclusion criteria.

The great majority of trials were interventional (68/71, 96%), with only 3 observational trials (3/71, 4%). Most trials used an RCT design (57/71, 80%).
Specific methodological features of clinical evaluation of a connected medical device

Qualitative outcomes (not specified) were evaluated in 17 trials (17/71, 24%).
Control group assignment was divided into standard care (30/59, 51%), active treatment (26/59, 44%), and waitlist (3/59, 5%).
The fields covered included:
- Mental health (12/71, 17%);
- Cardiovascular conditions (8/71, 11%);
- Diabetes (8/71, 11%);
- Cancer (7/71, 10%).
Of these trials, only 7 had been completed at the time this retrospective review was conducted (7/71, 10%).
The average study duration, defined as the time from recruitment to complete data collection, was 20 months. The median sample size was 112 (20 trials had a sample size of 0-49 (20/71, 28%), 33 had a sample size of 101-499 (33/71, 47%), and 8 had sample sizes of over 500 participants (8/71, 11%)). The largest trial had 12,000 participants.
The authors conclude that mHealth evaluation methodology has not deviated from common methods and stress the need for robust clinical evaluation to evaluate the significant impact of health technologies.

Conclusion
A literature review analysing the methodologies of clinical trials involving mobile health apps was identified.
The designs of the clinical studies identified in the ClinicalTrials.gov clinical trial database do not differ from those usually described in assessment applications for the various medical device categories submitted to the CNEDiMTS.
Most of the studies were interventional and involved an RCT (80%). The authors stress the importance of robust clinical evaluation to provide evidence of the impact of these health technologies.

3.4.3 Supplementary data
The documentary search strategy identified 6 publications addressing the issue of clinical evaluation of digital health technologies and proposing methodological adaptations (10-15).
The specific features associated with digital health technologies identified by the authors related to:
- rapid technological development (10-15) and regular upgrades of health technologies (11);
- interactions between participants. These health technologies involve a complex set of components entailing a dynamic interplay with participants at any stage of the process (delivery, receipt, acceptance, and intention to use). These specific features infer that evaluation should not focus exclusively on delivery; it should be multifactorial. An assessment of human factors, such as acceptance and user adoption, should be included in monitoring to ensure that the health technology is implemented as planned (10);
- automatic data collection and communication via the Internet and other health technologies. These specific features infer the need to establish procedures to monitor the health indicators collected as outcome data so that changes in the patient’s health are detected and patient safety is ensured (10). The data collected may be biological, behavioural, or environmental. These include sensors that monitor phenomena with higher precision, improved sampling frequency, fewer missing data, or greater convenience (11). These data are collected in real
time without requiring user intervention, making it possible to carry out more timely analyses, with an ensuing reduction in the time needed to complete trials (11). The collection may apply to multiple parameters, enabling exploratory analyses on potential linkages between parameters, according to the context and environment of use (11);

- big data enabling continuous automated collection of parameters previously measured intermittently and in a time-limited fashion. Analysing these data can increase the discriminative power of any experimental study design to examine effects on variances, both between- and within-subjects (11);
- the appeal of these health technologies for some patients, facilitating patient recruitment in studies. The number of face-to-face interactions and the number of journeys can be reduced or even avoided altogether, allowing the patient to take part in the study remotely, from their home, instead of travelling to an academic research centre (10, 11);
- adaptation and tailoring according to the patients’ profile and in real time, helping increase patient engagement and increase the power of the intervention (11).

In order to address these specific features, the methodological adaptations identified are listed below.

RCTs are generally presented as the gold standard for providing evidence of the effectiveness of health interventions. Nevertheless, the obstacles reported in the literature in relation to the conduct of mobile health technology RCTs apply to:
- the time lag associated with conducting the study: 5 years on average, between patient recruitment and publication of the study. This renders the technology obsolete at the end of the study;
- the cost;
- randomisation for treatment assignment;
- the level of treatment adherence required.

In some cases, the rapidly evolving nature of the technology and uptake by patients and care teams may entail continuous modifications of certain components of the intervention group during the trial. This rapid development may lead developers of the study to move quickly from pilot to dissemination or avoid full-scale RCTs, limiting the clinical validity of the outcomes (11).

Kumar et al. (11) note that recent work in mHealth, and the data “revolution” that it augers, suggests that mHealth’s capabilities may change the strengths, weaknesses, and feasibility of existing research methods and may even enable development of new, more efficient designs. Nevertheless, many open questions exist about use of current methods in mHealth research. For example, the option of obtaining multiple repeated measures on a few participants, rather than a few measures on many participants. This option would render RCTs more efficient (i.e., quicker with a smaller sample).

With this in mind, potential research designs to evaluate the efficacy and effectiveness of mHealth interventions are described. The authors point out that, for so-called “mature” interventions, i.e. with previously validated prior or quasi-experimental data supporting their efficacy, RCTs are appropriate (10-15).

The methodological adaptations or alternatives to the RCT design proposed are:

- **Regression discontinuity design** (11); Participants are assigned to treatment or control based on a criteria cut-off score. The assignment variable may be any variable measured before treatment. The design is most powerful when the cut-off is placed at the mean of the assignment variable. This study design allows only a fraction of the participants to be used for the analysis (11).

Other statistical methods aimed at reducing confounding biases associated with observational studies are cited, such as the propensity score, stratifications and multivariable regression modeling (15).
Specific methodological features of clinical evaluation of a connected medical device

- **Stepped-wedge designs** (11)
  The intervention group can be compared with both their pre-test measures and with measures from other subjects who have not yet received the treatment, who form an independent and homogeneous control group. In this design, all participants are told that they will receive intervention, which ensures participants are not denied intervention. This design is appropriate if the intervention is going to be implemented with all individuals (or at all sites) (11, 15).

Developers might wish to provide upgrades on a regular basis during the evaluation of their technologies. However, changes to an intervention during a research study threaten internal validity of the results (11).

In order to address continuously evolving technologies, the CEEBIT “Continuous Evaluation of Evolving Behavioural Intervention Technologies” method (19) is proposed as one method for testing evolving interventions (11). This method would make it possible to retain the most efficacious version, based on *a priori* criteria. It would thus be well suited to ongoing evaluation of interventions as they go to scale, and continuously improve over time.

The authors point out that the RCT may still be applied if the level of inference in the intervention group is made around a package of robust intervention features or functions whose delivery will naturally adapt to changing technology environments and preferences over time and across contexts (11).

Finally, some authors suggest using model-based designs to adapt the intervention group using the mHealth technology.

Tailoring and personalisation in the intervention group require a better understanding of within- and between-subject differences. Statistical methods that better specify these differences are needed (11).

Strategies have been developed to address tailoring and optimisation of the intervention group. The example of the **Multiphase Optimisation Strategy (MOST)** is cited, helping identify the promising components of an intervention in a screening phase. These components are subsequently evaluated in a randomised trial.

For refining the intervention, **Sequential Multiple Assignment Randomised Trials (SMART)** can be used. Individuals are randomly assigned to various intervention choices over time. In SMART, researchers decide which aspects of treatments require investigation and then randomise individuals at each treatment decision based on feasibility, ethical issues, or other factors.

The MOST and SMART strategies are also cited by other authors (12-14) as approaches enabling efficient use of resources and helping identify the most appropriate intervention, prior to setting up an RCT involving mobile apps.

---

**Conclusion**

Specificities associated with CMDs are identified by several authors. These specificities concern rapid technology development, dynamic interplay with participants at all stages of the process, automated collection, processing and delivery of data. The challenge consists of aligning the clinical evaluation with these specificities, and with the pace of upgrades in particular. A number of converging proposals have emerged, particularly in terms of ensuring the reliability, validity and quality of digital health technology outcomes and accounting for a multifactorial assessment including acceptance and user adoption of the health technology.

The authors point out that, despite the difficulties encountered in respect of set-up, RCTs are the preferred study type to assess the benefit of a digital health technology. Methodological adaptations are nonetheless suggested.
3.4.4 Data on in silico trials

In silico research consists of developing mathematical modelling of a disease based on physiology and biology (e.g. the immune system, bronchial tree), with a view to forming a virtual patient population. The aim of this approach is to optimise conventional in vitro and in vivo studies. For example, it can be used to explore hypotheses using computer simulation, before launching conventional clinical trials, in order to increase the likelihood of their success or identify failures at an earlier stage.

A specific search was conducted on these trials. Of the 30 references listed on Medline, no guidelines, health technology assessment reports, meta-analyses or systematic literature reviews relating to in silico trials in the field of medical devices were identified.

Research and areas of study, particularly in the United States and in Europe, were searched with a view to identifying the development prospects of these trials.

- Data from the Food and Drug Administration (FDA), United States

In 2017, FDA launched a digital health innovation action plan (20), with, in particular, a pilot programme for software pre-certification aimed at adapting and enhancing the efficiency of the regulatory process in respect of these technologies and the creation of a “Center of Excellence for Digital Health”.

FDA\(^6\) recognises the public health benefits offered by modelling and simulation, including those in the area of in silico clinical trials. These trials are defined as the use of individualised computer simulation in development and or regulatory evaluation of medical products, medical devices, or medical interventions.

FDA advocates for the development of modelling and simulation and recommends their use in the following scenarios:
1) predicting clinical outcomes;
2) informing clinical trial designs;
3) supporting evidence of effectiveness;
4) identifying the most relevant patients to study;
5) predicting product safety.

FDA points out that in silico clinical trials can be used to replace human clinical trials, especially those that are intended to evaluate the risk of drug interactions.

A public-private partnership bringing together representatives of the FDA, industry, non-profits, and patient organisations, the Medical Device Innovation Consortium (MDIC), will particularly have the objective of studying the prospects of replacing clinical trials by modelling and simulation for the clinical evaluation of medical devices (21)\(^7\).

- Avicenna Consortium, Seventh Framework Program for Research and Technological Development (FP7), Europe

In the absence of guidelines, the European Commission funded the Avicenna support action between 2013 and 2015 with the objective of drafting a roadmap (22) for in silico clinical trials relating to biomedical products (medicinal products and medical devices).

The roadmap was drafted over an 18-month period using a panel of 525 international experts (35 countries) from industry, academic research, CROs (Contract Research Organisations), in silico

\(^6\) [https://www.fda.gov/ScienceResearch/AboutScienceResearchatFDA/ucm616822.htm](https://www.fda.gov/ScienceResearch/AboutScienceResearchatFDA/ucm616822.htm) [viewed on 04/01/2019].

clinical trial developers, agencies (NICE, FDA, EMA, etc.), patient associations, etc. It features in a publication (Viceconti et al., 2016 (23)).

This roadmap provides an overview of the current clinical development strategy in respect of biomedical products and expressed possibilities for the use of *in silico* trials (through 36 recommendations).

As regards of the current clinical strategy in respect of biomedical products, the Consortium points out that the only conclusive way to ensure the safety and efficacy of a biomedical product is to test it on living organisms, i.e. animals (in the preclinical assessment phase), and humans (in the clinical assessment phase). The preclinical testing process represents an essential step in the development of any potential biomedical product. Due to the hugely complex nature of human diseases, and to the significant differences between individuals, it is pointed out that it is not unusual for a product to perform exceptionally well in tightly controlled laboratory tests, but show some serious problems during clinical trials.

The Consortium has established a consensual definition of *in silico* clinical trials (ISCT) as the use of individualised computer simulation in the development and or regulatory evaluation of a medicinal product, medical device, or medical intervention. It is a subdomain of *in silico medicine*, the discipline that encompasses the use of individualised computer simulations in all aspects of the prevention, diagnosis, prognostic assessment, and treatment of disease.

Computer modelling and simulation is already being used in the development of biomedical products to estimate the pharmacokinetic and pharmacodynamic aspects of new compounds. For medical devices, these studies are used to assess fluid dynamics to predict how blood or other bodily fluids move inside and around the device being tested, or structural finite element analysis to make sure that the forces exchanged between the body and the device will not cause any harm.

*In silico* trials could help to:

- reduce the size and the duration of clinical trials, for example, by identifying which patients might be at greater risk of complications or providing earlier confirmation that the product is working as expected. *In silico* clinical trials might thus be used to leverage a smaller clinical trial population, by identifying those patients that will respond to the treatment;
- refine clinical trials through information on potential outcomes and by interpreting any adverse effects that might emerge, as well as better understanding how the tested product interacts with the individual patient anatomy and physiology, and predicting long-term or rare effects that clinical trials are unlikely to reveal;
- partially replace clinical trials in those situations where *in silico* clinical trials can generate scientifically robust evidence, particularly by replacing animal models with *in silico* models under appropriate conditions;
- complement clinical trials by offering the ability to test various multiple-disease scenarios (for example, where the patient has conditions interfering with the disease under investigation, such as diabetes associated with a heart rhythm disorder).

The roadmap states that *in silico* clinical trials will never entirely replace the clinical development and assessment of a biomedical product. The experts agree with the statement that these trials will never fully replace clinical trials on humans, but will help reduce and refine conventional clinical trials.

The Avicenna Consortium recommends the development and validation of *in silico* models that extrapolate clinical outcomes in humans.

It is recommended (recommendation 23) that developers contribute to collaborations with the aim of:
• establishing an *in silico* assessment framework for each family of devices, enabling research groups to refine predictors for the various failure modes;
• evaluating retrospectively a number of biomedical products, for which the clinical outcome is well known, in order to build confidence in the methods;
• running double-blind ISCT in parallel with existing *in vivo* clinical trials, in order to compare clinical outcomes. The Consortium specifies that precedence should be given to critical areas such as paediatric and rare diseases, in particular.

The recommendations aimed at regulatory bodies are as follows:

• Regulatory bodies should embrace *in silico* trials and, in collaboration with academic and industrial experts, develop the framework of standards, protocols, and shared resources required to evaluate the safety and the efficacy of biomedical products using *in silico* clinical trials (recommendation 29);
• Regulators should consider also regulating *in silico services* to be used for the assessment of biomedical products as medical devices in their own right (i.e. software as a medical device) (recommendation 30);
• It is recommended that European regulators follow the approach used by the USA Food and Drug Administration, which has recognised the strategic potential of *in silico* trials (recommendation 31).

**Conclusion**

*In silico* trials, based on mathematical modelling of a disease according to physiology and biology, show promise. They are currently used for exploratory purposes to assess pharmacokinetic and pharmacodynamic predictions in respect of a biomedical product, toxicological aspects or targeting of subjects to be included in conventional clinical trials. At the present time, these trials are among the tools available for the clinical development of biomedical products, used to complement existing tools in order to optimise the clinical development phase of a biomedical product. They are not a replacement for clinical trials on humans. *In silico* model development and validation are recommended by the Avicenna Consortium and the Food and Drug Administration, in order to identify potential extrapolation possibilities of their outcomes for the clinical evaluation of biomedical products. However, no trials of this type were identified in clinical trial databases or publications.
4. General conclusion and opinion of the working group

Based on the data provided by the other evaluation agencies contacted, applications submitted for which a clinical evaluation has been carried out by the CNEDiMTS and the clinical trial designs identified on the ClinicalTrials.gov clinical trial database, no specific methodological features for the clinical evaluation of connected medical devices were identified. The majority of the clinical studies reported were randomised controlled trials.

Nevertheless, the complexity of the assessment of a CMD was highlighted, due to its rapid development, impact on patients and on the organisation of care, as well as its interaction with other systems. The assessment for this type of MD indicated the importance of accounting for the overall and organisation assessment of the health technology as a whole.

The data from the literature reviews on the specific features of the clinical evaluation of CMDs are based on 3 methodological guidance documents and 1 literature review.

- As regards the guidance documents, the clinical study methodologies applicable to a connected medical device do not differ from those of other medical devices, with RCTs deemed to be the gold standard for assessing the effectiveness of digital health technologies.

The scope of health technologies covered by these guidance documents is wide-ranging and does not align completely with the scope of devices covered by the CNEDiMTS project. Nevertheless, the study methodologies proposed are dependent on the CMD’s intended ultimate purpose and its stage of clinical development. The types of clinical studies listed in these guidance documents include quantitative studies (experimental and observational), qualitative studies, and mixed-method studies (both quantitative and qualitative). The drafting report of one of the guidance documents identified other studies, particularly engineering studies and so-called “innovative” study designs, such as instrumental variables estimation and the use of big data.

- As regards the literature review, the methodologies of the clinical trials involving mobile health apps listed in the ClinicalTrials.gov database do not differ from those usually described in assessment applications for the various medical device categories submitted to the CNEDiMTS. Most of the studies were interventional and involved an RCT.

The 8 additional publications addressing the issue of the clinical evaluation of digital health technologies identified specificities associated with CMDs and proposed methodological adaptations. These specificities concern rapid technology development, dynamic interplay with participants at all stages of the process, automated collection, processing and delivery of data. The challenge consists of aligning the clinical evaluation with these specificities, and with the pace of upgrades in particular. A number of converging proposals have emerged, particularly in terms of ensuring the reliability, validity and quality of digital health technology outcomes and accounting for a multidimensional assessment including acceptance and user adoption of the health technology.

New methodologies are being developed, particularly simulation and modelling methods. At the present time, these methods are no replacement for conventional clinical trials, and are among the tools available for the clinical development of biomedical products, used to complement existing tools in order to optimise the clinical development phase of a biomedical product. No trials of this type were identified in clinical trial databases or publications.

The working group has not identified any specific clinical study methodology for the evaluation of CMDs; specificities common to some CMDs liable to influence the clinical development plan have nonetheless been identified. The importance of the different stages of the evaluation of the CMD, the rapidly evolving nature of the technology and upholding outcome quality is highlighted.

The resulting guide (see “Guide to the specific features of clinical evaluation of a CMD in view of its application for reimbursement”) has been reviewed and approved by the CNEDiMTS and submitted
to the HAS College, for information purposes. It is intended to help companies manufacturing or operating CMDs and seeking to have their health technology reimbursed by the French National Health Insurance to optimise the clinical studies supporting their future application for reimbursement in their development strategy.

This guide is thus intended for companies preparing a clinical study intended to provide evidence of the benefit of their CMD.
5. Appendices
Annexe 1. Work method

The main stages are described below.

- **Framing phase**
  During the framing phase, meetings were arranged with:
  - institutional representatives (Directorate-General of Health, DSS, Health Insurance Funds – CNAMTS, RSA, MSI, DGOS, ANSM, INCa);
  - manufacturers and manufacturer representatives (SNITEM);
  - service providers and service provider representatives (UPSADI, FEDEPSAD, SIDIV, UNPDM, SNADOM).
  A roadmap was drafted and submitted to the CNEDiMTS for approval prior to publication on the HAS website (24).

- **Data collection and review**
  - A systematic scientific literature review was conducted: the documentary search applied is detailed in Appendix 2.
  - Manufacturer and service provider associations were requested for the specific methodological features that they have already identified for CMDs, particularly where:
    - the CMD impacts healthcare organisation.
    - and where it includes a learning algorithm.
  The bibliographic report was sent to the working group prior to the first meeting.

- **Working group formation**
  A multidisciplinary working group, made up of representative healthcare professionals, was set up. This group includes methodologists, clinicians, a patient and a biostatistician. A call for applications was also broadcast on the HAS website and on Twitter.

- **Drafting and sending of stakeholder questionnaire**
  A questionnaire (Appendix 3) was sent to the identified stakeholders prior to the first working group meeting.
  The information provided by the stakeholders was passed onto the working group.

- **Working group opinion**
  The working group was requested, at the meetings, to:
  - approve (optionally supplement) the bibliographic report;
  - review the responses sent by the stakeholders following the issue of the questionnaire;
  - draft a guide to the specific methodological features of clinical evaluation of CMDs.
  The working group approves the documents drafted.

- **Guide review stage**
  The guide produced from the working group’s deliberations was sent to the various parties contacted during the framing phase as well as to the identified stakeholders (see review grid in Appendix 4 and list of guide review contributions in Appendix 5). Following the review stage, the various comments were reviewed by the working group for any amendments and approval of the definitive version.

- **Review by the CNEDiMTS**
  The finalised documents (guide and preparatory report) were reviewed and approved by the CNEDiMTS.

- **HAS College review**
  The finalised documents (guide and preparatory report) were submitted to the HAS College for information purposes.

- **Document publication**
  The finalised documents (guide and preparatory report) are published on-line on the HAS website.
Annexe 2. Documentary search

1. Method
A systematic documentary search was conducted from January 2013 to October 2018, by querying the following medical bibliographic databases:

- Medline (National Library of Medicine, United States);
- The Cochrane Library (Wiley Interscience, United States);
- websites publishing guidelines, health technology or economic assessment reports (see list of websites consulted);
- websites of learned societies or other organisations with expertise in the field studied (see list of websites consulted).

The search was restricted to publications in English and French. Health technology assessment reports, guidelines, consensus conferences, meta-analyses, systematic reviews, and clinical studies were searched. This automated search was supplemented by a manual search conducted based on the references cited in the selected publications. The experts’ bibliography and industry data were also taken into consideration as sources of data. If the studies provided by these sources met the selection criteria defined for the documentary search, they were included in the bibliography.

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

Table 6 shows the Medline database search strategy. The numbers of references indicated correspond to single references identified.

<table>
<thead>
<tr>
<th>Study type/Subject</th>
<th>Search period</th>
<th>Number of references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connected MD assessment</td>
<td>01/2013 – 10/2018</td>
<td>1,071</td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meta-analyses, literature reviews</td>
<td>01/2013 – 05/2018</td>
<td>204</td>
</tr>
<tr>
<td>Step 1 AND Step 2 AND</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Specific methodological features of clinical evaluation of a connected medical device

Step 3

<table>
<thead>
<tr>
<th>Clinical trials</th>
<th>01/2013 – 05/2018</th>
<th>335</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health technology assessment</td>
<td>01/2013 – 05/2018</td>
<td>64</td>
</tr>
<tr>
<td>In silico clinical trials</td>
<td>01/2013 – 05/2018</td>
<td>30</td>
</tr>
</tbody>
</table>

Total number of references obtained: 1,163
Total number of articles analysed: 75

3. Sites consulted
Latest consultation: 10/2018

Information in French:
Bibliothèque médicale Lemanissier
Catalogue et index des sites médicaux francophones – CISMeF
Comité d’évaluation et de diffusion des innovations technologiques – CEDIT
Évaluation des technologies de santé pour l'aide à la décision (Fédération hospitalière de France) – ETSAD

Information in English:
Adelaide Health Technology Assessment – AHTA
Agency for Healthcare Research and Quality – AHRQ
Alberta Heritage Foundation for Medical Research – AHFMR
Alberta Medical Association
Allied Health Evidence
Australian Clinical Practice Guideline
Australian Safety and Efficacy Register of New Interventionsal Procedures – Surgical – ASERNIPS
Australia and New Zealand Horizon Scanning Network
Blue Cross Blue Shield Association – BCBS - Technology Evaluation Center
BMJ Clinical Evidence
British Columbia guidelines
California Technology Assessment Forum – CTAF
Canadian Agency for Drugs and Technologies in Health – CADTH
Canadian Task Force on Preventive Health Care
Centers for Disease Control and Prevention – CDC
Specific methodological features of clinical evaluation of a connected medical device

Belgian Health Care Knowledge Centre – KCE  
Centre for Clinical Effectiveness – CCE  
Centre for Effective Practice  
Centre for Reviews and Dissemination databases  
CMA Infobase  
Cochrane Library  
College of Physicians and Surgeons of Alberta – CPSA  
Euroscan  
Guidelines International Network – GIN  
Health Services Technology Assessment Text – HSTAT  
Institute for Clinical Evaluative Sciences – ICES  
Institute for Clinical Systems Improvement – ICSI  
Institute for Health Economics Alberta – IHE  
Institut national d’excellence en santé et en services sociaux – INESSS  
International Medical Devices Regulator Forum - IMDRF  
International Network of Agencies for Health Technology Assessment – INAHTA  
Medical Services Advisory Committee – MSAC  
National Coordinating Centre for Health Technology Assessment – NCCHTA  
National Guideline Clearinghouse – NGC  
National Health and Medical Research Council – NHMRC  
National Health Services Evidence  
Horizon Scanning Research & Intelligence Centre  
National Institute for Health and Clinical Excellence – NICE  
New Zealand Guidelines Group – NZGG  
New Zealand Health Technology Assessment – NZHTA  
Ontario Health Technology Advisory Committee – OHTAC  
Public Health Agency of Canada – Diseases Prevention and Control Guidelines  
Scottish Intercollegiate Guidelines Network – SIGN  
Singapore Ministry of Health  
Tripdatabase  
National Institutes of Health  
U.S. Food and Drug Administration - FDA  
Veterans affairs, Dep. Of Defense Clinical practice guidelines  
West Midlands Health Technology Assessment Collaboration – WMHTA  
World Health Organization

4. Clinical trials
Clinical trials were searched on ClinicalTrials.gov:  
A search was run with the term mhealth (415 references).  
Another search was run with algorithm (482 references).

5. Monitoring
In addition to the search, monitoring was carried out until 10/12/18 on the Internet and on journal contents, particularly the contents of the following journals:
- Journal of Medical Internet Research Jmir;  
- JMIR Research Protocols;  
- JMIR mHealth and uHealth.

Results
Number of single references identified (websites consulted and Medline): 22 on websites and 1,163 in Medline  
Numbers of references reviewed: 75  
Number of references selected: 10
Annexe 3. Questionnaire sent to stakeholders

We are contacting you to obtain your feedback or opinion on any specific methodological features to be factored into the clinical evaluation of CMDs. This request is being made within the framework of a project aimed at producing a specific methodological guide for clinical evaluation, to inform companies preparing to set up a clinical study intended to provide evidence of the benefit of their CMD.

We invite you to enter your responses directly in the document and would like to express our thanks in advance for your contribution to this project.

The clinical data provided by companies are reviewed according to the criteria of evidence-based medicine. The CNEDiMTS determines the clinical relevance and transferability to the French healthcare system.

The study type offering the best level of evidence is the double-blind randomised controlled trial design (or at least with an independent observer). However, in some scenarios, it may not be possible to conduct this type of study for MD evaluation. The Committee takes this context into account and may adapt its requirements, subject to explanations and justifications from the companies in such a scenario.

The CNEDiMTS has already listed scenarios in which it is not possible to conduct conventional randomised controlled trials and has identified the methods and conditions enabling high-quality clinical evaluation in “Methodological choices for the clinical development of medical devices”.

The CNEDiMTS has also issued an information document intended to provide some practical benchmarks regarding the CNEDiMTS’s principles of assessment in the methodological guidance document “Assessment principles established by the Medical Device and Health Technology Evaluation Committee (CNEDiMTS) to determine the reimbursement eligibility of medical devices for individual use.”

The CNEDiMTS’s expectations may vary according to:

- the MD category: the definition of a medical device\(^8\) involves a wide diversity in the MD sector, including products as diverse as, for example, dressings, prescription glasses, cardiac pacemakers or medical imaging devices. The Committee accounts, for example, for the inherent context relating to assistive products, whether they are MDs or not.

Clinical evaluation of the products, focusing on the disabled subject’s needs, lifestyle, life plans and environment, is always the preferred choice. However, for a number of product categories (energy storage and return prosthetic feet, for example), a prior technical assessment, in compliance with standards, is required.


European Regulation 2017/745 defines a medical device (MD) as “any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes:

- diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease;
- diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability;
- investigation, replacement or modification of the anatomy or of a physiological or pathological process or state;
- providing information by means of in vitro examination of specimens derived from the human body, including organ, blood and tissue donations;

and which does not achieve its principal intended action by pharmacological, immunological or metabolic means, in or on the human body, but which may be assisted in its function by such means.”
• **the MD lifecycle**: this may be of shorter duration than the clinical investigation.
  
  For all devices, it is essential, insofar as possible, to anticipate incremental developments in the clinical development plan of the MD. This is a key factor when applying clinical data to new ranges (new models).

• **restricted scope for recruitment in a clinical trial**, due to the small target population.
  
  Conducting international multicentre trials helps increase the number of subjects liable to be included. In any case, the CNEDiMTS can only issue an approval in view of the expected patient benefit if the inherent risk of these MDs is sufficiently characterised and limited according to the type of MD.

As regards certain CMDs in particular, the specific features already identified that are liable to influence their clinical evaluation may be associated with their mode of action and with interactions with the environment of use, the increasingly rapid development cycle, ongoing upgrades of the product or the algorithm integrated in the CMD, automated data collection, the patient’s ability to use the CMD, organisational aspects resulting from the use of the CMD impacting patients, their quality of life and/or organisation of care.

The aim is to identify the extent to which these specific features are liable to influence the clinical evaluation of certain CMDs and the method that can be used to address this.

**Section 1 – Short lifecycle of some CMDs**

The time required to conduct the clinical trial is liable to exceed the lifecycle of some CMDs.

**Question 1.1 – Are there any clinical trial methodologies quicker than a conventional clinical trial but nonetheless robust that can be applied to CMDs? If so, which?**

Response

- YES ☐
- NO ☐
- DON’T KNOW ☐

Justification

**Question 1.2 – Which clinical trial outcomes would be impacted by these specific methodological features?**

Response

- ☐ Type
- ☐ Duration
- ☐ Endpoints
- ☐ Sample size
- ☐ Randomisation method
- ☐ Outcome collection and analysis method
- ☐ Other

Justification
Question 1.3 – Have these clinical trial methodologies previously been applied for some CMDs? If so, what are the study references?

Response

YES □ NO □ DON’T KNOW □

Justification

Section 2 – Ongoing upgrades of some CMDs or algorithm(s)

Regardless of the CMD or type of algorithm used (deterministic or supervised learning mode), continuous upgrades may be applied.

Question 2.1 – Are there any clinical trial methodologies that can be used to account for ongoing upgrades of some CMDs or algorithms(s)? If so, which?

Response

YES □ NO □ DON’T KNOW □

Justification

Question 2.2 – Which clinical trial outcomes would be impacted by these specific methodological features?

Response

☐ Type
☐ Duration
☐ Endpoints
☐ Sample size
☐ Randomisation method
☐ Outcome collection and analysis method
☐ Other

Justification
**Question 2.3** – Have these clinical trial methodologies previously been applied for some CMDs? If so, what are the study references?

<table>
<thead>
<tr>
<th>Response</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>YES □</td>
<td>NO □</td>
<td>DON’T KNOW □</td>
<td></td>
</tr>
</tbody>
</table>

**Section 3 – Rapidly evolving nature of algorithms based on supervised machine learning processes**

One of the specificities of some CMDs lies in the rapidly evolving nature of included algorithms based on machine learning processes (this only applies to learning algorithms based on supervised learning). These algorithms are designed in such a way that their behaviour evolves over time, based on the data furnished. Their results thus tend to shift and are dependent at all times on the learning base provided, which evolves with use.

**Question 3.1** – Are there any clinical evaluation methods that can be used to account for the rapidly evolving nature of algorithms?

<table>
<thead>
<tr>
<th>Response</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>YES □</td>
<td>NO □</td>
<td>DON’T KNOW □</td>
<td></td>
</tr>
</tbody>
</table>

**Question 3.2** – If so, which?

<table>
<thead>
<tr>
<th>Response</th>
<th></th>
</tr>
</thead>
</table>

**Question 3.3** – Have these clinical trial methodologies previously been applied for some CMDs? If so, what are the study references?

<table>
<thead>
<tr>
<th>Response</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>YES □</td>
<td>NO □</td>
<td>DON’T KNOW □</td>
<td></td>
</tr>
</tbody>
</table>

**Question 3.4** – Is it possible to “freeze” the algorithm at a time T to enable clinical evaluation?

<table>
<thead>
<tr>
<th>Response</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>YES □</td>
<td>NO □</td>
<td>DON’T KNOW □</td>
<td></td>
</tr>
</tbody>
</table>
Question 3.5 – Is it possible to ensure the non-regressiveness of the algorithm before the new clinical evaluation?

Response

YES ☐  NO ☐  DON’T KNOW ☐

Question 3.6 – If so, which procedures would be applied (requirements, criteria, etc.)?

Response


Section 4 – Clinical data collection

One of the specificities of some CMDs is linked with the fact that they can generate patient data continuously via apps or sensors, in particular.

Question 4.1 – With a view to ensuring that the analysis is exhaustive and relevant, is it possible to stipulate specifications in respect of the data collected in the clinical trial protocol?

Response

YES ☐  NO ☐  DON’T KNOW ☐


Some CMDs can generate big data.

Question 4.2 – Will the “big data” approach affect the robustness of conventional CMD clinical evaluation methods?

Response

YES ☐  NO ☐  DON’T KNOW ☐

Question 4.3 – If so, which methodologies can be used to ensure that their clinical evaluation is robust? Which clinical trial outcomes would be impacted by these specific methodological features?

Response

☐ Type
☐ Duration
☐ Endpoints
☐ Sample size
☐ Randomisation method
☐ Outcome collection and analysis method
☐ Other

Justification
Specific methodological features of clinical evaluation of a connected medical device

Question 4.4 – Have these clinical trial methodologies previously been applied for some CMDs? If so, what are the study references?

Response

YES ☐ NO ☐ DON’T KNOW ☐

Justification

Section 5 – Mode of action and interactions of certain CMDs with environment of use

Question 5.1 – Should the effect of the CMD in isolation or the effect of the CMD and the environment of use be taken into account?

Response

Section 6 – Organisational aspects entailed with the use of certain CMDs (impacting patients, their quality of life and/or organisation of care)

Some CMDs require user training, user adoption of the CMD and engagement over time.

Question 6.1 – Are there any clinical trial methodologies that can be used to include these specificities? If so, which?

Response

YES ☐ NO ☐ DON’T KNOW ☐

Justification

Question 6.2 – Which clinical trial outcomes would be impacted by these specific methodological features?

Response

☐ Type
☐ Duration
☐ Endpoints
☐ Sample size
☐ Randomisation method
☐ Outcome collection and analysis method
☐ Other

Justification
Specific methodological features of clinical evaluation of a connected medical device

**Question 6.3** – Have these clinical trial methodologies previously been applied for some CMDs? If so, what are the study references?

**Response**

YES ☐  NO ☐  DON’T KNOW ☐

**Justification**

**Question 6.4** – Besides clinical morbi-mortality endpoints, other endpoints are key for assessing the effectiveness of some CMDs, such as satisfaction, usability, compliance and patient quality of life. Are there any specific endpoints / ratings for CMDs?

**Response**

YES ☐  NO ☐  DON’T KNOW ☐

**Justification**

**Question 6.5** – Of the endpoints cited above, could be envisaged to pre-define certain thresholds, such as minimum use thresholds, number of app logins, etc.?

**Response**

YES ☐  NO ☐  DON’T KNOW ☐

**Justification**

**Question 6.6** – Is it possible to stipulate a strict description of organisational procedures in respect of remote patient follow-up in the clinical trial protocol so as to establish a gold standard?

**Response**

YES ☐  NO ☐  DON’T KNOW ☐

**Section 7 – Other specific features applicable to CMDs**

**Question 7.1** – Have you identified other specificities applicable to some CMDs potentially influencing their clinical evaluation (e.g. associated with the type of algorithm used)?

**Response**

YES ☐  NO ☐  DON’T KNOW ☐

**Question 7.2** – If so, which?

**Response**
**Question 7.3** – Which clinical trial outcomes would be impacted by these specific features?

<table>
<thead>
<tr>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Type</td>
</tr>
<tr>
<td>☐ Duration</td>
</tr>
<tr>
<td>☐ Endpoints</td>
</tr>
<tr>
<td>☐ Sample size</td>
</tr>
<tr>
<td>☐ Randomisation method</td>
</tr>
<tr>
<td>☐ Outcome collection and analysis method</td>
</tr>
<tr>
<td>☐ Other</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES ☐ NO ☐ DON’T KNOW ☐</td>
</tr>
</tbody>
</table>

**Justification**

**Please provide your contact details:**

- Last name First name: [Field]
- Position: [Field]
- Entity: [Field]
- Telephone: [Field]
- E-mail: [Field]
**Annexe 4. Review grid**

Draft methodological guide to the specific features of clinical evaluation of a connected medical device (CMD) in view of its application for reimbursement

**Questionnaire sent to review group**

| Entity:…………………………… |
| Last name, First name:………….. |
| Position:……………………… |
| Telephone:……………………… |
| E-mail:…………………………… |

| COMMENTS ON FORM  
THE DOCUMENT WILL UNDERGO DTP FORMATTING |
<table>
<thead>
<tr>
<th>(+) points</th>
<th>(-) points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| COMMENTS ON “INTRODUCTION” (PAGES 4&amp;5) |</p>
<table>
<thead>
<tr>
<th>Comments:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<p>| COMMENTS ON “CMDs EVALUATED BY THE CNEDIMTS AND COVERED IN THIS METHODOLOGICAL GUIDE” (PAGE 6) |</p>
<table>
<thead>
<tr>
<th>Comments:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<p>| COMMENTS ON “CNEDIMTS EVALUATION CRITERIA FOR ACCESS TO REIMBURSEMENT OF A MD” (PAGE 7) |</p>
<table>
<thead>
<tr>
<th>Comments:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>
COMMENTS ON “CMD COMMON FEATURES” (PAGES 8 TO 10)

Comments:

COMMENTS ON “ADAPTATION OF THE CLINICAL DEVELOPMENT OF CMDS” (PAGES 11 TO 13)

Comments:

COMMENTS ON “HAS EVALUATION CHALLENGES” (PAGES 13 TO 15)

Comments:

COMMENTS ON “CONCLUSION” (PAGE 16)

Comments:
Annexe 5. List of guide review contributions

The guide was submitted to 64 individuals or entities concerned by this topic. We would like to express our thanks to the 25 reviewers listed below for their contribution to this guide.

- Christine Balagué, Institut Mines Telecom BS
- Laure Belaya, Engineer
- Pierre-Yves Benhamou, Expert clinician
- Jean-Jacques Carre, CD Healthcare consulting
- Mehdi Cheraitia, Neogia
- David Coulon, @Health
- Cécile Delval, Air liquide santé international
- Fabrice Denis, Expert clinician
- Claire Desforcés, Fédération française des diabétiques
- Anne-Aurélie Epis de Fleurian, Syndicat national de l’industrie des technologies médicales (SNITEM)
- Michel Galinier, Expert clinician
- Laëtitia Gambotti, French National Cancer Institute (INCa)
- Olivier Goeau-Brissonniere, Fédération des spécialités médicales
- Enguerrand Habran, Fédération hospitalière de France
- Stéphane Korsia-Meffre, Association France côlon
- Olivier Lalaude, Fédération des prestations de santé à domicile (FEDEPSAD)
- Fabien Leclercq, AFCROs DM
- Claude-Fabien Litre, Expert clinician
- Laure Millet, Institut Montaigne
- Gill Morisse, Quantmetry
- Eric Renard, Expert clinician
- Nicolas Roche, Société de pneumologie de langue française
- Michel Vicaire, Association des insuffisants respiratoires
- Mobin Yasini, DMD Santé
- Sarah Zohar, Methodologist
Annexe 6.  List of tables

Table 1. Data provided by INAHTA network agencies ................................................................. 12
Table 2. CMD categories assessed by the CNEDIMTS, following the submission of applications by companies ............................................................................................................. 14
Table 3. Functional classification according to intended use of technology ........................................ 22
Table 4. Evidence standards for functional classification tiers 3A and 3B ........................................ 24
Table 5. Example of mobile app evaluation strategy ........................................................................ 33
Table 6. Documentary search strategy ......................................................................................... 46
Annexe 7. List of figures

Figure 1. Study designs mentioned in the YHEC report, 2016 (16) .................................................................27
Figure 2. Study designs mentioned in the WHO guide, 2016 (9) .................................................................34
References


18. Greaves F, Joshi I, Campbell M, Roberts S, Patel N, Powell J. What is an appropriate level of evidence for a
Specific methodological features of clinical evaluation of a connected medical device


Participants – Working Group

The working group includes the following professionals:

► **Jean Paul Beregi**, Radiologist, Nîmes (30)
► **Alain Bernard**, Thoracic and cardiovascular surgeon, Dijon (21)
► **Pierre-Yves Boelle**, Methodologist, Paris (75)
► **Dominique Costagliola**, Methodologist, Paris (75)
► **Marie-Cécile Fournier**, Biostatistician, Nantes (44)
► **Marie-Christine Jaulent**, Methodologist, Paris (75)
► **Pascal Sellier**, Patient and healthcare user association member, Paris (75)
► **Franck Semah**, Neurologist, Lille (59)
► **Rodolphe Thiebaut**, Methodologist, Bordeaux (33)
► **Valéry Trosini Desert**, Respiratory Medicine specialist, Paris (75)

The members of the working group were appointed by the board of the CNEDiMTS.

Pursuant to decree No. 2004-1139 dated 26 October 2004 (art. R. 161-84 to R. 161-86 of the French Social Security Code), all the members of the working group completed a declaration of interest, mentioning direct or indirect links with any business or organisation involved in the scope of the responsibilities of the HAS. These declarations of interest were published on the HAS website.

The declarations of interest were reviewed as per the criteria of the HAS “Guidance on declarations of interest and management of conflicts of interest” (adopted by the HAS College on 24 July 2013). A summary table of declared interests was reviewed by the board of the CNEDiMTS, who made the final decision on the membership of the working group. The interests declared by the selected experts were all deemed by the board of the CNEDiMTS to be “non-major”.

The summary table of declared interests was set out and, if applicable, updated, based on the experts’ updated declarations of interest at the start of each working group meeting and at the presentation of the opinion of the working group to the CNEDiMTS.
# Fact sheet

<table>
<thead>
<tr>
<th>Title</th>
<th>Specific methodological features of clinical evaluation of a connected medical device (CMD). Preparatory report of the guide to specific features of clinical evaluation in view of its application for reimbursement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work method</td>
<td>Review of the literature, data from other evaluation agencies, applications submitted to the CNEDiMTS, clinical trial databases, consultation of various stakeholders and opinion of a multidisciplinary working group.</td>
</tr>
<tr>
<td>Purpose(s)</td>
<td>This assessment was aimed at identifying the specific methodological features of clinical evaluation of connected medical devices (CMDs) in view of their application for reimbursement.</td>
</tr>
<tr>
<td>Requested by</td>
<td>Ministry of Solidarity and Health.</td>
</tr>
<tr>
<td>Sponsor</td>
<td>Haute Autorité de santé.</td>
</tr>
<tr>
<td>Project management</td>
<td>Estelle Piotto-Peylan, Doctor of Pharmacy, Medical device assessment department (SED) project manager (Head of Department: Hubert Galmiche; Deputy: Corinne Collignon). Secretarial duties: Yakaré Tounkara, assistant.</td>
</tr>
<tr>
<td>Participants</td>
<td>See list of participants.</td>
</tr>
<tr>
<td>Documentary search</td>
<td>Conducted by Frédérique Pages, Head of Documentation-Public Information Department, assisted by Maud Lefevre, assistant archivist.</td>
</tr>
</tbody>
</table>
| Authors of justification | Estelle Piotto-Peylan, SED project manager.  
The clinical trial database (section 3.3) and some of the data obtained from applications submitted to the CNEDiMTS (section 3.2) were reviewed by Cyril Olivier, SED project manager.  
Under the responsibility of Hubert Galmiche (Head of Department) and Corinne Collignon (Deputy Head of Department). |
| Approval | Review by the Medical Device and Health Technology Evaluation Committee (CNEDiMTS): January 2019.  
Submission to the HAS College, for information purposes. |
| Other formats | No other formats. |
| Accompanying document | Guide to the specific features of clinical evaluation of a CMD in view of its application for reimbursement.  
The objective of this guide is to help companies manufacturing or operating a CMD to include clinical trials for determining the benefit of the CMD in view of its reimbursement by national solidarity in their development strategy. |