



HAUTE AUTORITÉ DE SANTÉ

ÉVALUATION DES DISPOSITIFS MÉDICAUX CONNECTÉS

Spécificités méthodologiques d'évaluation clinique d'un dispositif médical connecté (DMC)

Rapport d'élaboration du guide sur ses spécificités d'évaluation
clinique, en vue de son accès au remboursement

Date de validation par la CNEDiMTS: 29 janvier 2019

Ce dossier est téléchargeable sur

www.has-sante.fr

Haute Autorité de santé

Service communication – information

5, avenue du Stade de France – F 93218 Saint-Denis La Plaine Cedex

Tél. : +33 (0)1 55 93 70 00

Comment citer ce document :

Haute Autorité de santé. Spécificités méthodologiques d'évaluation clinique d'un dispositif médical connecté. Rapport d'élaboration du guide. Saint-Denis La Plaine : HAS ; 2019.

Sommaire

Introduction	6
1. Périmètre des dispositifs médicaux concernés	7
1.1 Produits concernés.....	7
1.2 Produits non concernés.....	8
2. Méthode.....	9
2.1 Recherche documentaire et critères de sélection des articles	9
2.2 Résultats de la recherche.....	10
3. Évaluation – Analyse des données	12
3.1 Données transmises par les agences d'évaluation technologique dans les autres pays	12
3.2 Données issues des dossiers ayant été soumis à la CNEDiMTS	14
3.3 Données à partir des bases de données d'essais cliniques.....	19
3.4 Données de la littérature	22
4. Conclusion générale et position du groupe de travail	45
5. Annexes.....	47
Annexe 1. Méthode de travail	48
Annexe 2. Recherche documentaire.....	49
Annexe 3. Questionnaire envoyé aux parties prenantes	52
Annexe 4. Grille de lecture.....	61
Annexe 5. Liste des contributions à la relecture du guide	63
Annexe 6. Liste des tableaux.....	64
Annexe 7. Liste des figures.....	65
Références	66
Participants – Groupe de travail.....	68
Fiche descriptive.....	69

Abréviations et acronymes

ANSM	Agence nationale de sécurité du médicament et des produits de santé
CNAMTS	Caisse nationale de l'assurance maladie des travailleurs salariés
CNEDiMTS	Commission nationale d'évaluation des dispositifs médicaux et des technologies de santé
CSS	Code de la sécurité sociale
DGOS	Direction générale de l'offre des soins
DGS	Direction générale de la santé
DM	Dispositif médical
DMC	Dispositif médical connecté
DSS	Direction de la sécurité sociale
ECR	Essai contrôlé randomisé
EMA	<i>European medicines agency</i>
FDA	<i>Food and Drug Administration</i>
FEDEPSAD	Fédération des prestataires de santé à domicile
HAS	Haute Autorité de santé
HTA	<i>Health Technology Assessment</i>
INAHTA	<i>International Network of Agencies for Health Technology Assessment</i>
INCa	Institut national du cancer
LPPR	Liste des produits et prestations remboursables
MOST	<i>Multiphase Optimisation Strategy</i>
MSA	Mutualité sociale agricole
NICE	<i>National Institute for Health and Care Excellence</i>
NHS	<i>National Health Service</i>
OMS	Organisation mondiale de la santé
PREMs	<i>Patient Reported Experience Measures</i>
PROMs	<i>Patient Reported Outcome Measures</i>
RSI	Régime social des indépendants
SED	Service évaluation des dispositifs (de la HAS)
SIDIV	Syndicat de l'industrie du diagnostic in vitro
SNITEM	Syndicat national de l'industrie des technologies médicales
UNPDM	Union nationale des prestataires de dispositifs médicaux
UPSADI	Union des prestataires de santé à domicile indépendants
SMART	<i>Sequential Multiple Assignment Randomised Trials</i>
SNADOM	Syndicat des associations d'assistance à domicile
YHEC	<i>York Health Economic Consortium</i>

Glossaire

Algorithme

« Description d'une suite finie et non ambiguë d'étapes ou d'instructions permettant d'obtenir un résultat à partir d'éléments fournis en entrée. » (1)

Un algorithme apprenant a la capacité de faire évoluer au cours du temps et de façon autonome les paramètres des instructions qui le composent en fonction des résultats qu'il a obtenus antérieurement. Il s'oppose aux algorithmes classiques qui n'ont pas la capacité de se modifier eux-mêmes.

Apprentissage machine supervisé

« L'algorithme apprend de données d'entrée qualifiées par l'humain et définit ainsi des règles à partir d'exemples qui sont autant de cas validés. » (1)

L'apprentissage supervisé est composé de deux étapes. Au cours de la première, l'algorithme apprend, à partir de cas résolus, à ajuster ses paramètres pour maximiser ses performances de prédiction (e.g. classification des cas). La deuxième étape consiste à vérifier que l'algorithme obtenu est généralisable en l'appliquant à un ensemble de nouveaux cas, le résultat pour ces cas devant être validé par un expert. On vérifie ainsi si l'algorithme a appris correctement, si l'apprentissage a été biaisé (surapprentissage) ou, à l'inverse, s'il a besoin de plus d'exemples (sous-apprentissage).

Apprentissage machine non supervisé

« L'algorithme apprend à partir de données brutes et élabore sa propre classification qui est libre d'évoluer vers n'importe quel état final lorsqu'un motif ou un élément lui est présenté. Pratique qui nécessite que des instructeurs apprennent à la machine comment apprendre. » (1)

Dans le cas de l'apprentissage non supervisé, les données fournies à l'algorithme sont des cas non résolus (on ne connaît pas la conclusion). L'algorithme s'appuie sur des fonctions de similarité, de distance, etc. entre les cas pour les regrouper en classes. L'expert vérifie *a posteriori* que les classes obtenues classifient bien les données (en comparant le résultat de l'algorithme et sa propre décision sur de nouvelles données) afin de décider si l'algorithme est valide et peut être utilisé en vie réelle.

Apprentissage machine (ou apprentissage automatique, machine learning)

« Branche de l'intelligence artificielle, fondée sur des méthodes d'apprentissage et d'acquisition automatique de nouvelles connaissances par les ordinateurs, qui permet de les faire agir sans qu'ils aient à être explicitement programmés. » (1)

Les principales méthodes sont dites supervisées ou non supervisées. Ces méthodes peuvent être combinées entre elles. Lors de son déploiement en situation réelle, l'algorithme peut être temporairement figé, puis mis à jour séquentiellement et régulièrement par les personnes chargées de son développement selon une gestion de versions établie à l'avance. L'algorithme peut également ne pas être figé et se mettre à jour dynamiquement lors de son utilisation et du traitement de nouvelles données.

Big data

« Désigne la conjonction entre, d'une part, d'immenses volumes de données devenus difficilement traitables à l'heure du numérique et, d'autre part, les nouvelles techniques permettant de traiter ces données, voire d'en tirer par le repérage de corrélations des informations inattendues. » (1)

Le développement du numérique a permis de générer, dans le contexte du développement d'Internet, d'immenses volumes de données. Le développement parallèle de nouvelles approches de traitement de données et d'augmentation de la puissance des ordinateurs a généré un courant international pour le traitement massif de données hétérogènes appelé « big data ».

Essai de régression

« Essai exigé pour s'assurer qu'un changement d'un composant du système n'a pas altéré la fonctionnalité, la fiabilité ou les performances et n'a pas entraîné de défauts supplémentaires. »¹

¹ Norme ISO62304 <https://www.iso.org/obp/ui/#iso:std:iec:62304:ed-1:v1:fr> [consulté le 21/12/2018].

Introduction

La Commission nationale d'évaluation des dispositifs médicaux et des technologies de santé (CNEDiMTS) est la commission de la HAS qui évalue notamment les dispositifs médicaux (DM) et autres produits de santé en vue de leur remboursement par l'Assurance maladie. Son rôle est de donner aux autorités de santé un avis consultatif recommandant ou non la prise en charge des DM (inscription sur la liste des produits et prestations remboursables ou LPPR), de contribuer à la détermination des conditions de leur bon usage et leur place dans la stratégie thérapeutique, diagnostique ou de prévention. Le cas échéant, elle se prononce sur les conditions permettant d'optimiser l'utilisation de la technologie en matière de compétence de l'utilisateur et d'environnement nécessaire².

Comme pour toute autre catégorie de DM, tout dispositif médical connecté (DMC) éligible à la LPPR peut faire l'objet, sur initiative de l'entreprise qui en assure l'exploitation, d'une demande d'inscription sous nom de marque. Une telle demande doit être appuyée par des données permettant d'apprécier l'intérêt du produit et son intérêt de santé publique, soutenant l'argumentaire de Service Attendu (SA) et d'Amélioration du Service Attendu (ASA). En effet, ces critères d'évaluation réglementaires³ s'imposent à la CNEDiMTS, quel que soit le type de dispositif médical.

Toutefois, les DMC peuvent avoir des particularités liées notamment à leur mode d'action, à leur impact sur les patients, les aidants, les professionnels ou l'organisation des soins. Leur développement clinique doit prendre en compte ces particularités.

L'objectif de ce projet entrepris en 2018 est d'identifier les éventuelles spécificités méthodologiques dans l'évaluation clinique des DMC, afin de permettre aux industriels qui déposent un dossier à la CNEDiMTS de les intégrer dans leur stratégie de développement clinique et de préparer leur prise en charge par la solidarité nationale. Ces travaux s'inscrivent dans le programme de travail de la HAS⁴.

Un « **Guide sur les spécificités d'évaluation clinique d'un dispositif médical connecté (DMC) en vue de son accès au remboursement** » a été produit. Ce guide est disponible et téléchargeable sur le site de la HAS (www.has-sante.fr).

Ce rapport, complémentaire du guide, détaille la méthode d'élaboration du guide.

² Haute Autorité de santé. Règlement intérieur de la commission nationale d'évaluation des dispositifs médicaux et des technologies de santé. Saint-Denis La Plaine : HAS ; 2018. https://www.has-sante.fr/portail/upload/docs/application/pdf/reglement_cepp.pdf [consulté le 21/12/2018].

³ [Article R. 165-11 du Code de la sécurité sociale](#) [consulté le 21/12/2018].

⁴ https://www.has-sante.fr/portail/upload/docs/application/pdf/2018-02/programme_de_travail_has_vd.pdf [consulté le 21/12/2018].

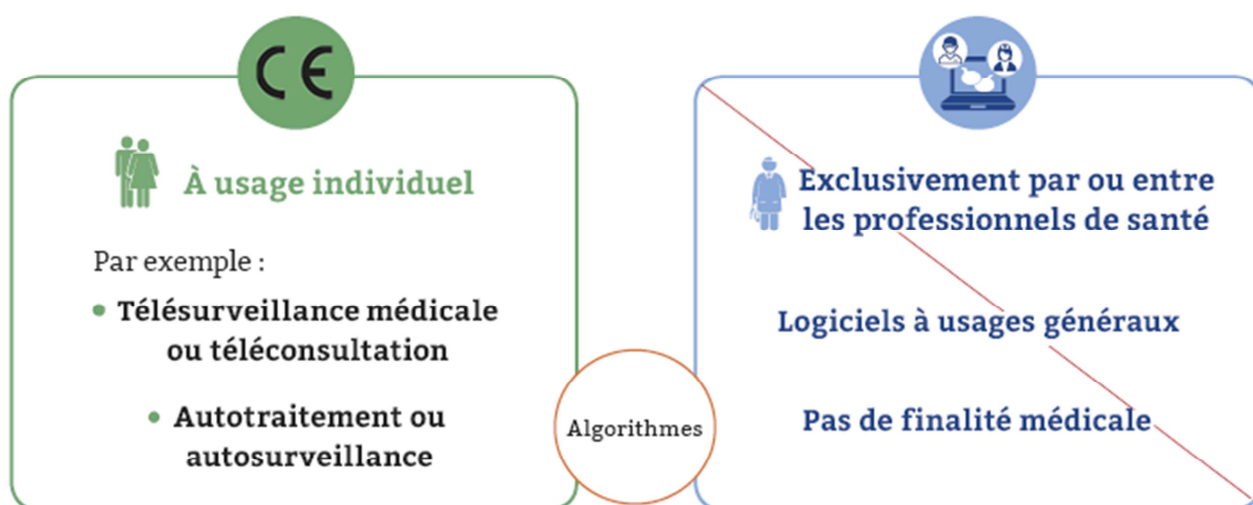
1. Périmètre des dispositifs médicaux concernés

1.1 Produits concernés

Le champ des dispositifs médicaux connectés est vaste, mais en raison des missions de la CNEDiMITS, les dispositifs médicaux (DM) qu'elle évalue ne représentent qu'une **étroite partie** des dispositifs médicaux connectés (DMC).

Le périmètre de ces travaux est totalement sous-tendu au périmètre des DM évalués par la CNEDiMITS, c'est-à-dire ceux qui sont marqués CE et sont candidats à un financement individualisé par l'Assurance maladie (en sus des actes ou forfaits de soins existants).

Dispositifs médicaux connectés évalués par la CNEDiMITS



Les DMC évalués par la CNEDiMITS remplissent les quatre conditions suivantes.

1. Ils sont destinés à être utilisés à des **fins médicales**⁵, leur finalité impliquant qu'ils soient marqués CE.
2. Ils sont à **usage individuel** (implantés ou utilisés par le patient lui-même).
3. Ils disposent d'une **fonction de télécommunication**.
4. **L'industriel a déposé un dossier de demande de prise en charge par la solidarité nationale.**

C'est cette quatrième condition qui déclenche l'évaluation d'un DMC par la CNEDiMITS : l'initiative de l'inscription d'une nouvelle technologie sur la liste des produits et prestations remboursables (LPPR) appartient à l'industriel (article **R. 165-7 du CSS**).

Pour illustrer ce périmètre, quelques exemples de DMC susceptibles de faire l'objet d'un remboursement individualisé par l'Assurance maladie ; des DM :

- utilisés à des **fins de télésurveillance médicale** (par exemple, un défibrillateur cardiaque implantable associé à une télésurveillance médicale) ;

⁵ Fins médicales visées dans la définition du dispositif médical : [article 2 du règlement \(UE\) 2017/745 du Parlement européen et du Conseil du 5 avril 2017 relatif aux dispositifs médicaux.](#)

- générant une action du patient à des **fins d'autotraitement ou d'autosurveillance** (par exemple, un neurostimulateur utilisé dans le traitement de la douleur connecté avec une application smartphone qui permet au patient de gérer lui-même son traitement) ;
- générant ou recevant des informations en vue de l'optimisation du traitement (par exemple, une pompe à insuline couplée à un capteur de mesure en continu du glucose interstitiel exploitant le carnet électronique du patient pour optimiser son traitement).

1.2 Produits non concernés

De nombreux produits utilisés en santé, bien qu'ils soient connectés, ne sont pas concernés par ces travaux :

- **les applications et objets connectés qui ne sont pas des dispositifs médicaux marqués CE** (par exemple, des logiciels ou applications destinés à augmenter son activité physique en calculant le nombre de pas par jour) ;
- **les dispositifs médicaux qui ne sont pas à usage individuel et qui ne font pas l'objet d'un financement individualisé** (par exemple, les balances, thermomètres ou tensiomètres connectés utilisés dans les établissements de santé qui ne sont pas destinés à un seul patient) ;
- **les dispositifs médicaux utilisés exclusivement par un professionnel de santé ou entre professionnels de santé** (par exemple, les dispositifs d'aide à la décision professionnelle, les logiciels d'aide à la prescription ou à la dispensation, ceux de télé-expertise, les dispositifs d'imagerie médicale d'aide au diagnostic ou à la décision thérapeutique, etc.) ;
- **les logiciels destinés à des usages généraux, même lorsqu'ils sont utilisés dans un environnement de soins** (par exemple, les logiciels de gestion administrative).

2. Méthode

La méthode mise en œuvre repose sur :

- une enquête conduite par questionnaire sur les méthodes et exigences mises en œuvre dans d'autres pays ;
- l'analyse des dossiers ayant été soumis à la CNEDiMITS concernant ce type de technologie ;
- l'analyse des bases de données d'essais cliniques ;
- l'analyse critique des données de la littérature scientifique ;
- la position d'experts réunis dans un groupe de travail.

Le groupe de travail est constitué des services de la HAS et d'experts sans conflit d'intérêts majeur (au sens du « Guide des déclarations d'intérêts et de gestion des conflits d'intérêts » de la HAS) choisis au regard de leur expérience en matière d'évaluation des dispositifs médicaux et de leur expertise en e-santé ou en méthodologie.

En amont du groupe de travail, pour le cadrage de ce projet et en relecture du guide qui en découle, des parties prenantes ont été sollicitées, parmi :

- les représentants des fabricants et des distributeurs ;
- la fédération des spécialités médicales et les sociétés savantes spécialisées en e-santé dans le secteur des dispositifs médicaux ;
- les prestataires de services et distributeurs de matériel à domicile, spécialisés en e-santé dans le secteur des dispositifs médicaux ;
- les associations de patients concernées par les expérimentations actuellement en cours ou ayant mis en place des plateformes d'évaluation ;
- les groupes de réflexion ou experts spécialisés en e-santé.

Ces travaux ont été analysés et validés par la CNEDiMITS puis présentés au Collège de la HAS, pour information. La Commission peut proposer des modifications complémentaires aux travaux élaborés par le groupe d'experts.

Le détail de cette méthode est décrit en annexe 1.

2.1 Recherche documentaire et critères de sélection des articles

2.1.1 Recherche documentaire

La recherche documentaire a été limitée aux publications en langue anglaise et française. Elle a porté sur la période de janvier 2013 à octobre 2018. Une veille a été réalisée jusqu'en décembre 2018 sur les sites Internet consultés.

Les évaluations technologiques, les recommandations, les conférences de consensus, les guides, les méta-analyses, les revues systématiques, les études cliniques ont été recherchés.

La recherche a été réalisée en interrogeant plusieurs bases de données bibliographiques, dont *Medline* et divers sites Internet tels que la *Cochrane Library*, les sites Internet publiant des recommandations, les sites Internet publiant des rapports d'évaluation technologique, les sites Internet des sociétés savantes compétentes dans le domaine étudié.

La stratégie de recherche et la liste des sources interrogées sont détaillées dans l'annexe 2.

Cette recherche automatisée a été complétée par une recherche manuelle réalisée à partir des références citées dans les articles retenus. La bibliographie des experts et les données des industriels ont également été considérées comme sources d'information. Lorsque les études transmises

par ces derniers répondaient aux critères de sélection définis pour la recherche documentaire, elles ont été intégrées à la bibliographie.

2.1.2 Critères de sélection

Pour être sélectionnées, les données devaient répondre aux critères suivants.

► Type d'étude

Recommandations, guides ou rapports d'évaluation technologique, méta-analyses, revues systématiques, études cliniques.

► Objectif principal

Évaluer des dispositifs médicaux connectés.

Les recommandations, revues systématiques, guides ou rapports d'évaluation technologique devaient identifier des méthodologies d'études cliniques pour les dispositifs médicaux connectés.

Les études cliniques ou méta-analyses devaient utiliser une méthodologie d'évaluation clinique spécifique autre que celle habituellement décrite dans les dossiers d'évaluation pour les différentes catégories de dispositifs médicaux soumis à la CNEDiMTS.

► Thématique traitée ou dispositifs utilisés

Dispositifs médicaux connectés rentrant dans le champ de ces travaux (partie 1.1).

► Population étudiée

Patients bénéficiant d'un dispositif médical connecté.

► Critères de jugement

Tous les critères de jugement étaient pris en compte, dès lors qu'ils permettaient l'évaluation de l'intérêt individuel et qu'ils étaient spécifiques aux dispositifs médicaux connectés.

► Critères d'exclusion

Ont été exclus de l'analyse les articles généraux, articles narratifs de type éditorial ou avis d'auteurs, les études *in vitro*, biomécaniques, histologiques, menées chez l'animal, les études évaluant des DMC sans spécificités cliniques (en termes de design, critères de jugement, durée, effectifs, etc.), les doublons d'étude.

► Période de sélection

De 2013 à 2018.

2.2 Résultats de la recherche

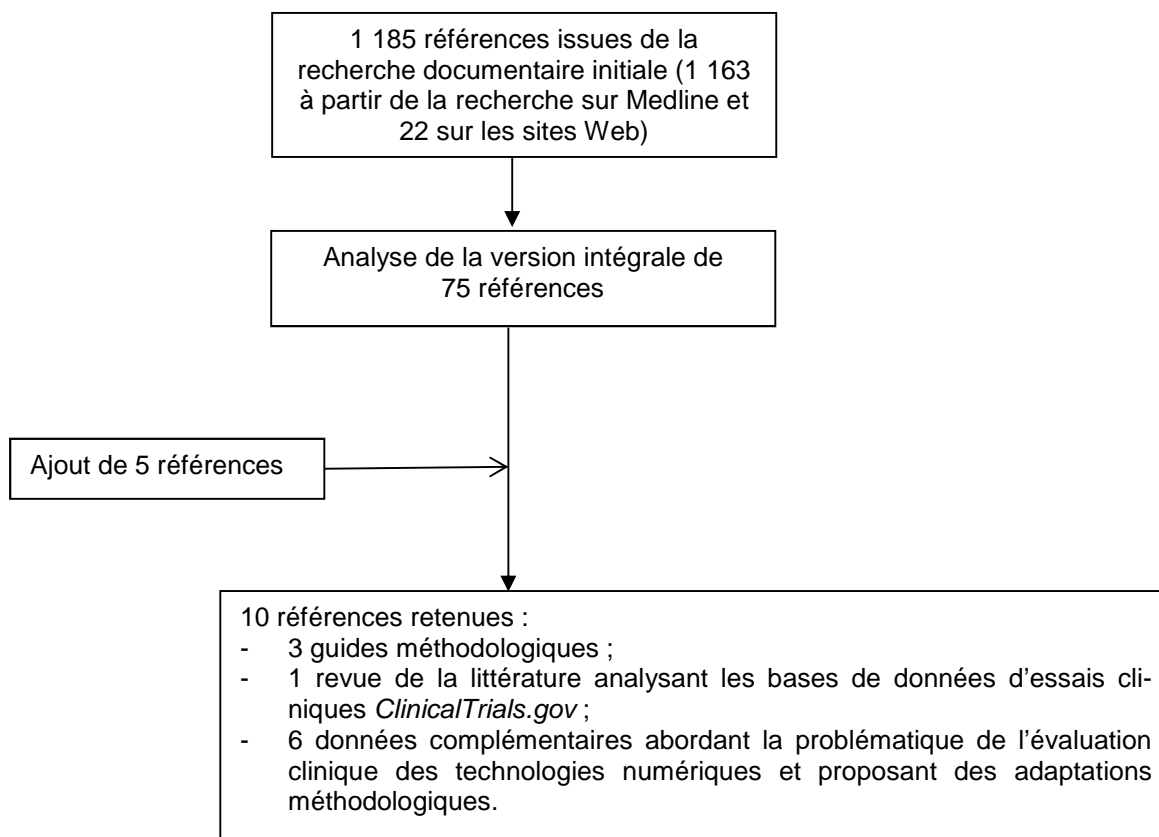
Au total, 1 163 références bibliographiques ont été identifiées à partir de la recherche systématique effectuée sur *Medline* et 22 sur les sites Web. Parmi ces références, 75 ont été analysées et 10 ont été retenues.

Le processus de sélection est illustré dans la figure ci-dessous.

Ces références se répartissent ainsi :

- 3 guides méthodologiques ;
- 1 revue de la littérature analysant les bases de données *ClinicalTrials.gov* ;
- 6 données complémentaires abordant la problématique de l'évaluation clinique des technologies numériques et proposant des adaptations méthodologiques.

Organigramme de sélection des articles



3. Évaluation – Analyse des données

Les données cliniques, issues de la littérature ou fournies par les parties prenantes, ont été sectionnées selon les critères exposés dans le paragraphe 2.1.2, et portent sur les spécificités méthodologiques d'évaluation clinique des dispositifs médicaux connectés.

3.1 Données transmises par les agences d'évaluation technologique dans les autres pays

Une enquête a été conduite au sein du [réseau INAHTA](#) – *International Network of Agencies for Health Technology Assessment* (49 agences d'évaluation – 30 pays).

Les données transmises par les autres agences d'évaluation technologique à l'étranger ont été analysées.

Les items sur lesquels portait le questionnaire étaient les suivants :

- champ de l'évaluation et types de DMC concernés ;
- nombre de DMC évalués ;
- identification de spécificités éventuelles lors de l'évaluation clinique des DMC, par rapport à l'évaluation des autres dispositifs médicaux ;
- existence de guidelines sur les DMC (recommandations d'un design particulier ou guide sur les méthodes d'évaluation) ;
- prévision ou planification d'une adaptation des méthodes d'évaluation des DMC.

Les réponses des 12 agences issues de 8 pays ayant répondu sont détaillées dans le Tableau 1.

Tableau 1. Données transmises par les agences du réseau INAHTA

Agences	Réponses
JAZMP (Slovénie) <i>Javna agencija Republike Slovenije za zdravila in medicinske pripomočke</i>	Aucune agence n'est responsable de l'évaluation des dispositifs médicaux. Aucune expérience en termes d'évaluation clinique des DMC.
RedETS (Espagne) <i>Red Española de Agencias de Evaluación de Tecnologías Sanitarias y Prestaciones del Sistema Nacional de Salud</i> SESCS (Espagne) <i>Servicio Evaluación del Servicio Canario de la Salud (HTA unit of the Regional Government of the Canary Islands)</i>	En Espagne, les dispositifs médicaux sont évalués par le réseau des agences HTA (RedETS) qui coordonne 8 agences HTA régionales (AQUAS, SESCO, AETSA, OSTEB, AVALIA-T, IACS, AETS, UETS-M). Le champ de leur évaluation intègre tous les types de DM (utilisés par le patient ou entre professionnels de santé). Trois documents publiés en espagnol uniquement sont transmis. Ils concernent « <i>Les technologies de l'information et communication en santé mentale (télépsychiatrie)</i> », « <i>Téléneurologie vs. consultation en face-à-face. Évaluation de son efficacité et de son rapport coût-efficacité</i> », « <i>Conception, évaluation et mise en œuvre de la télémédecine.</i> » Aucune spécificité en termes d'évaluation clinique n'est identifiée et aucune adaptation méthodologique n'est envisagée.
AETS (Espagne) <i>Agencia de Evaluación de Tecnologías Sanitarias</i>	Aucune donnée relative à l'évaluation clinique d'un DMC n'est disponible. Aucune spécificité en termes d'évaluation clinique n'est identifiée, l'évaluation clinique est similaire quel que soit le dispositif médical. Aucun guideline n'est disponible.
AQuAS (Espagne) <i>Agència de Qualitat i Avaluació Sanitàries de Catalunya</i>	Certains DM ont été évalués, tels que les dispositifs d'assistance ventriculaire ou défibrillateurs cardiaques implantables. Aucune spécificité en termes d'évaluation clinique n'est identifiée et aucune adaptation méthodologique n'est envisagée. Aucun guideline n'est disponible.
ASSR (Italie) <i>Agencia Sanitaria e Sociale Regionale dell'Emilia-Romagna (Regional Agency for Health and Social Care),</i>	Les DM connectés peuvent être évalués au niveau des agences HTA régionales ou au niveau national. Il n'y a pas d'agence spécifique évaluant les DM connectés. Le champ de l'évaluation intègre tous les types de DM (utilisés par le patient ou entre professionnels de santé). Aucune évaluation de DMC n'est transmise. Aucun guideline n'est disponible.

Agences	Réponses
<p>IQWiG (Allemagne) <i>Institute for Quality and Efficiency in Health Care</i></p>	<p>Le G-BA (Federal Joint Committee) est responsable de l'évaluation des nouvelles technologies. L'IQWiG est mandaté par le G-BA pour en évaluer le bénéfice. Le G-BA évalue les nouvelles technologies dans des indications spécifiques, l'utilisation large et non spécifique des dispositifs utilisés dans le cadre de la télésanté ne relève pas du G-BA et d'IQWiG. Un rapport préliminaire ne montrant pas de bénéfice de la télésurveillance est cité. Il concerne la télésurveillance des dispositifs cardiaques implantables actifs chez des patients ayant une arythmie ventriculaire et une insuffisance cardiaque. Aucune spécificité en termes d'évaluation clinique n'est identifiée, la méthodologie standard s'applique quel que soit le dispositif médical. Il n'y a pas de besoin d'ajustement méthodologique spécifique. Aucun guideline n'est disponible.</p>
<p>SV (Autriche) <i>Hauptverband der österreichischen Sozialversicherungsträger</i></p>	<p>Aucune agence n'est spécifiquement responsable de l'évaluation des DMC. Aucune expérience en termes d'évaluation clinique des DMC. Aucun guideline n'est disponible.</p>
<p>GOeG (Autriche) <i>Gesundheit Österreich GmbH,</i></p>	<p>Il n'y a pas de document spécifique sur les DMC. Les méthodes d'évaluation sont identiques à celles des autres DM. Aucune adaptation des méthodes d'évaluation n'est envisagée.</p>
<p>LBI-HTA (Autriche) <i>Ludwig Boltzmann Institute for Health Technology Assessment</i></p>	<p>LBI-HTA a évalué des technologies dans le cadre de la télécardiologie et télédermatologie. Une évaluation de la télémédecine à l'horizon 2020 est prévue. Il existerait un besoin de réaliser des recommandations sur l'évaluation clinique des DMC. Néanmoins, les méthodes d'évaluation des applications/services en Autriche sont les mêmes que celles des autres DM (PICO), avec un focus sur les données organisationnelles. Aucun guideline n'est disponible. Aucune adaptation des méthodes d'évaluation n'est envisagée.</p>
<p>DEFACTUM (Danemark) <i>Social & Health Services and Labour Market</i></p>	<p>Aucune agence spécifique n'est responsable des DMC. DEFACTUM évalue les DM, en collaboration avec différentes organisations professionnelles. Le champ de l'évaluation intègre tous les types de DM (utilisés par le patient ou entre professionnels de santé). Les spécificités liées à l'évaluation d'un DMC portent sur le fait qu'il s'agit de l'évaluation d'un ensemble, d'une chaîne de technologies dépendantes d'une organisation à mettre en place (et non d'un seul dispositif connecté). L'implémentation d'une technologie connectée varie souvent, ce qui rend l'évaluation plus complexe. Le contexte et l'organisation mise en place sont des éléments importants à prendre en compte dans l'évaluation. Ainsi, les composants de la chaîne peuvent s'influencer différemment selon les groupes de patients. Aucun guideline n'est disponible. Aucune adaptation des méthodes d'évaluation n'est envisagée.</p>
<p>NCPHA (Bulgarie) <i>National Center of Public Health and Analyses</i></p>	<p>Aucune agence spécifique n'est responsable des DMC. Les spécificités liées à l'évaluation d'un DMC portent sur leur impact indirect sur le comportement des patients. Aucun guideline n'est disponible. Aucune adaptation des méthodes d'évaluation n'est envisagée.</p>
<p>SFOPH (Suisse) <i>Swiss Federal Office of Public Health</i></p>	<p>En Suisse, les DMC utilisés à usage personnel à des fins médicales mais également au travers d'un acte professionnel sont évalués, en vue d'une décision de remboursement par l'assurance maladie. Les évaluations sont menées par le SFOPH puis validées par des comités d'évaluation permanents qui feront des recommandations à l'organe de décision (ministre de la Santé). Jusqu'à présent, deux dispositifs ont été évalués : un pour la surveillance de l'arythmie cardiaque (télémédecine) et un pour l'autosurveillance des taux de glucose chez les patients diabétiques. La même méthodologie est appliquée pour les évaluations des autres technologies et il n'y a pas de projet d'élaboration de recommandations ou guidelines.</p>

Conclusion

Les réponses apportées par les autres agences d'évaluation sollicitées ne permettent pas d'identifier de spécificités méthodologiques pour l'évaluation clinique des dispositifs médicaux connectés. Cette enquête ne fait ressortir aucune recommandation spécifique existante ni perspective d'adaptation méthodologique envisagée. Les méthodes d'évaluation des DMC seraient identiques à celles des autres dispositifs médicaux.

Parmi les éléments de réponses issus de l'enquête, la complexité de l'évaluation d'un DMC a été soulignée, en raison de son impact organisationnel et de son impact sur les patients. Ces réponses ont également fait ressortir que l'évaluation du DMC doit prendre en compte un ensemble de composants en constante interaction et dépendants d'une organisation à mettre en place, chaque élément du système pouvant s'influencer différemment selon les groupes de patients.

3.2 Données issues des dossiers ayant été soumis à la CNEDiMTS

Un certain nombre de catégories de dispositifs médicaux connectés ont déjà évaluées par la CNEDiMTS, suite à des dépôts de dossiers par les industriels. Ces catégories de DMC sont détaillées dans le tableau 2. À noter, à ce jour, aucun dossier ne mentionnait de DMC faisant appel à un traitement de données par intelligence artificielle.

Tableau 2. Catégories de DMC évaluées par la CNEDiMTS suite à des dépôts de dossiers par les industriels

Domaine d'application	Catégories de DMC	Indications générales/Commentaires
Appareillage loco-moteur	Dispositif d'assistance électrique à la propulsion pour fauteuil roulant manuel	<p><i>Indications :</i> Patients dont les capacités cognitives permettent d'assurer la maîtrise du système de propulsion assistée et qui, bien que capables de propulser eux-mêmes un fauteuil roulant manuel, ont besoin pour des raisons médicales d'assistance électrique à la propulsion de façon intermittente ou définitive.</p> <p><i>Particularités :</i></p> <ul style="list-style-type: none"> - commande par un bracelet « Pushtracker » attaché au poignet de l'utilisateur qui communique par Bluetooth avec l'unité motrice au moment de l'activation/désactivation de l'assistance électrique ; - l'ensemble des informations peut être visualisé à partir d'une application (moyenne quotidienne de poussées, nombre de km, etc.).
Audiologie	Implants à ancrage osseux (prothèse auditive ostéo-intégrée)	<p><i>Indications :</i></p> <ul style="list-style-type: none"> - surdités de transmission ou surdités mixtes pour lesquelles la chirurgie d'oreille moyenne ne peut être réalisée et l'appareillage traditionnel par voie aérienne ou osseuse est inefficace ou impossible ; - surdités neurosensorielles unilatérales au moins sévères. <p><i>Mode d'action :</i></p> <ul style="list-style-type: none"> - le processeur vocal externe transforme la pression acoustique en une force d'intensité variable ; - cette force est retransmise du pilier transcutané à l'implant ostéo-intégré ; - l'implant génère une déformation élastique de la corticale de l'os en fonction de la fréquence ; - cette vibration se répercute sur l'os temporal dans lequel se trouve la capsule labyrinthique qui comprend l'oreille interne. <p>Présence de data-logging (enregistreur de données).</p>
	Implant d'oreille moyenne	<p><i>Indications :</i></p> <p>Surdités de transmission ou surdités mixtes, unilatérales ou bilatérales, chez l'enfant et chez l'adulte, après échec ou impossibilité :</p> <ul style="list-style-type: none"> - d'une chirurgie d'oreille moyenne ; - et d'un appareillage traditionnel par voie aérienne ou osseuse ; - et d'une prothèse auditive ostéo-intégrée.

Domaine d'application	Catégories de DMC	Indications générales/Commentaires
		<p><i>Mode d'action :</i> Le principe repose sur la transmission directe des sons à une structure vibratoire de l'oreille moyenne par un transducteur électromagnétique, le FMT. L'audioprocasseur capte les sons par un microphone, traite le signal, et transmet l'information par voie transcutanée au composant implanté. Le récepteur interne VORP transfère le signal électrique via la liaison conductrice, au transducteur de sortie, le FMT. Le FMT convertit le signal électrique en vibrations mécaniques transmises à l'oreille moyenne. Ces vibrations sont ensuite transmises à l'oreille interne.</p>
	Implants du tronc cérébral	<p><i>Indications :</i> Surdités neurosensorielles (surdité de perception) bilatérales sévères à profondes, après échec ou inefficacité d'un appareillage acoustique conventionnel. Les indications sont limitées aux circonstances où la cause de la surdité ne permet pas l'implantation cochléaire : - soit du fait de l'exérèse des tumeurs (schwannome vestibulaire bilatéral) dans le cadre de la neurofibromatose de type 2 ; - soit en cas d'anomalies anatomique : ossification cochléaire bilatérale totale, malformations cochléaires majeures, fracture bilatérale des rochers, schwannome vestibulaire et surdité controlatérale. <i>Mode d'action :</i> Le principe de l'implant du tronc cérébral est d'émettre des impulsions électriques spécifiques destinées à stimuler directement les structures nerveuses de l'oreille interne de patients atteints de surdité neurosensorielle. La partie externe, composée d'un processeur, permet de traiter, numériser, coder et transmettre à l'implant les informations acoustiques extérieures reçues par un microphone. Les implants du tronc cérébral ont bénéficié des technologies développées pour les implants cochléaires, la stimulation se faisant dans ce cas sur le noyau cochléaire et non sur les cellules ciliées.</p>
	Implants cochléaires	<p><i>Indications :</i> Surdités neurosensorielles (surdité de perception) bilatérales sévères à profondes, après échec ou inefficacité d'un appareillage acoustique conventionnel. L'implantation cochléaire bilatérale chez l'enfant est également prise en charge pour les surdités neurosensorielles (surdités de perception) bilatérales profondes, après échec ou inefficacité d'un appareillage acoustique conventionnel. <i>Mode d'action :</i> Le processeur de son permet de capter les sons par un microphone multidirectionnel. Ce signal est traité, numérisé, codé et est transmis au niveau de l'antenne à la partie interne de l'implant par couplage inductif. Ce signal modulé parcourt les électrodes de l'implant cochléaire ou du tronc cérébral qui stimulent respectivement les fibres du nerf auditif ou le noyau cochléaire. Les potentiels d'action générés sont interprétés par le cerveau comme des sons.</p>
Cardiologie/rythmologie	<p>Système de télésurveillance (de défibrillateur cardiaque implantable, stimulateur ou de patients insuffisants cardiaque)</p> <p>Défibrillateur cardiaque implantable, Stimulateur cardiaque implantable</p>	<p><i>Indications :</i> Pour les systèmes de télésurveillance : surveillance des stimulateurs/défibrillateurs dans les indications ci-dessous.</p> <p>Pour les stimulateurs cardiaques implantables simple chambre : - stimulateur monochambre atrial asservi : dysfonction du nœud sinusal avec insuffisance chronotrope, si la conduction auriculo-ventriculaire est normale en l'absence de toute cardiopathie évolutive ; - stimulateur monochambre ventriculaire asservi : bloc auriculo-ventriculaire du second ou du troisième degré avec insuffisance chronotrope ; si l'oreillette n'est pas stimulable ou non détectable de façon prédominante.</p> <p>Pour les stimulateurs cardiaques implantables double chambre : Bloc auriculo-ventriculaire du 2^d ou du 3^e degré nécessitant une stimulation ventriculaire permanente ou intermittente (si la conduction auriculo-ventriculaire est préservée) : - chez le patient ayant une insuffisance chronotrope, lorsque le maintien d'une activité physique est possible, et si l'oreillette est stimulable de façon prédominante ; - dysfonction du nœud sinusal avec insuffisance chronotrope associée soit à</p>

Domaine d'application	Catégories de DMC	Indications générales/Commentaires
		<p>des anomalies de la conduction auriculo-ventriculaire, soit à une maladie rythmique atriale.</p> <p>Pour les défibrillateurs cardiaques implantables :</p> <ul style="list-style-type: none"> - arythmie ventriculaire entraînant une instabilité hémodynamique (mort subite récupérée, TV mal tolérée) et espérance de vie > 1 an avec un bon statut fonctionnel ; - patients avec insuffisance cardiaque symptomatique, NYHA II ou III, fraction d'éjection ventriculaire gauche (FEVG) \leq 35 %, malgré un traitement pharmacologique optimal \geq 3 mois et avec une espérance de vie > 1 an avec un bon statut fonctionnel, d'origine ischémique et > 40 jours après la phase aiguë d'un infarctus du myocarde ou d'origine non ischémique ; - maladie génétique à haut risque de mort subite par TV/fibrillation ventriculaire sans aucun autre traitement efficace connu. <p>Pour les systèmes implantables de resynchronisation cardiaque (stimulateurs ou défibrillateurs « triple chambre ») :</p> <ul style="list-style-type: none"> - patients insuffisants cardiaques chroniques symptomatiques, en classe NYHA II, III et IV ambulatoire, sous traitement médical optimal, avec fraction d'éjection \leq 35 %, en rythme sinusal : <ul style="list-style-type: none"> • avec une durée de QRS > 150 ms, • avec une durée de QRS comprise entre 120 et 150 ms et avec bloc de branche gauche ; - patients en fibrillation atriale permanente avec une insuffisance cardiaque chronique, ayant un QRS \geq 120 ms et une FEVG \leq 35 %, en classe NYHA III et IV ambulatoire malgré un traitement médical optimal, à condition de pouvoir obtenir une stimulation biventriculaire proche de 100 %.
	Défibrillateurs cardiaques externes portables	<p><i>Indications :</i></p> <p>Après explantation d'un système de défibrillation implantable pour infection de la loge ou des électrodes, jusqu'à la réimplantation (guérison de l'infection). En attente de transplantation cardiaque. L'indication doit être réévaluée tous les 3 mois (évaluation du rapport bénéfices/risques et de l'observance). Après revascularisation myocardique si la fraction d'éjection ventriculaire gauche (FEVG) est inférieure à 30 %, jusqu'à la réévaluation de la FEVG et discussion de l'indication d'un défibrillateur automatique implantable au 3^e mois. Au décours d'un infarctus du myocarde aigu si la FEVG est inférieure à 30 % après les 48 premières heures, jusqu'à la réévaluation de la FEVG et discussion de l'indication d'un défibrillateur automatique implantable au terme du 1^{er} mois.</p>
	Moniteur cardiaque implantable	<p><i>Indications :</i></p> <p>Diagnostic étiologique des syncopes. Diagnostic étiologique des accidents ischémiques cérébraux.</p>
Diabétologie	Système de mesure en continu du glucose interstitiel (MCG) couplé à une pompe à insuline	<p><i>Indications :</i></p> <ul style="list-style-type: none"> - patients diabétiques de type 1 (adultes et enfants) dont l'équilibre glycémique préalable est insuffisant (taux d'HbA1c \geq 8 %) en dépit d'une insulinothérapie intensive bien conduite par perfusion sous-cutanée continue d'insuline (pompe externe) depuis plus de 6 mois et d'une auto-surveillance glycémique pluriquotidienne (\geq 4/j) ; - patients diabétiques de type 1 (adultes et enfants) ayant présenté des hypoglycémies sévères ayant conduit à des interventions médicales en urgence, dans les 12 mois précédents, en dépit d'une insulinothérapie intensive bien conduite par perfusion sous-cutanée continue d'insuline (pompe externe) depuis plus de 6 mois et d'une autosurveillance glycémique pluriquotidienne (\geq 4/j). <p>Ces systèmes sont réservés aux patients ayant reçu une éducation thérapeutique, ainsi qu'une formation spécifique à l'utilisation du système de mesure en continu du glucose.</p> <p><i>Mode d'action :</i></p> <p>Le système MCG + pompe permet l'administration d'insuline (fonction pompe) et la mesure en continu de la concentration de glucose dans le tissu interstitiel (fonction moniteur du glucose). Le couplage des deux fonctions rend possible l'arrêt temporaire de la perfusion d'insuline en cas d'hypoglycémie avérée ou prévisible et en l'absence</p>

Domaine d'application	Catégories de DMC	Indications générales/Commentaires
		d'intervention du patient. La fonction d'arrêt avant hypo estime le risque de franchissement prévisible (dans les 30 minutes) d'un seuil prédéterminé de la glycémie estimée.
	Lecteurs de glycémie permettant le téléchargement en Bluetooth des données glycémiques	<i>Indications :</i> - chez le patient diabétique de type 1 : au moins 4 ASG/j ; - chez la femme ayant un diabète gestationnel : au moins 4 ASG/j ; - chez certains diabétiques de type 2 définis ci-dessous : • les patients insulino-traités : • les patients traités par insulinosécréteurs (sulfamides hypoglycémisants ou glinides, seuls ou associés à d'autres traitements antidiabétiques) afin de rechercher ou confirmer une hypoglycémie et d'adapter si besoin la posologie de ces médicaments. La prise en charge est toutefois limitée à 200 bandelettes par an pour ces patients, • les patients pour lesquels il est recherché une amélioration de l'équilibre glycémique lorsque l'objectif n'est pas atteint, comme instrument d'éducation permettant d'apprécier l'effet de l'activité physique, de l'alimentation et du traitement. La prise en charge est toutefois limitée à 200 bandelettes par an pour les patients non insulino-requérants. <i>Mode d'action :</i> Mesure de la glycémie à partir d'un prélèvement capillaire, à l'aide d'un algorithme intégré dans le lecteur.
	Système flash d'autosurveillance du glucose	<i>Indications :</i> Mesure du glucose interstitiel dans le traitement des patients atteints d'un diabète de type 1 ou de type 2 (adultes et enfants âgés d'au moins 4 ans) traités par insulinothérapie intensifiée (par pompe externe ou ≥ 3 injections par jour) et pratiquant une autosurveillance glycémique pluriquotidienne ($\geq 3/j$). Ce système est réservé aux patients ayant reçu une éducation thérapeutique ainsi qu'une formation spécifique à l'utilisation du système flash d'autosurveillance du glucose interstitiel. <i>Mode d'action :</i> Mesure du glucose interstitiel. Le lecteur permet d'indiquer des tendances sur les variations glycémiques à l'aide d'algorithmes.
	Logiciel et prestation associée d'aide au traitement par insuline en schéma basal-bolus couplé à une télésurveillance médicale	<i>Indications :</i> Patient adulte diabétique de type 1 (diagnostiqué depuis plus de 1 an) non contrôlé ($HbA1c \geq 8\%$) par une insulinothérapie en schéma basal-bolus administrée par multi-injections ou par pompe (depuis au moins 6 mois). La solution est réservée aux patients ayant reçu une formation spécifique à son utilisation. <i>Mode d'action :</i> La Solution est un logiciel couplé à une télésurveillance médicale et une prestation technique d'apprentissage de l'utilisation du logiciel. Le logiciel vise à aider le patient dans le calcul quotidien des doses d'insuline lente et rapide, selon des objectifs prédéfinis par le médecin prescripteur. Il est disponible via une application sur terminal mobile (smartphone ou tablette) pour le patient, et via un portail Web pour les professionnels de santé impliqués dans sa prise en charge (médecin prescripteur et infirmier). Le calculateur de dose d'insuline rapide propose une recommandation de dose d'insuline rapide fondée sur les données enregistrées par le patient (glycémies, événements glycémiques, consommation de glucides et activité physique). Ces données sont prises en compte sur la base de paramètres algorithmiques définis initialement par le professionnel de santé à travers la fonctionnalité de configuration du traitement à distance.
	Système de pompe implantable pour administration d'insuline par voie intrapéritonéale	<i>Indications :</i> Patients adultes diabétiques de type 1 non contrôlés par insuline administrée par voie sous-cutanée (y compris via une pompe) et présentant des épisodes hyperglycémiques et/ou hypoglycémiques sévères, fréquents ou non expliqués. <i>Mode d'action :</i> Pompe implantable reliée à un cathéter et un composant portable qui permet au médecin et au patient de communiquer avec la pompe par radiofréquence. Le composant portable permet de programmer les débits, activer les bolus, et garde en mémoire les informations concernant les différentes programmations.
Neurologie	Systèmes implantables rechargeables ou	<i>Indications :</i> 1- Douleur chronique d'origine neuropathique, après échec des alternatives thérapeutiques secondaires à :

Domaine d'application	Catégories de DMC	Indications générales/Commentaires
	non de stimulation médullaire	<ul style="list-style-type: none"> - un syndrome douloureux chronique radiculaire persistant depuis au moins 1 an en postopératoire ; - un syndrome douloureux chronique tronculaire (d'origine diabétique, zostérienne, traumatique ou chirurgicale) persistant depuis au moins 1 an ; - un syndrome régional douloureux complexe de type I ou II persistant depuis au moins 6 mois. <p>2- Douleur d'origine ischémique, en échec des alternatives thérapeutiques secondaires à la maladie de Buerger.</p> <p><i>Mode d'action de l'algorithme :</i></p> <ul style="list-style-type: none"> - indicateur de charge ou pile ; - programmes de stimulation ; - pour une gamme : détection et adaptation position debout/assis et adaptation de la délivrance de la thérapie.
	Systèmes implantables de stimulation du ganglion spinal	<p><i>Indications :</i></p> <p>Douleur chronique d'origine neuropathique, après échec des alternatives thérapeutiques secondaires à un syndrome régional douloureux complexe de type I ou II persistant depuis au moins 6 mois.</p> <p><i>Mode d'action de l'algorithme :</i></p> <ul style="list-style-type: none"> - indicateur de charge ou pile - programmes de stimulation
	Systèmes implantables de stimulation du nerf vague	<p><i>Indications :</i></p> <p>Enfant ou adulte atteint d'une épilepsie avérée (crise enregistrée en électroencéphalogramme - EEG) invalidante et pharmaco-résistante pour laquelle l'indication d'un traitement chirurgical intracrânien n'a pas été retenue. Les épilepsies pharmaco-résistantes sont définies par la persistance de crises après 2 ans sous traitement adapté, c'est-à-dire utilisation préalable en monothérapie séquentielle d'au moins 2 médicaments antiépileptiques et d'au moins une association de 2 médicaments antiépileptiques pendant une durée suffisante pour permettre d'en apprécier l'efficacité.</p> <p><i>Mode d'action de l'algorithme :</i></p> <ul style="list-style-type: none"> - indicateur de charge ou pile ; - programmes de stimulation ; - pour un DM : algorithme de détection automatique de crise fondé sur le monitoring du rythme cardiaque.
	Systèmes implantables de stimulation cérébrale profonde	<p><i>Indications :</i></p> <ul style="list-style-type: none"> - traitement des symptômes de la maladie de Parkinson idiopathique, lors de l'apparition d'une gêne fonctionnelle certaine (score de Schwab & England ≤ 70 % à certains moments de la journée) malgré un traitement médicamenteux optimisé ; - traitement des tremblements invalidants sévères, rebelles au traitement médical, en particulier dans le cadre du tremblement essentiel ; - traitement de la dystonie primaire chronique généralisée pharmaco-résistante, chez des patients âgés de 7 ans ou plus, n'ayant pas des capacités cognitives et ne bénéficiant pas d'un encadrement familial compatibles avec l'utilisation d'un système de neurostimulation implantable rechargeable, ou en primo-implantation. <p><i>Mode d'action de l'algorithme :</i></p> <ul style="list-style-type: none"> - indicateur de charge ou pile ; - programmes de stimulation.
Ophthalmologie	Implants rétinien	<p><i>Indications :</i></p> <p>Patients souffrant de dégénérescence rétinienne externe sévère à profonde bénéficiant d'une perception résiduelle de la lumière. S'il n'existe aucune perception résiduelle de la lumière, la rétine doit être capable de répondre à une stimulation électrique.</p> <p><i>Mode d'action de l'algorithme :</i></p> <ul style="list-style-type: none"> - indicateur de charge ; - traitement d'images et transformation en impulsions électriques.

Le champ des DMC entrant dans le périmètre d'évaluation de la CNEDiMTS est ainsi très hétérogène, la faculté d'émettre et de recevoir des données ayant des finalités différentes. L'évaluation de ces différentes catégories de DMC pour lesquels un dossier de demande d'inscription sur la

liste des produits et prestations remboursables a été soumis n'a pas mis en évidence de spécificité méthodologique dans le schéma des études cliniques mises en œuvre.

Le contenu des dossiers impliquant la transmission de données à des fins de télésurveillance médicale et ayant fait l'objet d'une évaluation par la CNEDiMTS a été spécifiquement analysé.

- **Principaux domaines thérapeutiques concernés** : cardiologie et diabète

Les DMC étaient des applications installées sur des smartphones ou des boîtiers communicants permettant une télésurveillance via la collecte, la transmission et l'analyse de données.

- **Types d'études**

Les schémas des études cliniques ne se distinguaient pas de ceux permettant l'évaluation d'autres technologies.

Les études cliniques analysées étaient principalement des essais contrôlés randomisés ou méta-analyse d'ECR.

Les autres types d'études cliniques recensés étaient les suivants :

- étude prospective comparative non randomisée ;
- étude observationnelle non comparative ;
- étude rétrospective non comparative.

- **Critères d'évaluation**

Le critère de jugement principal était clinique (critères cliniques validés) ou biologique (critère de substitution validé, en l'occurrence il s'agissait du taux d'HbA1c).

D'autres critères étaient évalués :

- la qualité de vie ;
- le temps passé par les médecins en consultation avec les patients (consultation conventionnelle ou téléphonique) et le temps passé par les patients en déplacement ou lors d'une attente à l'hôpital ;
- le nombre de consultations ;
- la satisfaction des patients ;
- le nombre d'hospitalisations ;
- l'observance.

Conclusion

Les dossiers concernant les dispositifs médicaux connectés soumis et ayant fait l'objet d'une évaluation clinique par la CNEDiMTS n'ont pas permis d'identifier de spécificités méthodologiques dans le schéma des études cliniques mises en œuvre.

3.3 Données à partir des bases de données d'essais cliniques

La recherche a été effectuée en deux temps en utilisant successivement deux mots clés différents : i) algorithm ; ii) mhealth.

► Base de données CLINICAL TRIALS - algorithm

La recherche a été effectuée en utilisant le mot clé *algorithm* et a répertorié 482 études.

Technologies

Les études identifiées impliquent des technologies d'une grande diversité. En effet, si certains algorithmes, tels les algorithmes apprenants, nécessitent l'usage de solutions informatiques, d'autres sont des arbres décisionnels affichés sur écran ou imprimés.

Types d'études

482 études ont été enregistrées dont 304 sont interventionnelles et 178 sont observationnelles. 258 sont décrites comme étant achevées (*completed*).

Parmi les études interventionnelles (n = 304), 192 sont comparatives. Le schéma de l'étude est randomisé entre deux bras ou plus pour 176 d'entre elles (58 %) ; les 16 autres études comparent les résultats (avant/après) à partir d'un seul groupe de patients. Pour 31 études interventionnelles, le design était un cross-over. Des schémas de randomisations séquentielles (n = 5) sont décrits.

Les principales pathologies ou domaines sur lesquels ont porté ces études sont :

- les maladies cardio-vasculaires (24 %) ;
- le diabète (16 %) ;
- la pneumologie (8 %) ;
- la cancérologie (7 %).

Pour le diabète, les études portent essentiellement (96 %) sur l'utilisation d'algorithmes à visée thérapeutique (par exemple, la régulation de l'administration d'insuline ou la sélection des patients potentiellement répondeurs). À l'inverse, en cancérologie, les algorithmes sont majoritairement (58 %) à visée diagnostique (par exemple, les algorithmes d'analyse d'image ou ceux de diagnostic reposant sur les résultats de tests non invasifs).

Critères

Pour les études comparatives, les critères d'évaluation portent principalement sur des critères cliniques ou biologiques, que les algorithmes soient apprenants ou non.

Pour les études non comparatives, elles sont majoritairement destinées à valider des algorithmes ; la mesure de la spécificité et celle de la sensibilité de la méthode sont des critères décrits.

► Base de données CLINICAL TRIALS - mhealth

La recherche effectuée en utilisant le mot clé *mhealth* (celui-ci était associé au terme *mobile health*) a répertorié 415 études.

Technologies

Les technologies concernées sont quasi exclusivement des applications installées sur des smartphones (*mobile apps*) qui sont associées ou non à des capteurs. Les éléments descriptifs des solutions mises en œuvre ne sont parfois pas présents dans la base de données ou sinon ils sont peu détaillés. Les dispositifs peuvent avoir un usage ponctuel : recueil d'un questionnaire ou d'une photo. Mais le plus souvent ils permettent un usage itératif actif ou passif de l'individu.

Types d'études

415 études ont été enregistrées dont 380 interventionnelles et 35 observationnelles. 151 sont décrites comme étant achevées (*completed*).

Parmi les études interventionnelles (n = 380), le schéma de l'étude est randomisé entre deux bras ou plus pour 300 d'entre elles. Pour 15 d'entre elles, le design était un cross-over. On note que de nombreux schémas proposent que le bras contrôle porte également le dispositif sans avoir accès aux résultats de la fonctionnalité évaluée (mesures masquées). Des schémas de randomisation séquentielle (n = 9) sont décrits.

Parmi les autres études interventionnelles (n = 80), 12 sont comparatives et 68 ne le sont pas.

Les études observationnelles (n = 35) sont essentiellement des études ne portant que sur un seul groupe de patients, celui utilisant le dispositif.

Les principales pathologies ou domaines sur lesquels ont porté ces études sont :

- les maladies cardio-vasculaires (13 %) ;

- les troubles psychiques (13 %) pour moitié liés aux addictions, les autres étant à part égale d'ordre psychiatrique ou liés à l'anxiété ou la dépression ;
- le diabète (11 %) ;
- les maladies sexuellement transmissibles (10 %) ;
- la nutrition (9 %) ;
- la cancérologie (9 %).

Critères

Parmi les critères spécifiques de la santé mobile, on retrouve ceux évaluant les applications elles-mêmes :

- la fréquence d'utilisation de l'application ;
- la mesure de l'acceptabilité de l'usage de l'application ou de sa qualité (par exemple avec l'échelle MARS, *Mobile App Rating Scale*, échelle spécifique aux applications mobiles) ;
- le temps passé sur l'application ;
- l'observance à l'utilisation de l'application ;
- la satisfaction d'utiliser l'application.

Les critères d'évaluation sont mesurés par le système directement, par les capteurs embarqués du smartphone (par exemple, l'accéléromètre mesure le nombre de levers ou le nombre de pas) ou des capteurs qui lui sont associés (par exemple des capteurs d'activité ou encore des capteurs de glucose ou de tension artérielle).

Les autres critères peuvent être cliniques ou biologiques (impact sur l'HbA1c), ou comportementaux (nombre de cigarettes consommées, nombre de prophylaxies post-exposition prises).

De nombreuses études favorisent l'autoévaluation du patient et cherchent à évaluer par des autoquestionnaires le degré d'anxiété, de stress, de dépression mais également la qualité de vie.

D'autres critères sont recherchés, et portent sur l'organisation du suivi du patient :

- le nombre de consultations ou d'hospitalisations réalisées (en urgence ou non) ;
- le délai avant consultation.

Effectifs des études

L'effectif indiqué est, selon les cas, le nombre de sujets nécessaires, le nombre de sujets inclus à la dernière mise à jour ou le nombre de sujets inclus en fin d'étude.

À noter, certaines études portent sur des effectifs conséquents, ce qui est rarement retrouvé pour les DM en général : qu'il s'agisse des études terminées ou en cours, 10 % d'entre elles incluent plus de 1 000 patients. De nombreuses études de faisabilité sur de faibles effectifs sont référencées.

Durée des études

La durée prévue des études recensées est extrêmement variable ; elle est d'une journée à plusieurs mois ou années.

Conclusion

Les schémas d'essais cliniques recensés dans la base de données d'essais cliniques *ClinicalTrials.gov* ne se distinguent pas de ceux habituellement décrits dans les dossiers d'évaluation pour les différentes catégories de dispositifs médicaux soumis à la CNEDiMETS. Aucune étude n'est commune aux deux analyses effectuées.

Pour les études comparatives impliquant des algorithmes, les critères de jugement évaluant l'impact de chaque méthode (bras avec algorithme ou bras comparateur) sont cliniques ou biologiques. Pour la validation d'algorithmes, la sensibilité et la spécificité sont principalement recherchées.

Pour les études relatives à la santé mobile, seuls quelques critères spécifiques aux applications sont décrits (par exemple le score MARS, l'utilisabilité, le temps ou la fréquence d'usage). Concernant les effectifs et la durée de suivi, en raison des objectifs différents de

ces études, il n'est pas possible de distinguer une spécificité propre à l'utilisation de dispositifs connectés.

Parmi les études interventionnelles, une majorité des études sont contrôlées et randomisées (58 % pour les études impliquant des algorithmes et 79 % pour les études liées à la santé mobile).

3.4 Données de la littérature

Aucune recommandation ni de rapport d'évaluation technologique ou méta-analyse relative aux spécificités méthodologiques d'évaluation clinique des dispositifs médicaux connectés n'ont été identifiés.

La recherche a permis d'identifier 1 revue de la littérature (2) et 7 guides (3-9).

Trois guides de la *Food and Drug Administration* (FDA) aux États-Unis ont été identifiés :

- le guide relatif aux applications médicales mobiles « Mobile medical applications » (5). Ce guide ne détaille pas de méthode d'évaluation spécifique à la réalisation d'essais cliniques. Il comprend notamment des définitions et des exigences réglementaires se rapportant à ces produits ;
- le projet de guide concernant les dispositifs ayant plusieurs fonctionnalités « Multiple function device products » (3). Ce projet de guide définit ces produits et précise les principes et le contenu des dossiers à soumettre à la FDA. Il ne détaille pas de méthode d'évaluation spécifique à la réalisation d'essais cliniques ;
- le guide relatif aux logiciels en tant que dispositifs médicaux « Software as a Medical Device (SAMd) » (4). Il reprend les recommandations de l'International Medical Device Regulators Forum (IMDRF). Ce guide détaille néanmoins une grille de criticité en fonction des fonctionnalités du logiciel et de leur impact clinique potentiel. La partie relative à l'évaluation clinique « clinical evaluation » ne décrit pas de méthodes d'évaluation spécifiques à la réalisation d'essais cliniques.

Le guide publié sur le site du ministère de la Santé de la Nouvelle-Zélande concerne les applications mobiles en santé « health apps » (7). Ce guide n'a pas décrit de spécificités méthodologiques pour la réalisation des essais cliniques. Néanmoins, sont cités des exemples d'outils permettant à tout utilisateur d'évaluer une application mobile, tels que les échelles « Mobile App Rating Scale » (MARS) et « App Chronic Disease Checklist » (ACDC) permettant d'attribuer un score selon la fonctionnalité, l'aspect esthétique, la qualité de l'information, la facilité d'utilisation, etc. Des sites Web sont également cités répertoriant des applications dans le domaine de santé, classées par thématiques. Ce guide précise que des travaux sont en cours sur le cadre de l'évaluation des applications cliniques « clinical apps » (le calendrier prévisionnel n'est pas précisé).

D'autre part, 6 publications abordant la problématique de l'évaluation clinique des technologies numériques et proposant des adaptations méthodologiques ont été analysées (10-15). Ces données complémentaires sont détaillées dans une partie spécifique.

Enfin, la recherche spécifique réalisée sur les essais « *in silico* » dans le domaine des dispositifs médicaux n'a pas permis d'identifier de recommandation, rapport d'évaluation technologique, méta-analyse ou revue systématique de la littérature. Les travaux et pistes de réflexion, notamment aux États-Unis et en Europe, ont été détaillés afin d'identifier les perspectives de développement de ces essais.

3.4.1 Guides méthodologiques

Trois guides méthodologiques ont été retenus :

- le guide méthodologique sur les technologies numériques « digital health technologies », élaboré par le National Institute for Health and Care Excellence (NICE) en 2018 (8) ;
- le guide méthodologique sur les applications numériques « digital applications », élaboré par le York Health Economic Consortium (YHEC) en 2016 (6). Ce guide repose sur un rapport répertoriant les différents types d'études concernées par l'évaluation de la santé digitale « digital health » (16). Des exemples d'études incluant des applications mobiles sont détaillés (17) ;
- le guide méthodologique sur les technologies numériques « digital health interventions », élaboré par l'Organisation mondiale de la santé (OMS) en 2016 (9).

3.4.1.1. Guide du National Institute for Health and Care Excellence (NICE), Royaume-Uni, 2018

Ce guide porte sur les technologies numériques « digital health technologies », telles que les applications mobiles, les dispositifs portables « wearable devices », les logiciels intégrant des algorithmes « software algorithms ».

Il a été réalisé entre juin et décembre 2018, sans analyse systématique de la littérature et sollicitation d'un groupe de travail multidisciplinaire.

L'objectif est d'identifier les éléments de preuve nécessaires en termes d'efficacité et d'impact économique, afin d'aider les développeurs à mettre en place leur programme de développement clinique.

Ce guide est fondé sur une approche par niveau de risque selon les fonctionnalités des outils numériques. Les niveaux de preuve nécessaires en termes d'efficacité découlent d'une classification fonctionnelle hiérarchisée (1, 2, 3A, 3B). Cette classification est détaillée dans le tableau 3.

Tableau 3. Classification fonctionnelle selon la finalité d'usage de la technologie

Classification fonctionnelle selon la finalité d'usage	Fonctionnalités	Description	Inclusions Exemples	Exclusions Exemples
Niveau 1 <i>Pas de bénéfice direct pour l'utilisateur</i>	Service support	Améliore l'efficacité sans mesure de paramètres de l'utilisateur.	Systèmes de prescription électronique, plateforme d'enregistrement de données de santé.	Systèmes qui proposent des conseils, un traitement ou un diagnostic.
Niveau 2 <i>Aide à la compréhension des règles hygiéno-diététiques et de la pathologie</i> <i>Pas de mesure de paramètres de l'utilisateur</i>	Information	Transmet des informations générales sur les conditions de vie, règles hygiéno-diététiques, pathologie.	Applications proposant des conseils hygiéno-diététiques (recettes).	Outils qui collectent des données ou qui traitent une maladie. Applications permettant la communication (entre utilisateurs ou un tiers).
	Simple surveillance	Permet aux utilisateurs d'enregistrer des paramètres afin de créer des carnets de suivi, sans transmission de données à un tiers.	Carnets de suivi électroniques (symptômes, état de santé, etc.).	Systèmes qui partagent des informations (entre utilisateurs ou un tiers), outils de traitement d'une pathologie.

Classification fonctionnelle selon la finalité d'usage	Fonctionnalités	Description	Inclusions Exemples	Exclusions Exemples
	Communication	Permet une communication entre utilisateur et professionnel, soignant, etc. Des conseils peuvent être prodigués par le professionnel mais pas par le système.	Applications de messagerie instantanée. Logiciel de consultation de type vidéoconférence. Plateformes de communication avec les soignants.	Systèmes qui fournissent par eux-mêmes un contenu clinique (tels que des programmes comportementaux cognitifs pour la dépression).
Niveau 3A <i>Aide à la prévention et surveillance d'une pathologie, par une mesure des paramètres de l'utilisateur</i>	Actions préventives sur le comportement	Visé à changer le mode de vie, le comportement (tabac, alimentation, alcool, etc.).	Systèmes permettant l'arrêt du tabac, la perte de poids.	Systèmes qui permettent un diagnostic ou un traitement, qui fournissent un conseil sur le mode de vie de l'utilisateur.
	Autogestion « self-manage »	Visé à aider les personnes atteintes d'une maladie diagnostiquée dans la gestion de leur traitement. Peut inclure une fonction connectée de suivi des symptômes par un professionnel de santé.	Systèmes qui permettent aux utilisateurs d'enregistrer et d'envoyer éventuellement des données à un professionnel de santé afin d'améliorer la gestion de leur maladie.	Systèmes qui permettent un diagnostic ou un traitement, applications qui surveillent et rapportent automatiquement les données à un professionnel de santé ou à une organisation tierce.
Niveau 3B <i>Outils utilisés pour le traitement et le diagnostic. Il est possible que ces systèmes soient qualifiés comme des dispositifs médicaux</i>	Traitement	Fournit le traitement d'une pathologie diagnostiquée (thérapies cognitives comportementales dans l'anxiété).	Systèmes pour traiter la santé mentale. Applications qui donnent des conseils sur les traitements.	Applications qui fournissent des conseils de santé généraux aux patients, systèmes destinés à améliorer les parcours de soins.
	Surveillance active	Enregistrement automatique et transmission de données à un professionnel, tiers ou autre, sans intervention du patient.	Dispositifs médicaux tels que les implants, capteurs dont les données sont automatiquement transmises.	Systèmes permettant à l'utilisateur de choisir si ses données seront transmises.
	Calcul	Outils permettant le calcul de performances cliniques impactant la décision médicale.	Systèmes utilisés par les professionnels ou utilisateurs pour calculer des paramètres (logiciels).	Systèmes qui permettent de diagnostiquer ou fournir un traitement.
	Diagnostic	Permet le diagnostic d'une pathologie ou aide à une décision de diagnostic d'une maladie.	Utilisation de données cliniques pour diagnostiquer une pathologie.	Systèmes permettant de fournir une liste de symptômes et signes d'une maladie.

Les technologies prises en compte sont extrêmement diverses, toutes ne relevant pas du statut de dispositif médical au vu de de la finalité d'usage. Pour le niveau de classification 3B correspondant aux technologies pouvant être marquées CE, le niveau de preuve attendu doit être considéré comme complémentaire aux exigences réglementaires formulées dans le cadre du marquage CE.

Pour chaque catégorie, une flexibilité sur le niveau de preuve est proposée (niveau de preuve minimal et niveau de preuve optimal « best practice standards »). Les niveaux de preuve sont cumulatifs pour répondre aux différentes questions soulevées par l'utilisation d'une technologie.

Par exemple, les services supports et les technologies permettant uniquement la transmission d'informations (sans mesure de paramètres de l'utilisateur) sont classés en niveau 1 et 2, respectivement. Ce qui signifie que ces technologies nécessitent la démonstration de la fiabilité, reproductibilité, sécurité, de l'exactitude du contenu, de la faisabilité de l'implémentation de la technologie dans le système de soins et de son acceptabilité par les utilisateurs, incluant la conformité aux standards et référentiels existants.

Concernant les dispositifs d'autosurveillance et les systèmes de télésurveillance médicale classés en niveau 3B, ces dispositifs nécessitent à la fois de répondre aux exigences requises pour les niveaux précédents (1 à 3A) et à la fois démontrer leur efficacité à l'aide d'études expérimentales ou quasi expérimentales de haute qualité, le niveau de preuve optimal étant l'ECR.

Des méthodologies d'études cliniques sont décrites pour les niveaux de classification fonctionnelle 3A et 3B. Les exigences requises pour les niveaux de classification fonctionnelle 3A et 3B sont détaillées dans le tableau 4.

Tableau 4. Exigences pour les niveaux de classification fonctionnelle 3A et 3B

Exigences	Niveau de preuve minimal	Niveau de preuve optimal
Niveau de classification fonctionnelle 3A		
Démonstration de l'efficacité	<p>Études observationnelles ou quasi expérimentales de haute qualité sur des critères pertinents. Ces études devraient être comparatives.</p> <p>Les comparaisons pourraient inclure :</p> <ul style="list-style-type: none"> • un groupe contrôle sur des critères pertinents ; • un groupe contrôle historique ; • une collecte systématique de données. <p>Les critères pertinents peuvent inclure :</p> <ul style="list-style-type: none"> • des critères comportementaux ou liés à la condition de vie, tels que la réduction du tabagisme ou l'amélioration de la gestion des conditions de vie ; • un changement de comportement positif ; • la satisfaction des usagers. 	<p>Études quasi expérimentales ou expérimentales de haute qualité qui intègrent un groupe de comparaison, montrant des améliorations sur des critères pertinents, tels que :</p> <ul style="list-style-type: none"> • les critères centrés sur les préférences et l'expérience du patient (patient reported outcome) en utilisant des outils validés, y compris la qualité de vie ; • autres mesures cliniques de sévérité ou d'invalidité de la maladie ; • mesures physiologiques ; • satisfaction et adhésion des usagers ; • utilisation des ressources de santé et de soins sociaux. <p>Le comparateur doit être la stratégie de soins de référence.</p>
Utilisation de techniques pertinentes pour le changement de comportement des utilisateurs	<p>Être en mesure de montrer que les techniques utilisées dans les technologies numériques sont :</p> <ul style="list-style-type: none"> • conformes aux recommandations (NICE ou organisations professionnelles pertinentes) ; • adaptées à la population cible. 	<p>Études qualitatives ou quantitatives publiées montrant que les techniques utilisées sont :</p> <ul style="list-style-type: none"> • basées sur des techniques de changement de comportement efficaces publiées et reconnues ; • alignées sur la pratique recommandée ; • adaptées à la population cible.
Niveau de classification fonctionnelle 3B		
Démonstration de l'efficacité	<p>études expérimentales ou quasi expérimentales de haute qualité montrant des améliorations sur des critères pertinents, tels que :</p> <ul style="list-style-type: none"> • fiabilité diagnostique ; • critères rapportés par le patient (PRO), y compris la qualité de vie ; • autres critères cliniques de sévérité ou handicap ; • mesures physiologiques ; • satisfaction et adhésion des usagers. <p>Le comparateur doit être la stratégie de référence.</p>	<p>Études contrôlées randomisées (ECR) de haute qualité réalisées dans le contexte d'utilisation de la technologie numérique, comparant la technologie à un comparateur pertinent et démontrant un bénéfice significatif « consistant » incluant un critère clinique et validé.</p> <p>En alternative : une méta-analyse bien menée d'ECR (s'il y a suffisamment d'études disponibles sur la technologie).</p>

Un commentaire se rapportant à la publication du guide du NICE a été publié (18). Il est notamment précisé que les spécificités de ces technologies concernent leur très forte rapidité d'évolution et la capacité de fournir des résultats en temps réel. Ces caractéristiques ont impliqué le développement de méthodologies telles que la *Multiphase Optimisation Strategy* (MOST) et les essais *Sequential Multiple Assignment Randomised Trials* (SMART).

Ce commentaire précise que les études à mettre en place doivent respecter les mêmes exigences cliniques que les autres technologies, incluant la transparence des méthodes, les plans d'analyse *a priori*, et la publication de tous les résultats de l'analyse. L'auteur précise qu'au-delà de l'efficacité, d'autres dimensions sont à explorer, notamment les aspects éthiques et de sécurité.

3.4.1.2. Guide et rapport associé du York Health Economic Consortium (YHEC), Royaume-Uni, 2016

3.4.1.2.1. Guide du York Health Economic Consortium (YHEC)

Le guide élaboré en 2016 par le York Health Economic Consortium (6) vise à aider les développeurs d'applications à comprendre et à identifier les éléments de preuve nécessaires pour démontrer les bénéfices apportés par leur applications.

Le York Health Economic Consortium est notamment chargé de fournir des avis axés sur la recherche en économie de la santé pour le National Health Service (NHS). Les décisions sur l'adoption d'une technologie innovante au NHS sont fondées sur une approche HTA « Health Technology Assessment ».

Le domaine d'application de ce guide porte sur les applications numériques « digital applications », incluant une application mobile « mobile app », une application Web « web-based application » ou, dans certains cas, un service numérique « digital service ».

Le guide est structuré sur six dimensions (détaillées ci-dessous) devant être étayées par des données. Le développeur doit comparer la solution technologique à la pratique actuelle et considérer l'ensemble du système et non l'application isolément de ces composants. Pour chaque critère, des questions spécifiques et des réponses à ces questions de recherche sont détaillées, afin de guider les développeurs de technologies de santé à recueillir les preuves pertinentes.

Ce guide précise que tous les critères sont importants. Néanmoins, pour certaines applications, un développeur peut être en mesure de démontrer que des critères spécifiques ne s'appliquent pas à son application. Par conséquent, tous les développeurs ne seront pas tenus de fournir l'exhaustivité des preuves sur tous les critères, mais les preuves requises seront proportionnelles à l'impact potentiel de l'application sur chaque critère et au risque d'un tel impact.

Une synthèse des **types d'études proposés dans le guide** pour chaque critère est détaillée ci-dessous.

► Sécurité/tolérance

La solution technologique doit être en mesure de démontrer qu'elle réduit ou n'a pas d'impact sur les patients et les soignants, en termes de sécurité et de tolérance, par rapport à la pratique actuelle.

Ce guide précise que les études cliniques sont peu susceptibles de détecter à elles seules tous les préjudices potentiels pour les patients et les utilisateurs. Aussi, les développeurs d'applications sont encouragés à identifier les problèmes potentiels le plus tôt possible, via un retour régulier des cliniciens et des patients utilisateurs.

Les méthodologies des études de sécurité/tolérance sont variées. Les **études de cohorte et séries de cas** ont généralement une durée plus longue qu'un essai contrôlé randomisé (ECR) et ont idéalement un bras comparateur. Ce sont souvent des études pragmatiques utilisées en routine. La fonction de la solution technologique peut déterminer le type d'étude requis. Néanmoins, pour les études de sécurité, **il n'y a pas de « gold standard »**.

Lorsque la technologie est diffusée et commercialisée, les développeurs d'applications sont invités à intégrer la collecte de données sur la tolérance/sécurité, notamment via des **registres** permettant une collecte systématique de données sur la sécurité.

► Efficacité

La solution technologique doit être en mesure de démontrer qu'elle améliore la santé des utilisateurs, ou n'a aucun impact sur celle-ci, par rapport à la pratique actuelle. S'il n'y a pas de bénéfices cliniques, des bénéfices sur les autres critères sont à démontrer.

Les conceptions d'études dans ce domaine doivent se focaliser sur l'impact de l'intervention sur les résultats cliniques et de bien-être. L'organisation mise en place pour l'utilisation de la solution technologique est un élément important impactant l'efficacité. Il est ainsi utile de savoir si l'application est utilisée conformément à la pratique recommandée. Le guide précise que le taux d'abandon pour ces technologies est plus élevé que pour plupart des autres interventions médicales. Ainsi, la façon dont l'application est utilisée dans la pratique et l'impact de celle-ci sur les résultats cliniques doivent être évalués.

Les études cliniques doivent identifier la combinaison optimale de chaque composant de l'intervention. Les méthodologies proposées portent sur la **stratégie d'optimisation multiphase (MOST) ou les essais randomisés à assignations multiples séquentielles (SMART)**. Le guide précise **qu'une fois la solution technologique optimisée, d'autres essais, généralement des ECR, sont toujours nécessaires pour identifier les bénéfices cliniques. Un ECR est le « gold standard » pour évaluer l'efficacité d'un produit de santé, mais d'autres études, y compris des études de cohorte et des séries de cas, peuvent fournir des preuves suffisantes, à condition que les biais potentiels puissent être maîtrisés.**

Lorsque la technologie est diffusée et commercialisée, les développeurs d'applications sont invités à réaliser des **études « en vie réelle »** qui recueillent des résultats dans l'environnement d'utilisation, notamment via des **registres ou des études de cohorte**. En effet, les ECR et les autres études quantitatives sont menés sur un échantillon. Les études « en vie réelle » permettent notamment d'identifier si les résultats de cet échantillon sont représentatifs de l'ensemble des patients utilisateurs.

Pour les applications conçues pour être utilisées par les patients, il est important que les développeurs mesurent des critères qui reflètent l'impact de l'application sur leur santé et leur bien-être. L'utilisation de **questionnaires validés spécifiques ou génériques** permet de démontrer l'impact de l'utilisation de l'application sur des résultats de santé (en utilisant, par exemple, le questionnaire générique EQ-5D ou des critères plus spécifiques d'une pathologie via les *Patient Reported Outcome Measures* (PROMs)). Par conséquent, les développeurs d'applications sont invités à fournir des preuves à l'aide de tels outils (susceptibles d'être incorporés dans des évaluations utilisant des ECR, des cohortes ou des séries de cas).

► Intérêt pour le patient/l'utilisateur

La solution technologique doit être en mesure de démontrer qu'elle est utile pour les patients/utilisateurs, améliorant leur état de santé et leur expérience en matière de santé à partir de résultats pertinents, par rapport à la pratique actuelle.

Le guide précise que les éléments de preuve peuvent notamment porter sur l'utilisabilité des applications « usability testing ». L'implication du patient est nécessaire en amont du développement de l'application. Cette implication peut prendre la forme d'observations, d'entretiens, de groupes de discussion, de questionnaires ou de tests d'utilisateurs structurés. Le recueil de l'expérience du patient lors de l'utilisation de l'application via les *Patient Reported Experience Measures* (PREMs) est un élément important à recueillir dans les études cliniques.

Les bénéfices sur le bien-être des patients sont plus difficiles à mesurer, bien qu'il existe des échelles validées pour mesurer les aspects du bien-être (échelles non précisées dans le guide). Les développeurs d'applications peuvent choisir de spécifier quels sont les avantages attendus, même s'il n'y a pas de données empiriques à l'appui. Les méthodologies d'étude peuvent concerner **divers types d'études, à la fois qualitatives et quantitatives.**

► Organisation des soins

La solution technologique doit être en mesure de démontrer qu'elle optimise ou n'a pas d'impact sur le parcours de soins du patient ou sur l'optimisation des soins apportés au patient.

Les éléments de preuve peuvent provenir **d'études économiques**, mais également **d'études qualitatives et quantitatives**.

► **Efficiences**

La solution technologique doit être en mesure de démontrer qu'elle peut réduire, ou non, le coût des ressources nécessaires pour dispenser des soins de qualité centrés sur le patient, comparativement à la pratique actuelle. Si le coût est plus important, il doit être justifié par des bénéfices sur d'autres critères.

Les preuves peuvent provenir **d'études économiques** associées aux **études qualitatives et quantitatives (ECR)**.

► **Équité**

La solution technologique doit être en mesure de démontrer qu'il existe peu d'obstacles à son utilisation selon le handicap, l'âge, la race, l'appartenance ethnique, le sexe, l'orientation sexuelle ou la religion.

Les éléments de preuve peuvent provenir **d'études interventionnelles** (par exemple, en détaillant les caractéristiques des patients inclus), des **études évaluant l'utilisabilité**, des **données collectées directement via l'application**, des retours des patients à partir de questionnaires, interviews et groupes de parole. Il n'y a **pas de type d'étude spécifique**.

3.4.1.2.2. Rapport du York Health Economic Consortium (YHEC)

Le guide méthodologique élaboré par le York Health Economic Consortium (YHEC) repose sur un rapport répertoriant les différents types d'études concernées par l'évaluation de la santé digitale « digital health » (16).

Ce rapport a été réalisé en 2016 par le York Health Economic Consortium (YHEC) pour le National Institute for Health and Care Excellence (NICE). La méthode d'élaboration du rapport n'est pas renseignée (analyse systématique de la littérature, critères de recherche, mise en place d'un groupe de travail, etc.).

Le domaine d'application porte sur la santé digitale « digital health » qui inclut la santé mobile « mHealth », les technologies de l'information en santé « health information technologies », les dispositifs intégrant des capteurs « wearable devices », la télésanté « telehealth », télémedecine, et médecine personnalisée.

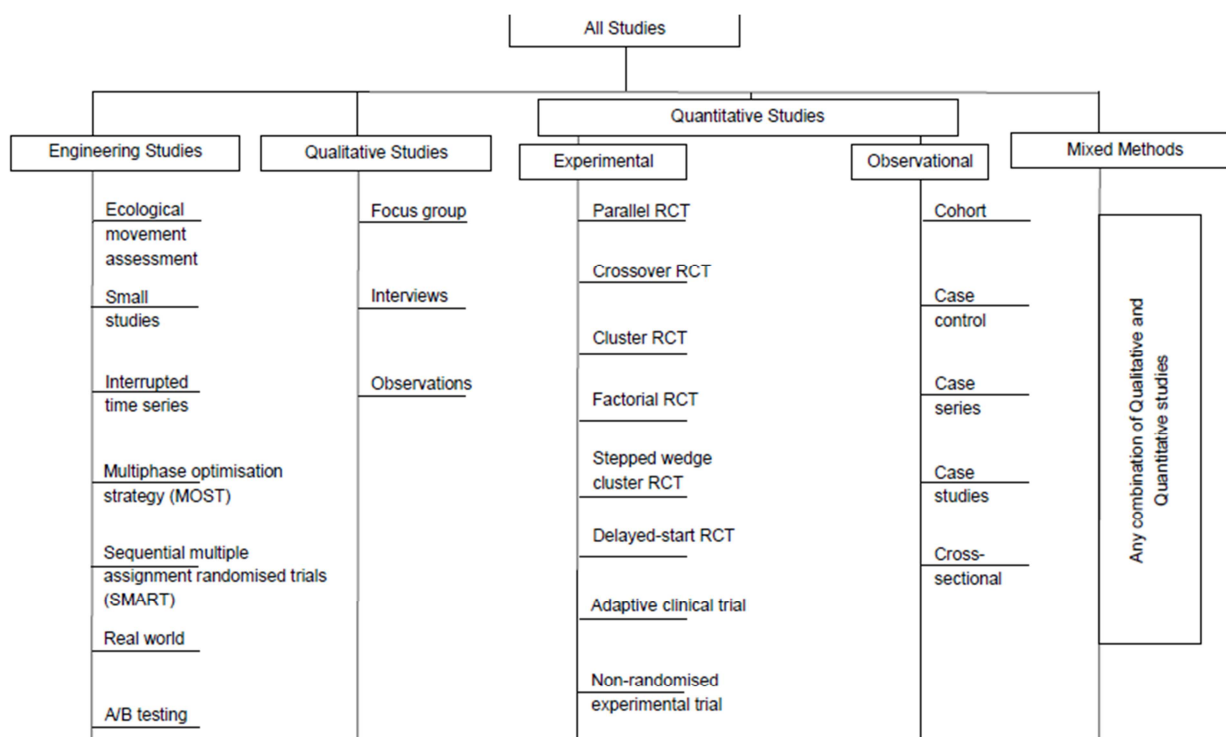
L'objectif de ce rapport était de fournir des informations sur les avantages et inconvénients de chaque type d'étude, à différents stades du cycle de développement du produit, ainsi que les limites d'interprétation des résultats.

Les auteurs précisent que les concepteurs des produits seront amenés à produire une base cumulative de niveaux de preuve à chaque stade du développement, avant une adoption à large échelle de la technologie. En raison de la nécessité probable de plusieurs plans d'étude sur la durée de vie du produit afin de recueillir les preuves nécessaires, les types d'études ne sont pas reliés à des questions de recherche spécifiques. Le rapport fournit une gamme de types d'études disponibles pour générer les preuves nécessaires pour répondre à chaque question de recherche. En outre, chaque produit étant différent, aucun plan d'étude unique n'est recommandé. Il est souligné l'importance de l'évaluation du système dans son ensemble et non le produit isolé de ses composants.

Les études répertoriées regroupent les études quantitatives (expérimentales et observationnelles), les études qualitatives, les études mixtes (à la fois quantitatives et qualitatives) ainsi que les études d'ingénierie « engineering studies ».

La Figure 1 indique les différents types d'études proposés.

Figure 1. Schémas d'études mentionnés dans le rapport du YHEC, 2016 (16)



Source : York Health Economic Consortium, Craig J, Shore J. RX137: Evidence generation guide for apps and wearables developers: study designs including applied examples. York: University of York; 2016.

Pour l'évaluation du bénéfice/risque d'une nouvelle technologie, **l'ECR bien mené est à privilégier**. Les études qualitatives permettent d'explorer les comportements des patients, leurs expériences, attitudes et interactions, afin d'améliorer la compréhension de la technologie vis-à-vis des utilisateurs dans un contexte spécifique. Les deux méthodologies peuvent être associées, il s'agit des méthodes mixtes. Elles se définissent comme l'intégration dans une même étude de données qualitatives et quantitatives et de leur analyse.

Les designs d'études focalisées sur l'ingénierie « engineering studies » sont utilisés pour résoudre des problèmes d'optimisation. Ce type d'étude a pour objectif d'identifier un ensemble optimal de paramètres, qui maximise la probabilité que plusieurs objectifs puissent être atteints, soumis à des contraintes, qui peuvent inclure des aspects comportementaux et le coût du produit. Ces modèles ont été largement utilisés depuis les années 1970 et sont de plus en plus adoptés dans le domaine de la santé, en particulier avec des interventions comportementales. Ces types d'études sont réalisés en amont, avant l'élaboration d'un ECR.

Les schémas d'études d'ingénierie détaillés dans ce rapport sont les suivants.

► **Études d'évaluation écologique momentané « ecological momentary assessment (EMA) studies »**

Ces études permettent aux participants d'autoévaluer les changements de symptômes, par exemple la douleur, les humeurs ou le comportement, en temps quasi réel dans leur vie quotidienne, à partir de dispositifs électroniques. Ces études peuvent utiliser les données longitudinales pour corrélérer les changements de symptômes aux changements d'interventions au fil du temps. Initialement, ces études sont utilisées dans le domaine de la psychopharmacologie clinique. Elles

peuvent ainsi examiner les relations entre les variables au fil du temps et potentiellement développer des algorithmes afin d'indiquer les étapes de changement de comportement.

Les auteurs soulignent que les retours des données qualitatives du patient (p. ex. via des questionnaires, groupes de discussion) sont des éléments importants à recueillir, surtout au cours des premières étapes du développement des technologies, afin d'optimiser la solution technologique avant la réalisation d'un ECR.

Avantages identifiés :

- Fiabilité des données recueillies.
- Des données qualitatives peuvent être intégrées au cours des premiers stades du développement de l'intervention, afin d'identifier les déficiences des systèmes et les corriger avant la réalisation des essais contrôlés.
- Permet de collecter de grandes quantités de données.
- Les données de l'étude peuvent être utilisées par les cliniciens et les patients pour suivre les progrès du traitement et informer les personnes qui ont besoin de changements de médication ou d'intervention.
- Type d'étude incluant la collecte et l'analyse des données.

Inconvénients identifiés :

- Certains patients ou certains cliniciens peuvent ne pas être ouverts à ce type d'étude.
- Il peut être difficile de trouver un groupe contrôle approprié.
- La méthodologie est encore en développement.

► **Études à faible effectif « small studies and N-1 studies »**

Ces études impliquent un très petit nombre de participants (moins de 10) et se concentrent sur les modifications des comportements individuels ou les taux d'événements au fil du temps. Par conséquent, le cadre de ces études porte essentiellement sur l'étude d'une seule personne ou d'un petit groupe de personnes au fil du temps avec une mesure répétée du résultat.

Avantages identifiés :

- Les patients et les cliniciens peuvent reconnaître rapidement les interventions inefficaces, réduisant ainsi les coûts et l'exposition potentielle aux effets indésirables.
- Elles peuvent aider à faire participer les patients à leurs propres soins et ainsi améliorer les résultats.
- Elles permettent de mener des recherches dans le cadre de la pratique clinique.

Inconvénients identifiés :

- Les bénéfices pour le patient et le clinicien doivent être clairement indiqués dès le départ.
- Les comités d'éthique peuvent ne pas accepter ce design comme une forme valide de recherche exploratoire.
- Les analyses statistiques doivent être développées.
- Les interventions doivent être soigneusement sélectionnées en évitant les effets à long terme.
- Faible validité externe.

► **Séries à temps interrompus « interrupted times series »**

Une série chronologique est une séquence de points de données réalisée sur un intervalle de temps continu, en utilisant des mesures successives à travers cet intervalle et à la même unité de temps entre chaque mesure consécutive. Les exemples peuvent concerner un relevé quotidien de la pression artérielle ou des scores de douleur. Ces exemples se rapportent à l'expérience d'un

individu, mais un tableau chronologique peut aussi mesurer des événements au niveau de la population comme les taux d'incidence des nouvelles maladies ou la mortalité ou l'adoption de nouvelles interventions de santé publique.

Une série chronologique d'observations répétées d'un événement particulier collecté au fil du temps est divisée en deux segments. Le premier segment comprend le taux de l'événement avant l'intervention et le deuxième segment est le taux après l'intervention. Une analyse statistique de régression segmentée («segmented regression») est utilisée pour mesurer les variations des résultats et de la tendance dans la période postérieure à l'intervention par rapport à la période précédant l'intervention. Ces études peuvent également inclure un bras de comparaison.

Avantages identifiés :

- L'utilisation de l'analyse des tendances dans le temps peut éviter les erreurs d'attribution d'un bénéfice lié à un effet continu de l'intervention. Le changement peut être à tort lié à l'intervention alors que, en fait, il était lié à d'autres facteurs.
- Des analyses stratifiées pour évaluer l'impact différentiel d'une intervention sur des sous-populations de patients (par exemple, selon l'âge, le sexe, la race) sont possibles.
- Ce type d'étude permet de fournir des informations claires sur les modifications et la tendance des changements dans la période postérieure à l'intervention par rapport à la période précédant l'intervention.

Inconvénients identifiés :

- Nécessite des connaissances statistiques pour l'analyse de régression.
- L'hypothèse de causalité n'est valable que si aucune autre variable ou processus ne change en même temps.
- Un minimum de 8 observations est nécessaire pour fournir une puissance suffisante pour l'analyse de régression.
- L'interprétation est renforcée s'il y a un contrôle direct, mais complexifie l'étude.

► **Stratégie d'optimisation multiphase « Multiphase Optimisation Strategy (MOST) »**

Cette stratégie a pour objectif d'optimiser et évaluer les interventions comportementales multicomposants. MOST inclut une étude randomisée pour l'évaluation de l'intervention, mais elle comprend d'autres phases de recherche avant l'ECR. Ces phases antérieures visent à optimiser l'intervention en utilisant des critères prédéfinis afin d'atteindre un objectif spécifique.

L'approche MOST vise à utiliser des plans expérimentaux efficaces afin d'optimiser les différents composants de l'intervention en utilisant un feedback lié aux comportements des participants. Un composant est une partie d'une intervention qui peut raisonnablement être séparée pour l'étude. Ces derniers sont menés avant qu'un ECR soit entrepris.

MOST comprend trois phases :

- une phase de préparation avec la création d'un modèle théorique permettant d'identifier les composants de l'intervention à évaluer et le critère d'optimisation ;
- une phase d'optimisation où l'investigateur teste empiriquement les options permettant de déterminer quels composants constituent l'intervention qui répond au critère d'optimisation. Cette phase utilise un essai factoriel randomisé « randomised factorial trial ». Le résultat de cette phase est l'identification des composants qui constituent l'intervention optimisée, ainsi qu'une compréhension de son efficacité probable. À ce stade, une décision est requise pour savoir s'il y a suffisamment de preuves pour mettre en œuvre un ECR. Si c'est le cas, la phase d'évaluation commence. Si ce n'est pas le cas, l'investigateur peut revenir à la phase 1 et pilote de nouveaux composants. Par conséquent, ce processus peut être itératif ;

- une phase d'évaluation qui consiste notamment à mener un ECR, afin de comparer l'intervention optimisée à un groupe contrôle. Si l'ECR indique que l'intervention n'est pas efficace, l'investigateur peut revenir à la première phase.

Avantages identifiés :

- Elle permet de tester efficacement l'optimisation de l'intervention. Les ECR conventionnels testent si une intervention fonctionne ou non ; leurs résultats peuvent conduire à une décision de révision de l'optimisation de l'intervention, nécessitant un autre essai randomisé.
- Elle permet de tester les effets des différents composants individuellement et de leurs associations, contrairement à un ECR conventionnel qui ne fait que mesurer l'efficacité de l'intervention dans son ensemble.
- Cette stratégie est flexible avec peu de limites, en théorie, sur le nombre de composants à tester, sur la capacité d'utiliser des paramètres de substitution à court terme pour évaluer les changements comportementaux et sur la capacité à permettre une gamme de méthodologie d'études, y compris les expérimentations factorielles dans l'évaluation.
- Les mesures traditionnelles de significativité statistique peuvent être appliquées aux résultats, de sorte que les analyses ne perdent aucune rigueur malgré la souplesse.

Inconvénients identifiés :

- La mise en œuvre des études peut être difficile car il faut s'assurer que chaque participant reçoit les éléments appropriés, aux moments appropriés, et que les résultats sont enregistrés et analysés. Ainsi, une planification minutieuse, une bonne collecte de données et une équipe bien formée et entraînée sont indispensables.
- Les décisions peuvent nécessiter des compromis entre les résultats.
- Les organismes de financement peuvent ne pas être familiers avec MOST en privilégiant l'approche ECR, en particulier concernant les incertitudes du départ relatives aux modalités de l'intervention.

► **Essais randomisés par assignation multiple séquentielle « Sequential Multiple Assignment Randomised Trials (SMART) »**

Ces essais sont utilisés pour développer des interventions adaptatives afin d'évaluer des troubles de santé complexes tels que les cancers, les addictions et la dépression. La gestion de ces maladies entraîne souvent le séquençage et la répétition de différentes approches possibles en fonction des réponses des participants. Une intervention adaptative est une séquence de règles de décision qui spécifient les modalités des mesures, le type ou l'attribution des traitements à des stades de décision prédéfinis dans le protocole de soins.

L'objectif de SMART est de développer uniquement une stratégie adaptative optimale, mais l'évaluation clinique est toujours nécessaire. SMART inclut plusieurs stades d'ECR. Chaque participant d'un SMART va modifier ses traitements et la randomisation est réalisée à chaque étape de traitement.

Les auteurs soulignent que SMART est une forme de conception expérimentale factorielle, et par conséquent les méthodes d'analyse de données standard peuvent être appliquées.

Les avantages et inconvénients des essais SMART sont semblables à ceux de MOST (les essais MOST englobent SMART). Les auteurs précisent que ces schémas d'études sont les seuls qui peuvent adapter de manière fiable les interventions aux individus en raison de leurs caractéristiques de randomisation unique. Ces études ne nécessitent pas de grandes tailles d'échantillon.

► Essais en vie réelle « real world studies »

Les résultats des essais cliniques peuvent ne pas être toujours une aide utile pour la prise de décision, en particulier si les résultats ne mesurent pas la valeur de l'intervention dans la pratique, en conditions de vie réelle. Il n'y a pas de design spécifique pour ce type d'étude. L'intérêt de ces essais est la collecte « en vie réelle » des données pour démontrer l'efficacité dans l'environnement d'utilisation qui peut différer de l'environnement d'utilisation contrôlé d'un ECR.

Les sources de données les plus utilisées sont :

- les bases de données : les données des utilisateurs sont entièrement rendues anonymes et agrégées afin de mener les analyses. Les cliniciens peuvent également utiliser les données que les patients consignent dans le dispositif connecté pour éclairer la prise de décision sur les options de traitement, en soutenant une approche de traitement personnalisée ;
- les enquêtes auprès des patients et des populations ;
- les données observationnelles issues de cohortes et d'études de cas ;
- les registres, permettant l'analyse de tous les patients traités dans un centre particulier, en utilisant des champs prédéfinis pour recueillir les résultats cliniques et les événements indésirables.

Avantages identifiés :

- Ces essais sont susceptibles d'inclure une population de patients plus représentative qu'une étude expérimentale.
- Des critères cliniquement pertinents peuvent être évalués lorsque des paramètres de substitution sont utilisés à court terme dans les essais cliniques.
- Ces essais peuvent fournir des renseignements à plus long terme sur la sécurité, les conditions d'utilisation et les coûts.

Inconvénients identifiés :

- Le coût d'un registre ou d'une base de données est élevé.
- Les incitations à renseigner les données peuvent être faibles, de sorte que les données manquantes ou de mauvaise qualité peuvent être importantes.
- Les définitions des événements ou du seuil pour inclure les événements peuvent différer d'un utilisateur à l'autre.
- Les informations essentielles nécessaires pour interpréter des résultats peuvent ne pas être disponibles.
- Des biais peuvent interférer les données.

► Tests A/B « A/B testing »

Les auteurs précisent que la plupart des applications, en particulier celles qui sont utilisées sur le Web, auront fait l'objet de tests officiels ou informels A/B. Il s'agit d'une expérimentation randomisée avec deux variantes, A et B, le contrôle et la variation. Les réponses des utilisateurs à chaque version sont comparées pour informer sur leur efficacité relative. Avec les applications Web, les variables sont souvent des pages Web qui sont essentiellement les mêmes, sauf pour les variations d'éléments individuels comme les mises en page, les images ou les couleurs. Dans ces paramètres en ligne, l'objectif des tests A/B est d'identifier les modifications apportées aux pages Web qui augmentent ou maximisent les variables d'intérêt (par exemple, taux de clics pour une bannière publicitaire). La page Web actuelle est associée à l'hypothèse nulle.

Ces tests sont randomisés au même titre que les ECR, mais leurs objectifs sont très différents.

Avantages identifiés :

- Ce test peut être réalisé en continu.
- Les résultats sont immédiatement disponibles, ce qui permet d'actualiser régulièrement le site Web.
- La planification et la conduite de ces tests sont simples et peu coûteuses.
- Les utilisateurs de sites Web constituent des groupes de patients randomisés permettant directement leur recrutement à moindre coût.

Inconvénients identifiés :

- Un seul aspect peut être modifié à la fois et les tests de séquençage dans le temps peuvent être complexes.
- La collecte de données inefficaces, car aucune information sur le test précédent ne peut être réutilisée pour l'évaluation des autres variables à venir.

D'autres schémas d'étude sont détaillés, notamment des schémas d'étude innovantes « innovative study designs ».

► **Estimation de variables instrumentales**

L'estimation de variables instrumentales « instrumental variables estimation » est une technique utilisée pour contrôler les variables de confusion et les erreurs de mesure. Les études qui ne sont pas randomisées, telles que les études observationnelles, peuvent avoir des biais de confusion, ce qui limite l'interprétation des résultats de l'étude. La plupart des variables de confusion sont observables « observable » et peuvent être contrôlées par des modèles de régression, afin de minimiser leur impact sur les résultats. Cependant, certaines variables sont non observables « unobservable ». L'estimation de variables instrumentales identifie les variables ou instruments qui ont deux propriétés: 1) ils affectent la variable de traitement, c'est-à-dire l'intervention, mais 2) n'ont pas d'effet direct sur la variable de résultat. Une fois que la variable instrumentale a été identifiée, elle peut être intégrée dans un modèle de régression (avec les autres variables de confusion) pour déterminer l'effet causé par l'intervention.

► **Utilisation du big data en santé**

Le big data est défini dans ce rapport comme un immense volume de données qui dépassent la capacité de traitement des systèmes de base de données conventionnels et qui peut être défini par six caractéristiques en « V »: 1) volume, 2) vitesse, 3) variété, 4) valeur, 5) variabilité et 6) véracité.

En santé, le big data fait référence à de grands ensembles de données électroniques complexes, difficiles à manipuler avec les méthodes et les logiciels existants. Il peut inclure des données génomiques, cliniques et comportementales, des données de publication, des données administratives ou commerciales.

Des techniques d'analyse sophistiquées peuvent être appliquées à ces données intégrées pour fournir des profils de patients individuels ainsi que des marqueurs de coût et d'efficacité clinique, permettant d'identifier et de personnaliser les interventions les plus efficaces sur le plan clinique et les plus rentables (les techniques ne sont pas détaillées dans le rapport).

Les avantages identifiés sont les suivants :

- amélioration des résultats cliniques grâce à des diagnostics plus précis et personnalisés ;
- réduction des coûts grâce à l'identification précoce de la maladie ;
- amélioration de la santé publique, telle que la prévention de l'obésité et du diabète ;
- réduction des admissions hospitalières imprévues et urgentes grâce à l'optimisation des soins.

Ce type de données peut être collecté via le téléphone mobile du patient, permettant par exemple une surveillance de scores subjectifs (humeur, douleur), ainsi que des marqueurs cliniques (poids, pression artérielle, taux de glucose).

Les auteurs soulignent le potentiel de l'intégration du big data dans le système de santé, permettant de collecter des données via des applications ou d'autres appareils numériques, malgré les nombreux obstacles non encore résolus, tels que le budget et l'infrastructure, les problèmes liés à la confidentialité des données personnelles du patient, notamment.

Un exemple de stratégie d'évaluation d'une application mobile a été détaillé dans le Tableau 5.

Tableau 5. Exemple de stratégie d'évaluation d'une application mobile

	Research methods	Purpose	Example measures
Formative research	<ul style="list-style-type: none"> Focus groups; Online surveys. 	To inform the development of the intervention content and process.	<ul style="list-style-type: none"> How do the target audience use their phones? How does evidence and theory behind intervention fit with mobile phone usage? What would attract target audience to population?
Pretesting	<ul style="list-style-type: none"> Online surveys; Focus groups; Individual interviews. 	<ul style="list-style-type: none"> To determine acceptability of proposed intervention to target audience; To improve and refine intervention on basis of feedback. 	<ul style="list-style-type: none"> What styles/content/language do they prefer? Are messages useful and understandable?
Pilot study	Small and nonrandomised.	<ul style="list-style-type: none"> To test content and regimen of intervention; To test processes. 	<ul style="list-style-type: none"> Is the intervention acceptable? Are there any technical or process issues? Baseline measures for sample size calculations if necessary.
Randomised control trial	Pragmatic community-based randomised control trial.	To test the effect of the intervention in comparison with a control group.	<ul style="list-style-type: none"> Health-related outcomes; Objective measures; Participant satisfaction; Adverse / unintended effects.
Qualitative research	Semi-structured interviews.	<ul style="list-style-type: none"> To improve the intervention further; To determine implementation issues and methods. 	<ul style="list-style-type: none"> How can we improve the intervention? What aspects were un/helpful? Was theory conveyed? Were messages acted upon? What are the best methods for rollout and promotion?
Evaluation of implementation impact	<ul style="list-style-type: none"> Phone / online surveys; Semi-structured interviews. 	To determine the effect of the intervention once scaled up.	Reach, utilization, health-related outcomes, flow on effects on health / other services, unintended consequences, effect of different promotional methods.

Source : York Health Economic Consortium, Craig J, Shore J. RX137: Evidence generation guide for apps and wearables developers: study designs including applied examples. York: University of York; 2016.

3.4.1.3. Guide de l'Organisation mondiale de la santé (OMS), 2016

Le guide de l'OMS intitulé « Monitoring and evaluating digital health interventions » (9) vise à aider le lecteur tout au long du parcours du développement de la technologie, des exigences techniques aux tests de faisabilité jusqu'à l'évaluation clinique et la mesure de leur impact. Le domaine d'application concerne les applications santé mobile ou e-santé (mHealth ou eHealth). La méthode d'élaboration du rapport n'est pas renseignée (analyse systématique de la littérature, critères de recherche, mise en place d'un groupe de travail, etc.).

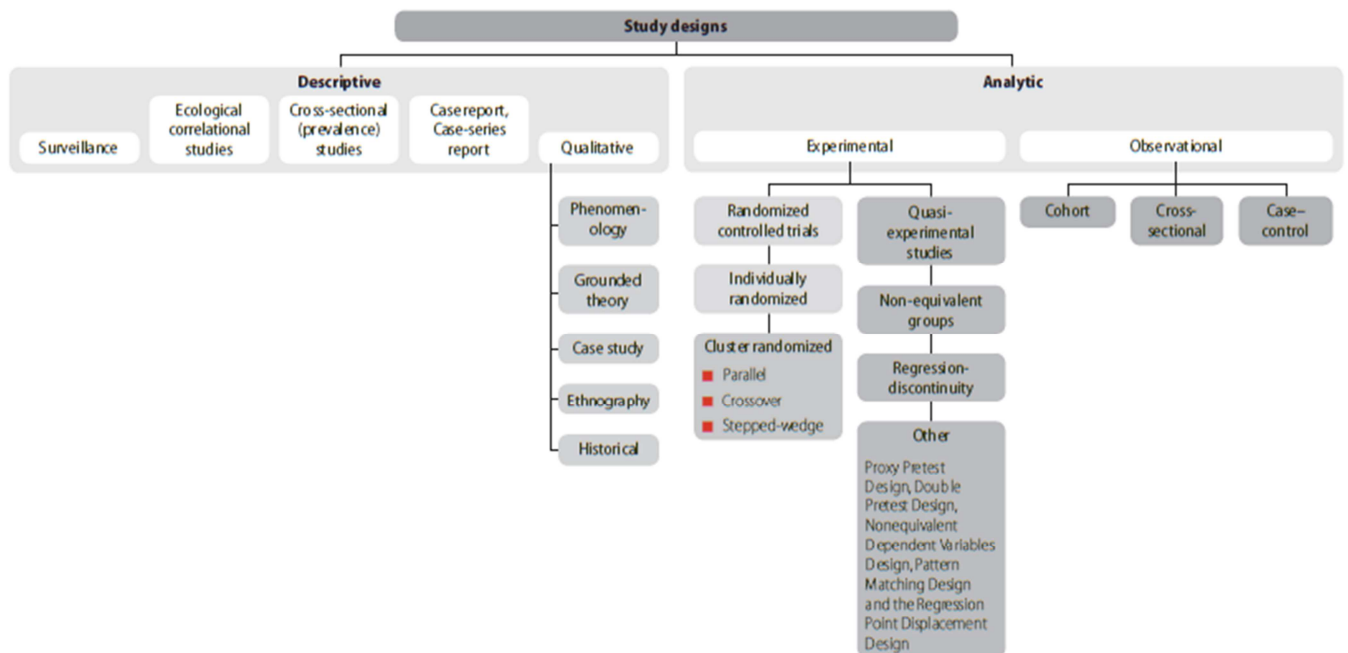
La rapidité d'évolution de ces technologies est soulignée.

Concernant l'évaluation clinique, les différentes étapes sont rappelées selon le degré de maturité de la technologie :

- faisabilité : détermine si la technologie numérique fonctionne comme prévu dans un contexte donné ;
- utilisabilité « usability »: détermine si la technologie est utilisée comme prévu ;
- efficacité: évalue si la technologie donne les résultats escomptés dans le cadre d'une étude contrôlée ou non ;
- mise en œuvre « implementation research »: évalue la technologie dans les conditions réelles d'utilisation incluant les pratiques médicales.

Les types d'études cliniques répertoriées regroupent les études quantitatives (expérimentales et observationnelles), les études qualitatives, les études mixtes (à la fois quantitatives et qualitatives). Elles sont regroupées dans la figure 2.

Figure 2. Schémas d'études mentionnés dans le guide de l'OMS, 2016 (9)



Source : Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. Geneva: World Health Organization; 2016. Licence: CC BY-NC-SA 3.0 IGO.

Les auteurs soulignent que le niveau de preuve optimal est l'ECR multicentrique. Pour évaluer la technologie numérique dans les conditions réelles d'utilisation (phase de mise en œuvre), une étude observationnelle ou quasi expérimentale incluant des critères qualitatifs et quantitatifs serait plus appropriée.

Conclusion

Trois guides ont été retenus et produits récemment, le guide de l'OMS en 2016 et 2 guides émanant du Royaume-Uni, le guide du York Health Economic Consortium en 2016 et celui du NICE en 2018. À noter : le guide du NICE ne reprend pas les éléments issus du NHS.

Ils visent à aider les développeurs à identifier les éléments de preuve nécessaires pour démontrer les bénéfices apportés par leurs technologies.

Le guide élaboré par le York Health Economic Consortium (YHEC) concerne les applications mobiles en e-santé et porte sur six dimensions (sécurité/tolérance, efficacité, critère centré sur le patient, impact organisationnel, efficience, équité). Pour chaque dimension, des méthodologies d'études sont suggérées.

Le guide du NICE concerne les technologies numériques et a défini une classification hiérarchisée de niveaux de preuve selon les fonctionnalités de la technologie. Pour les technologies répondant à la définition d'un dispositif médical, la démonstration de leur efficacité est nécessaire avec comme niveau de preuve optimal l'ECR.

Le guide de l'OMS concerne les technologies numériques en santé (mHealth ou eHealth) et a établi des niveaux de preuve selon le degré de maturité de la technologie.

Le champ des technologies couvert par ces guides est large et n'est pas superposable au périmètre des dispositifs concernés par le projet de la CNEDiMTS. Néanmoins, les méthodologies d'études se rapportant à un dispositif médical connecté ne se distinguent pas de celles des autres dispositifs médicaux, le « gold standard » pour évaluer l'efficacité des technologies numériques étant l'ECR.

Les types d'études cliniques répertoriées dans ces guides regroupent les études quantitatives (expérimentales et observationnelles), les études qualitatives et les études mixtes (à la fois quantitatives et qualitatives). Le rapport du YHEC a répertorié d'autres études, notamment des études d'ingénierie et des schémas d'étude dits « innovants », tels que l'estimation des variables instrumentales et l'utilisation du big data.

3.4.2 Revue de la littérature

La recherche a permis d'identifier 1 revue de la littérature analysant les méthodologies des essais cliniques impliquant des applications mobiles en santé (2).

L'étude de Pham *et al.* (2) a analysé, sur la base de données d'essais cliniques *ClinicalTrials.gov*, les essais cliniques impliquant des applications mobiles en santé « mHealth apps » entre novembre 2014 et novembre 2015.

La recherche a été réalisée en utilisant le terme « mobile application ».

Les critères d'inclusion étaient les suivants :

- évaluation d'une application mobile en santé ;
- mesure d'un critère clinique ;
- application exclusivement installée sur un smartphone et non sur un site Web « web-based app ».

Les critères d'exclusion étaient les suivants :

- interventions fondées sur l'envoi de SMS uniquement « SMS-based interventions » ou sur des appels téléphoniques ;
- études incluant dans un même groupe les interventions avec et sans applications mobiles ;
- applications permettant uniquement la prise de rendez-vous ;
- applications qui ne nécessitent pas l'intervention de l'utilisateur pour les données d'entrée.

La stratégie de recherche a identifié 137 essais cliniques parmi lesquels 71 répondaient aux critères d'inclusion.

La plupart des études étaient interventionnelles (68/71, 96 %), avec seulement 3 études observationnelles (3/71, 4 %). La majorité des essais étaient des ECR (57/71, 80 %).

Des critères qualitatifs (critères non détaillés) étaient évalués dans 17 essais (17/71, 24 %).

Le groupe contrôle concernait la stratégie de soins habituelle (30/59, 51 %), un traitement actif (26/59, 44 %) ou une liste d'attente « waitlist » (3/59, 5 %).

Les domaines concernés portaient sur :

- La santé mentale « mental health » (12/71, 17 %) ;
- Les pathologies cardio-vasculaires (8/71, 11 %) ;
- Le diabète (8/71, 11 %) ;
- Le cancer (7/71, 10 %).

Parmi ces essais, seuls 7 étaient terminés au moment de l'analyse rétrospective (7/71, 10 %).

La durée de l'étude, définie par le temps entre le recrutement des patients et la collecte complète des données, était en moyenne de 20 mois. Le nombre médian de sujets inclus était de 112 (20 essais avaient un échantillon de patients entre 0 et 49 (20/71, 28 %), 33 entre 101 et 499 (33/71, 47 %) et 8 > 500 participants (8/71, 11 %)). Le plus large essai portait sur 12 000 participants.

Les auteurs concluent que les méthodes d'évaluation des technologies mHealth n'étaient pas différentes des méthodes habituelles et soulignent la nécessité d'une évaluation clinique de qualité afin d'évaluer l'impact significatif des technologies de santé.

Conclusion

Une revue de la littérature analysant les méthodologies des essais cliniques impliquant des applications mobiles en santé a été identifiée.

Les schémas d'études cliniques recensés dans la base de données d'essais cliniques *ClinicalTrials.gov* ne se distinguent pas de ceux habituellement décrits dans les dossiers d'évaluation pour les différentes catégories de dispositifs médicaux soumis à la CNEDiMTS.

La plupart des études étaient interventionnelles et impliquaient un ECR (80 %). Les auteurs soulignent l'importance d'une évaluation clinique de qualité pour démontrer l'impact de ces technologies.

3.4.3 Données complémentaires

La stratégie de recherche documentaire a permis d'identifier 6 publications qui abordaient la problématique de l'évaluation clinique des technologies numériques et qui proposaient des adaptations méthodologiques (10-15).

Les particularités liées aux technologies numériques identifiées par les auteurs ont concerné :

- la rapidité d'évolution technologique (10-15) et les mises à jour régulières des technologies (11) ;
- les interactions entre les participants. Ces technologies font intervenir un ensemble complexe de composants qui ont une interrelation dynamique avec les participants à n'importe quelle étape du process (transmission, réception, acceptation, intention d'utilisation). Ces particularités impliquent que l'évaluation ne doit pas uniquement se focaliser sur la transmission, elle doit être multifactorielle. L'évaluation de facteurs humains, tels que l'acceptation et l'appropriation de la technologie par le patient, doit être intégrée

- dans le suivi afin de s'assurer que la technologie soit utilisée en conformité avec le plan de l'étude (10) ;
- la collecte automatique des données et la transmission des données via Internet et via d'autres technologies. Ces particularités impliquent l'établissement de procédures afin de surveiller les indicateurs de l'état de santé du patient, via des résultats permettant de garantir la sécurité du patient et de détecter des éventuels changements de l'état de santé du patient (10). Les données collectées peuvent être biologiques, comportementales, ou environnementales. Elles incluent des capteurs qui peuvent surveiller des paramètres avec une meilleure précision, améliorer la fréquence des échantillons, diminuer les données manquantes, ou augmenter la commodité d'emploi (11). Cette collecte de données est réalisée en temps réel sans intervention du patient, ce qui peut permettre de réaliser des analyses plus rapides, induisant une diminution du temps de réalisation des études (11). La collecte peut concerner plusieurs paramètres, ce qui permet de réaliser des analyses exploratoires sur des éventuels liens entre des paramètres, selon le contexte et l'environnement d'utilisation (11) ;
 - la masse importante de données collectées big data permettant le recueil automatique et continu de paramètres qui étaient auparavant mesurés de façon intermittente et limitée dans le temps. L'analyse de ces données peut augmenter le pouvoir discriminant des études expérimentales permettant d'évaluer les variabilités inter- et intra- individuelles (11) ;
 - l'attractivité de ces technologies par certains patients induisant une facilité de recrutement des patients dans les études. Le nombre d'interactions face-face et le nombre de déplacements peuvent être diminués voire évités, ce qui permet au patient de participer à l'étude à distance, chez lui, au lieu de se rendre dans un centre de recherche académique (10, 11) ;
 - l'adaptation et la personnalisation selon le profil des patients et en temps réel, permettant d'augmenter l'adhésion du patient et renforcer la puissance de l'intervention (11).

Afin de pallier ces particularités, les adaptations méthodologiques identifiées sont listées ci-dessous.

La réalisation d'un ECR est généralement présentée comme le « gold standard » pour démontrer l'efficacité des interventions en santé. Néanmoins, les freins reportés dans la littérature relatifs à la réalisation d'un ECR d'une « mobile health technology » portent sur :

- le décalage dans le temps lié à la réalisation de l'étude : 5 ans en moyenne, entre le recrutement du patient et la publication de l'étude. Ce qui rend obsolète la technologie à la fin de l'étude ;
- le coût ;
- la randomisation des traitements ;
- l'adhésion au traitement.

Dans certains cas, la rapidité d'évolution de la technologie et son adoption par les patients et l'équipe soignante peuvent impliquer des modifications en continu de certains composants du groupe intervention pendant l'essai. Cette évolution rapide peut conduire les concepteurs de l'étude à passer rapidement du projet pilote à la diffusion de la technologie ou éviter les ECR à large échelle, ce qui limite la validité clinique des résultats (11).

Kumar *et al.* (11) précisent que, compte tenu des travaux récents en mHealth et la révolution des données qu'elle augure, les capacités de ces technologies seraient susceptibles de faire changer les méthodologies cliniques existantes permettant le développement de nouveaux designs plus efficaces. Néanmoins, de nombreuses questions restent sans réponse concernant l'utilisation de méthodes d'évaluation adaptées pour ces technologies. Par exemple, la possibilité de réaliser des mesures multiples répétées sur un faible échantillon de patients au lieu de réaliser quelques mesures sur de nombreux participants. Cette éventualité permettrait d'optimiser la réalisation de l'ECR (c.-à-d., plus rapide avec un échantillon restreint).

Dans cette perspective, des conceptions de recherche potentielles pour évaluer l'efficacité des interventions mHealth ont été décrites.

Les auteurs soulignent que, lorsque la technologie est dite « mature », c'est-à-dire lorsque des données pilotes ou quasi expérimentales soutenant l'efficacité des technologies ont déjà été validées, **le type d'étude à privilégier reste l'ECR** (10-15).

Les adaptations méthodologiques ou designs alternatifs à l'ECR proposés sont :

- **le plan expérimental de discontinuité de la régression « regression discontinuity design »** (11) ;

Les participants sont assignés dans le groupe intervention ou dans groupe contrôle selon une valeur seuil. La variable d'affectation peut être toute variable mesurée avant traitement. Le design est plus robuste si la valeur seuil est la moyenne de la variable d'affectation. Ce plan expérimental permet d'inclure uniquement une fraction des participants pour l'analyse (11).

D'autres méthodes statistiques ayant pour objectif de réduire les biais de confusion liés aux études observationnelles sont citées, telles que le **score de propension, les stratifications et le modèle de régression multivariable « multivariable regression modeling »** (15).

- **les essais randomisés en cluster « Stepped-wedge designs »** (11)

Le groupe intervention peut être comparé à la fois avec les mesures initiales réalisées dans ce groupe, mais également avec des mesures d'autres patients qui n'ont pas reçu le traitement, qui forment un groupe témoin indépendant et homogène. Dans cet essai, tous les participants sont informés qu'ils recevront le traitement, ce qui garantit que les participants ne se voient pas refuser l'intervention. Cet essai est appropriée si l'intervention peut porter sur tous les individus (ou sur tous les sites) (11, 15).

Les concepteurs souhaiteraient apporter des mises à jour régulières de leurs technologies durant leur évaluation. Cependant, toute modification du groupe intervention durant l'essai risque de limiter la validité interne des résultats (11).

Afin de pallier l'évolution continue des technologies, **la méthode CEEBIT « Continuous Evaluation of Evolving Behavioral Intervention Technologies »** (19) est citée pour tester l'évolutivité des interventions (11). Cette méthode permettrait de retenir la version la plus efficace, basée sur des critères élaborés *a priori*. Elle serait ainsi adaptée pour les évaluations de technologies qui évoluent en continu, qui s'améliorent au fil du temps.

Les auteurs soulignent que l'ECR peut néanmoins toujours être réalisé si le niveau d'inférence « *level of inference* » dans le groupe intervention concerne un ensemble de fonctionnalités robustes qui s'adapteront naturellement à l'environnement et aux préférences des patients au fil du temps et selon le contexte d'utilisation (11).

Enfin, certains auteurs proposent d'utiliser des designs basés sur des modèles « model-based designs » pour adapter le groupe intervention utilisant la technologie mHealth.

L'adaptation et la personnalisation dans le groupe intervention impliquent une bonne connaissance de la variabilité intra- et interindividuelle. Des méthodes statistiques qui précisent ces variabilités sont nécessaires (11).

Afin d'ajuster et optimiser le groupe intervention, des stratégies ont été développées. L'exemple de la **stratégie d'optimisation multiphase (MOST « Multiphase Optimization Strategy »)** est cité, permettant d'identifier les composants prometteurs d'une intervention au niveau de la phase de screening. Ces composants sont par la suite évalués dans une étude randomisée.

Pour affiner l'intervention, des **essais randomisés par assignation multiple séquentielle (SMART « Sequential Multiple Assignment Randomized Trials »)** peuvent être utilisés. Les individus sont assignés au hasard à divers choix d'intervention au fil du temps. Dans les essais SMART, les chercheurs décident quels aspects des traitements nécessitent une investigation et ensuite randomisent les individus à chaque décision de traitement fondée sur la faisabilité, les aspects éthiques, ou autres facteurs.

Les stratégies MOST et SMART sont également citées par d'autres auteurs (12-14) comme approches permettant l'utilisation efficiente des ressources et l'identification de l'intervention la plus appropriée, avant la mise en place d'un ECR impliquant des applications mobiles.

Conclusion

Des particularités liées aux DMC sont identifiées par plusieurs auteurs. Ces particularités concernent la rapidité d'évolution de la technologie, les interrelations dynamiques avec les participants à n'importe quelle étape du process, la collecte automatique, le traitement et la transmission des données. L'enjeu étant de concilier l'évaluation clinique avec ces particularités, et notamment avec le rythme des évolutions. Un certain nombre de propositions convergentes ressortent, notamment la garantie de la fiabilité, la validité et la qualité des résultats issus des technologies numériques et la prise en compte d'une évaluation multifactorielle intégrant l'acceptation et l'appropriation de la technologie par le patient.

Les auteurs soulignent que malgré les difficultés rencontrées pour sa mise en place, l'ECR est le type d'étude à privilégier pour évaluer le bénéfice d'une technologie numérique. Des adaptations méthodologiques sont néanmoins suggérées.

3.4.4 Données sur les études *in silico*

La recherche *in silico* consiste à développer des modélisations mathématiques d'une maladie fondées sur la physiologie et la biologie (par ex. le système immunitaire, l'arbre bronchique), afin de constituer une population de patients virtuels. Cette approche vise à optimiser les études traditionnelles *in vitro* et *in vivo*. Par exemple, elle permet d'explorer des hypothèses par simulation numérique, avant le lancement d'essais cliniques conventionnels, pour augmenter leur probabilité de succès ou identifier plus précocement des échecs.

Une recherche spécifique a été réalisée sur ces essais. Parmi les 30 références répertoriées sur *Medline*, aucune recommandation ni de rapport d'évaluation technologique, méta-analyse ou revue systématique de la littérature relative aux études *in silico* dans le domaine des dispositifs médicaux n'ont été identifiés.

Des travaux et pistes de réflexion, notamment aux États-Unis et en Europe, ont été recherchés afin d'identifier les perspectives de développement de ces essais.

- Données de la *Food and Drug Administration* (FDA), États-Unis

La FDA a lancé en 2017 un plan d'action pour l'innovation dans le domaine de la santé numérique (20), avec notamment un programme pilote de pré-certification des logiciels visant à adapter et à améliorer l'efficacité du processus réglementaire de ces technologies et la création d'un centre d'excellence de la santé digitale « Center of Excellence for Digital Health ».

La FDA⁶ souligne les bénéfices en santé publique de la modélisation et la simulation, incluant les essais cliniques *in silico*. Ces essais sont définis comme l'utilisation de la simulation informatique individualisée pour le développement ou l'évaluation réglementaire de produits, de dispositifs médicaux ou d'interventions médicales.

La FDA encourage le développement de la modélisation et la simulation et préconise leur utilisation dans les situations suivantes :

- 1) prévision de résultats cliniques ;
- 2) information sur la conception des essais cliniques à mettre en œuvre ;
- 3) information sur les données d'efficacité ;
- 4) identification de la population cible la plus pertinente ;

⁶ <https://www.fda.gov/ScienceResearch/AboutScienceResearchatFDA/ucm616822.htm> [consulté le 04/01/2019].

5) prévision de la tolérance du produit.

La FDA précise que les essais cliniques *in silico* peuvent être utilisés pour remplacer les essais cliniques chez l'homme, en particulier ceux destinés à évaluer le risque d'interactions médicamenteuses.

Un partenariat public-privé réunissant des représentants de la FDA, de l'industrie, des organisations à but non lucratif et des associations de patients, le *Medical Device Innovation Consortium* (MDIC), aura notamment pour objectif d'explorer les perspectives de remplacement des essais cliniques par la modélisation et la simulation pour l'évaluation clinique des dispositifs médicaux (21)⁷.

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

En l'absence de recommandations, la Commission européenne a fondé une action de soutien « Avicenna » entre 2013 et 2015 dont l'objectif est d'élaborer une feuille de route (22) sur les essais *in silico* « in silico clinical trials » relatifs aux produits de santé (médicaments et dispositifs médicaux) « biomedical products ».

La feuille de route a été élaborée sur 18 mois à partir d'un groupe de 525 experts internationaux (35 pays) issus de l'industrie, de la recherche académique, des CRO (*Contract Research Organizations*), développeurs d'essais cliniques *in silico*, agences (NICE, FDA, EMA, etc.), associations de patients, etc. Elle a fait l'objet d'une publication (Viceconti *et al.*, 2016 (23)).

Cette feuille de route a établi un état des lieux de la stratégie de développement clinique actuelle des produits de santé et a formulé des perspectives d'utilisation des essais *in silico* (via 36 recommandations).

Concernant la stratégie de développement clinique actuelle des produits de santé, le Consortium souligne que la seule voie d'accès à l'évaluation de l'efficacité et de la sécurité d'un produit de santé est de le tester sur des organismes vivants, à savoir les animaux (en phase d'évaluation préclinique) et les humains (en phase d'évaluation clinique). L'évaluation préclinique est une étape essentielle pour le développement des produits de santé potentiels. En raison de la complexité des pathologies humaines et des variabilités individuelles, il est rappelé qu'un produit de santé peut donner d'excellents résultats en laboratoire mais non confirmés lors de la phase préclinique.

Le Consortium a établi une définition consensuelle des essais cliniques *in silico* qui est la suivante : utilisation de la simulation informatique individualisée pour le développement ou l'évaluation réglementaire de produits, de dispositifs médicaux ou d'interventions médicales. C'est un sous-domaine de la médecine *in silico*, discipline qui englobe l'utilisation de simulations informatiques individualisées dans tous les aspects de la prévention, du diagnostic, de l'évaluation pronostique et du traitement de la maladie.

Les études de simulation et modélisation sont déjà utilisées pour le développement de produits de santé, afin d'estimer les aspects pharmacocinétiques et pharmacodynamiques de nouveaux produits pharmaceutiques. Concernant les dispositifs médicaux, ces études sont utilisées pour évaluer la dynamique des fluides pour prédire comment le sang ou d'autres fluides corporels se déplacent à l'intérieur et autour du dispositif testé, ou pour analyser des éléments structurels afin de confirmer que les forces échangées entre le corps et le dispositif médical ne causeront aucun dommage.

⁷Computational modeling and simulation (CM&S): paving the way for evaluation of 21st century medical devices. Medical Device Innovation Consortium <https://mdic.org/program/computational-modeling-and-simulation-cms/> [consulté le 04/01/2019].

Les essais *in silico* pourraient permettre de :

- réduire le nombre de sujets impliqués et la durée de l'expérimentation clinique, par exemple en identifiant les patients les plus à risque de complications ou en confirmant précocement l'efficacité du produit. Les essais cliniques *in silico* pourraient ainsi être utilisés pour optimiser un essai clinique sur un faible échantillon, en identifiant les patients répondeurs au traitement ;
- optimiser les essais cliniques en fournissant des informations sur les résultats potentiels et l'interprétation des effets indésirables qui pourraient survenir, ainsi qu'une meilleure compréhension sur les interactions du produit testé avec l'anatomie et la physiologie du patient, et prédire à long terme les événements indésirables rares que les essais cliniques ont peu de chance de révéler ;
- remplacer partiellement les essais cliniques dans les situations où les essais cliniques *in silico* fourniraient des preuves scientifiques robustes, notamment via le remplacement de modèles animaux par des modèles *in silico* dans des conditions appropriées ;
- compléter un essai clinique en évaluant différents scénarios dans les situations multipathologiques (par exemple, lorsque le patient a des pathologies interférant avec la pathologie à évaluer, tel un diabète associé à un trouble du rythme cardiaque).

La feuille de route indique **qu'aucun essai clinique *in silico* n'a permis de remplacer l'intégralité du développement clinique et l'évaluation d'un produit de santé.** Les experts s'accordent à dire que **ces essais ne remplaceront jamais complètement les essais cliniques chez l'homme**, mais permettront de réduire et optimiser les essais cliniques conventionnels.

Le Consortium Avicenna recommande le développement et la validation de modèles *in silico* permettant l'extrapolation des résultats cliniques chez l'homme.

Il est recommandé (recommandation 23) une collaboration entre les développeurs afin :

- de mettre en place un cadre d'évaluation *in silico* pour chaque catégorie de dispositifs, permettant aux groupes de recherche d'optimiser les prédictions des différentes défaillances ;
- d'évaluer rétrospectivement un certain nombre de produits de santé pour lesquels le résultat clinique est bien connu, afin de renforcer la confiance dans les méthodes ;
- de réaliser un ISCT en double aveugle avec des essais cliniques *in vivo* existants, afin de comparer les résultats cliniques. Le Consortium précise que la priorité doit être donnée aux domaines critiques tels que les maladies rares et pédiatriques, notamment.

Les recommandations s'adressant aux organismes de régulation sont les suivantes.

- Les organismes de régulation doivent adopter les essais *in silico* et, en collaboration avec des experts universitaires et industriels, développer des normes, protocoles et partager les ressources nécessaires pour évaluer la sécurité et l'efficacité des produits de santé via les essais cliniques *in silico* (recommandation 29) ;
- Les organismes de régulation doivent considérer les essais *in silico* pour l'évaluation de produits de santé et notamment dans le domaine des dispositifs médicaux (i.e logiciels en tant que dispositifs médicaux) (recommandation 30) ;
- Les organismes de régulation européens sont invités à suivre l'approche adoptée par la *Food and Drug Administration* américaine, qui reconnaît le potentiel stratégique des essais *in silico* (recommandation 31).

Conclusion

Les essais *in silico*, fondés sur des modélisations mathématiques d'une maladie à partir de la physiologie et de la biologie, sont prometteurs. Ils sont d'ores et déjà utilisés à titre exploratoire pour évaluer la prédiction pharmacocinétique et pharmacodynamique d'un produit de santé, les aspects toxicologiques ou le ciblage des patients à inclure dans les essais cliniques conventionnels. À ce jour, ces essais font partie de l'arsenal disponible pour le développement clinique des produits de santé, en tant qu'outils complémentaires aux outils existants permettant d'optimiser la phase de développement clinique d'un produit de santé. Ils ne remplacent pas les essais cliniques chez l'homme. Le développement et la validation de modèles *in silico* sont recommandés par le Consortium *Avicenna* et la *Food and Drug Administration*, afin d'identifier les éventuelles perspectives d'extrapolation de leurs résultats pour l'évaluation clinique des produits de santé.

Aucun essai de ce type n'a toutefois été identifié dans les bases de données d'essais cliniques ou de publications.

4. Conclusion générale et position du groupe de travail

Les données transmises par les autres agences d'évaluation sollicitées, les dossiers soumis et ayant fait l'objet d'une évaluation clinique par la CNEDiMTS et les schémas d'essais cliniques recensés dans la base de données d'essais cliniques *ClinicalTrials.gov* ne permettent pas d'identifier de spécificités méthodologiques pour l'évaluation clinique des dispositifs médicaux connectés. La majorité des études cliniques reportées étaient des essais contrôlés randomisés.

Néanmoins, la complexité de l'évaluation d'un DMC a été soulignée, en raison de sa rapidité d'évolution, de son impact sur les patients et sur l'organisation des soins, ainsi que son interaction avec d'autres systèmes. L'évaluation pour ce type de DM a fait ressortir l'importance de prendre en compte l'évaluation globale et organisationnelle de la solution de technologie dans son ensemble.

Les données issues de l'analyse de la littérature sur les spécificités d'évaluation clinique des DMC reposent sur 3 guides méthodologiques et 1 revue de la littérature.

- Concernant les guides, les méthodologies d'études cliniques se rapportant à un dispositif médical connecté ne se distinguent pas de celles des autres dispositifs médicaux, le « gold standard » pour évaluer l'efficacité des technologies numériques étant l'ECR.

Le champ des technologies couvert par ces guides est large et n'est pas superposable au périmètre des dispositifs concernés par le projet de la CNEDiMTS. Néanmoins, les méthodologies d'études proposées dépendent de la finalité d'usage du DMC et de son stade de développement clinique. Les types d'études cliniques répertoriées dans ces guides regroupent les études quantitatives (expérimentales et observationnelles), les études qualitatives et les études mixtes (à la fois quantitatives et qualitatives). Le rapport d'élaboration d'un guide a répertorié d'autres études, notamment des études d'ingénierie et des schémas d'étude dits « innovants », tels que l'estimation des variables instrumentales et l'utilisation du big data.

- Concernant la revue de la littérature, les méthodologies des essais cliniques impliquant des applications mobiles en santé et recensés dans la base de données *ClinicalTrials.gov* ne se distinguent pas de celles habituellement décrites dans les dossiers d'évaluation pour les différentes catégories de dispositifs médicaux soumis à la CNEDiMTS. La plupart des études étaient interventionnelles et impliquaient un ECR.

Les 8 publications complémentaires abordant la problématique de l'évaluation clinique des technologies numériques ont identifié des particularités liées aux DMC et ont proposé des adaptations méthodologiques. Ces particularités concernent la rapidité d'évolution de la technologie, les interrelations dynamiques avec les participants à n'importe quelle étape du process, la collecte automatique, le traitement et la transmission des données. L'enjeu étant de concilier l'évaluation clinique avec ces particularités, et notamment avec le rythme des évolutions. Un certain nombre de propositions convergentes ressortent, notamment la garantie de la fiabilité, de la validité et de la qualité des résultats issus des technologies numériques et la prise en compte d'une évaluation multidimensionnelle intégrant l'acceptation et l'appropriation de la technologie par le patient.

De nouvelles méthodologies se développent, notamment les méthodes de simulation et de modélisation. À ce jour, ces méthodes ne remplacent pas les essais cliniques conventionnels, elles font partie de l'arsenal disponible pour le développement clinique des produits de santé, en tant qu'outils complémentaires aux outils existants permettant d'optimiser la phase de développement clinique d'un produit de santé. Aucun essai de ce type n'a été identifié dans les bases de données d'essais cliniques ou de publications.

Le groupe de travail n'a pas identifié de méthodologie d'étude clinique spécifique pour l'évaluation des DMC ; des spécificités communes à certains DMC susceptibles d'avoir un impact sur le programme de développement clinique ont néanmoins été identifiées. L'importance des différents

temps de l'évaluation du DMC, de son évolutivité et du maintien de la qualité des résultats est soulignée.

Le guide qui en résulte (cf. « Guide sur les spécificités d'évaluation clinique d'un DMC en vue de son accès au remboursement ») a été analysé et validé par la CNEDiMTS et présenté au Collège de la HAS, pour information. Il vise à aider les entreprises qui fabriquent ou assurent l'exploitation d'un DMC et qui souhaitent un remboursement de leur technologie par l'Assurance maladie à optimiser dans leur stratégie de développement les études cliniques soutenant leur future demande de remboursement.

Ce guide est ainsi dédié aux industriels qui s'appêtent à mettre en place une étude clinique visant à démontrer l'intérêt de leur DMC.

5. Annexes

Annexe 1. Méthode de travail

Les principales étapes sont décrites ci-dessous.

• Phase de cadrage

Pendant la phase de cadrage, des réunions ont été organisées avec :

- les représentants des institutions (DGS, DSS, Caisses d'Assurance maladie – CNAMTS, RSA, MSI, DGOS, ANSM, INCa) ;
- les fabricants et les représentants des fabricants (SNITEM) ;
- les prestataires et les représentants des prestataires (UPSADI, FEDEPSAD, SIDIV, UNPDM, SNADOM).

Une feuille de route a été élaborée et a été soumise à la CNEDiMTS pour avis avant publication sur le site Internet de la HAS (24).

• Recueil et analyse des données

- Une revue systématique de la littérature scientifique a été réalisée : la recherche documentaire mise en œuvre est détaillée en annexe 2.
- Les syndicats de fabricants et prestataires ont été sollicités sur les spécificités méthodologiques qu'ils auraient d'ores et déjà identifiées pour les DMC, et notamment lorsque :
 - le DMC a un impact sur l'organisation des soins.
 - et lorsqu'il intègre un algorithme apprenant.

Le rapport bibliographique a été envoyé au groupe de travail avant la première réunion.

• Constitution du groupe de travail

Un groupe de travail multidisciplinaire, constitué de professionnels de santé représentatifs, a été mis en place. Ce groupe est composé de méthodologistes, cliniciens, patient et biostatisticien. Un appel à candidature a également été diffusé sur le site Internet de la HAS et sur le réseau Twitter.

• Élaboration et envoi d'un questionnaire aux parties prenantes

Un questionnaire (annexe 3) a été envoyé aux parties prenantes identifiées avant la première réunion du groupe de travail.

Les informations fournies par les parties prenantes ont été communiquées au groupe de travail.

• Position du groupe de travail

Il a été demandé au groupe de travail, lors des réunions, de :

- valider (éventuellement compléter) le rapport bibliographique ;
- analyser les réponses transmises par les parties prenantes suite à l'envoi du questionnaire ;
- élaborer un guide sur les spécificités méthodologiques d'évaluation clinique des DMC.

Le groupe de travail valide les documents élaborés.

• Étape de relecture du guide

Le guide issu des réflexions du groupe de travail a été envoyé aux différents acteurs sollicités lors du cadrage ainsi qu'aux parties prenantes identifiées (cf. grille de lecture en annexe 4 et liste des contributions à la relecture du guide en annexe 5). Suite à l'étape de relecture, les différents commentaires ont été analysés par le groupe de travail pour modification éventuelle et validation du guide dans sa version définitive.

• Examen par la CNEDiMTS

Les documents finalisés (guide et son rapport d'élaboration) ont été analysés et validés par la CNEDiMTS.

• Examen par le Collège de la HAS

Les documents finalisés (guide et son rapport d'élaboration) ont été présentés au Collège de la HAS, pour information.

• Diffusion des documents

Les documents finalisés (guide et son rapport d'élaboration) sont mis en ligne sur le site Internet de la HAS.

Annexe 2. Recherche documentaire

1. Méthode

Une recherche documentaire systématique a été réalisée de janvier 2013 à octobre 2018, par interrogation des bases de données bibliographiques médicales :

- *Medline (National Library of Medicine, États-Unis)* ;
- *The Cochrane Library (Wiley Interscience, États-Unis)* ;
- les sites Internet publiant des recommandations, des rapports d'évaluation technologique ou économique (voir liste des sites consultés) ;
- les sites Internet de sociétés savantes et autres organismes compétents dans le domaine étudié (voir liste des sites consultés).

La recherche a été limitée aux publications en langue anglaise et française.

Les évaluations technologiques, les recommandations, les conférences de consensus, les méta-analyses, les revues systématiques, les études cliniques ont été recherchées.

Cette recherche automatisée a été complétée par une recherche manuelle réalisée à partir des références citées dans les articles retenus. La bibliographie des experts et les données des industriels ont également été considérées comme sources d'information. Lorsque les études transmises par ces derniers répondaient aux critères de sélection définis pour la recherche documentaire, elles ont été intégrées à la bibliographie.

2. Base de données bibliographiques *Medline*

La stratégie de recherche dans la base de données bibliographiques *Medline* est construite en utilisant, pour chaque sujet, soit des termes issus de thésaurus (descripteurs du MESH pour *Medline*), soit des termes libres (du titre ou du résumé). Ils sont combinés avec les termes décrivant les types d'études.

Le tableau 6 présente la stratégie de recherche dans la base de données *Medline*. Les nombres de références indiqués correspondent aux références uniques identifiées.

Tableau 6. Stratégie de recherche documentaire

Type d'étude/Sujet	Termes utilisés	Période de recherche	Nombre de références
Évaluation des DM connectés		01/2013 – 10/2018	1 071
Étape 1	(((((«Equipment and Supplies»[Majr] OR device[ti]) and («Telecommunications»[Majr] OR connected[ti] OR mhealth[ti] OR mobile[ti]))) OR («Artificial Intelligence»[MAJR] OR «deep learning»[Ti] OR «Mobile Applications»[MAJR] OR «Wearable Electronic Devices»[MAJR] OR wearable[ti] OR «Telecommunications»[MAJR] OR app[ti] OR apps[ti] OR «digital health»[ti] OR mhealth[ti] OR «mobile health»[ti] OR «mobile-health»[Ti] OR «Medical Informatics Applications»[Majr])))		
ET Étape 2	("Clinical Trials as Topic"[Mesh] OR "Research Design"[Mesh] OR "clinical evaluation"[ti] OR "clinical outcome"[ti] OR "Research Design"[ti] OR "trial design"[ti])		
Méta-analyses, revues de littérature		01/2013 – 05/2018	204

Étape 1 ET Étape 2 ET Étape 3	metaanalys*[TI] OR meta-analys*[TI] OR meta analysis[TI] OR systematic review*[TI] OR systematic overview*[TI] OR systematic literature review*[TI] OR systematical review*[TI] OR systematical overview*[TI] OR systematical literature review*[TI] OR systematic literature search[TI] OR pooled analysis[TI] OR meta-analysis[PT] OR cochrane database syst rev[TA]		
Essais cliniques		01/2013 – 05/2018	335
Étape 1 ET Étape 2 ET Étape 4	Clinical Trial[ptyp]		
Évaluation des technologies de santé		01/2013 – 05/2018	64
Étape 1 ET Étape 5	"Technology assessment, biomedical"[MAJR]		
Essais cliniques <i>in silico</i>		01/2013 – 05/2018	30
Étape 1 ET Étape 4 ET Étape 6	"Computer Simulation"[Majr] OR "in silico"[ti] OR in-silico[ti]		
	Nombre total de références obtenues		1 163
	Nombre total d'articles analysés		75

3. Sites consultés

Dernière consultation : 10/2018

Information francophone :

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 anglophone :

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 Interventional 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
Centre fédéral d'expertise des soins de santé – 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. Essais cliniques

Les essais cliniques ont été recherchés sur ClinicalTrials.gov :
Une recherche a été faite avec le terme mhealth (415 références).
Une autre recherche a été faite avec algorithm (482 références).

5. Veille

En complément de la recherche une veille a été réalisée jusqu'au 10/12/18 sur Internet et sur les sommaires des journaux, en particulier des sommaires de journaux suivants :

- *Journal of Medical Internet Research* JMIR;
- *JMIR Research Protocols* ;
- *JMIR mHealth and uHealth*.

Résultats

Nombre de références uniques identifiées (sites consultés et Medline) : 22 sur les sites Web et 1 163 dans Medline
Nombre de références analysées : 75
Nombre de références retenues : 10

Annexe 3. Questionnaire envoyé aux parties prenantes

Spécificités méthodologiques d'évaluation clinique des Dispositifs Médicaux Connectés Questionnaire parties prenantes

Vous êtes sollicité afin de partager votre expérience ou point de vue sur les éventuelles spécificités méthodologiques qui doivent être prises en compte pour l'évaluation clinique des DMC. Cette sollicitation intervient dans le cadre d'un projet dont l'objectif est de produire un guide méthodologique d'évaluation clinique dédié, afin d'éclairer les industriels qui s'appêtent à mettre en place une étude clinique visant à démontrer l'intérêt de leur DMC.

Nous vous proposons de répondre directement dans le document et vous remercions par avance pour votre collaboration sur ces travaux.

Les données cliniques transmises par l'industriel sont analysées selon les critères de la médecine fondée sur les preuves. La CNEDiMTS détermine la pertinence clinique et la transposabilité dans le système de soins français.

Une étude contrôlée randomisée conçue et réalisée en double aveugle (ou du moins avec un observateur indépendant) est l'étude offrant le meilleur niveau de preuve. Toutefois, dans certaines situations, ce type d'étude peut ne pas être possible pour évaluer les DM. La commission prend en compte ce contexte et peut adapter ses exigences, sous réserve que l'industriel explique et argumente cette impossibilité.

La CNEDiMTS a d'ores et déjà répertorié des situations dans lesquelles l'essai contrôlé randomisé conventionnel n'est pas possible et a identifié les méthodes et conditions permettant une évaluation clinique de qualité dans le « [Guide méthodologique pour le développement clinique des dispositifs médicaux](#) ».

La CNEDiMTS a également élaboré un document d'information dont l'objectif est de donner des points de repère pratiques sur les principes de l'évaluation de la CNEDiMTS dans le guide méthodologique « [Principes d'évaluation relatifs aux dispositifs médicaux à usage individuel en vue de leur accès au remboursement](#) ».

Les attentes de la CNEDiMTS peuvent varier selon :

- **la catégorie de DM** : la définition du dispositif médical⁸ implique une importante hétérogénéité du monde des DM comportant des produits aussi divers que, par exemple, des pansements, des lunettes de vue, des stimulateurs cardiaques ou des appareils d'imagerie médicale. La commission prend par exemple en compte le contexte inhérent à certains produits d'assistance, DM ou non.

⁸ Règlement (UE) 2017/745 du Parlement européen et du Conseil du 5 avril 2017 relatif aux dispositifs médicaux, modifiant la directive 2001/83/CE, le règlement (CE) n° 178/2002 et le règlement (CE) n° 1223/2009 et abrogeant les directives du Conseil 90/385/CEE et 93/42/CEE. 2017. Journal officiel de l'Union européenne 2017.

Le règlement européen 2017/745 définit un dispositif médical (DM) comme étant « *tout instrument, appareil, équipement, logiciel, implant, réactif, matière ou autre article, destiné par le fabricant à être utilisé, seul ou en association, chez l'homme pour l'une ou plusieurs des fins médicales précises suivantes :*

- *diagnostic, prévention, contrôle, prédiction, pronostic, traitement ou atténuation d'une maladie ;*
- *diagnostic, contrôle, traitement, atténuation d'une blessure ou d'un handicap ou compensation de ceux-ci ;*
- *investigation, remplacement ou modification d'une structure ou fonction anatomique ou d'un processus ou état physiologique ou pathologique ;*
- *communication d'informations au moyen d'un examen in vitro d'échantillons provenant du corps humain, y compris les dons d'organes, de sang et de tissus ;*

et dont l'action principale voulue dans ou sur le corps humain n'est pas obtenue par des moyens pharmacologiques ou immunologiques ni par métabolisme, mais dont la fonction peut être assistée par de tels moyens. »

L'évaluation clinique des produits, centrée sur les besoins de la personne en situation de handicap, ses habitudes de vie, son projet de vie et son environnement, est toujours privilégiée. Toutefois pour un certain nombre de catégories de produits (supports d'aide à la prévention d'escarres, pieds à restitution d'énergie par exemple), une évaluation technique, soumise au respect de normes, est prérequis.

- **le cycle de vie du DM** : il peut être plus court que celui de l'investigation clinique.
Pour tout dispositif, il est essentiel, dans la limite du possible, d'anticiper l'évolution incrémentale dans le plan de développement clinique du DM. C'est un élément clé pour la prise en compte des données cliniques aux nouvelles gammes (nouveaux modèles).
- **les possibilités de recrutement limitées dans une étude clinique**, en raison de la faible population cible.

La réalisation d'études multicentriques internationales permet d'augmenter le nombre de sujets susceptibles d'être inclus. Dans tous les cas, un avis favorable de la CNEDiMITS ne peut être émis au vu du bénéfice attendu pour le patient que lorsque le risque inhérent à ces DM est suffisamment caractérisé et limité selon le type de DM.

Concernant plus particulièrement certains DMC, les spécificités d'ores et déjà identifiées qui pourraient avoir un impact sur leur évaluation clinique peuvent être liées à leur mode d'action et aux interactions avec l'environnement d'utilisation, au cycle de développement encore plus rapide, à l'évolution permanente du produit ou de l'algorithme intégré dans le DMC, à la collecte automatique des données, à l'aptitude / capacité du patient à utiliser le DMC, aux aspects organisationnels engendrés par l'utilisation du DMC impactant les patients, leur qualité de vie et/ou l'organisation des soins.

L'objectif est d'identifier dans quelle mesure ces spécificités peuvent impacter l'évaluation clinique de certains DMC et les méthodes pour y remédier.

Partie 1 - Cycle de vie court de certains DMC

Le temps requis pour la réalisation de l'étude clinique est susceptible d'outrepasser le cycle de vie de certains DMC.

Question 1.1 – Existe-t-il des méthodologies d'études cliniques plus rapides qu'une étude clinique conventionnelle tout en restant robustes applicables aux DMC ? Si oui lesquelles ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Argumentaire

Question 1.2 – Quels seraient les paramètres de l'étude clinique qui seraient impactés par ces particularités méthodologiques ?

Réponse

- Type
- Durée
- Critères de jugement
- Taille de l'échantillon
- Méthode de randomisation
- Méthode de recueil et d'analyse des résultats
- Autre

Argumentaire

Question 1.3 – Ces méthodologies d'études cliniques ont-elles déjà été mises en place pour certains DMC ? Si oui quelles sont les références de ces études ?

Réponse

OUI NON NE SE PRONONCE PAS

Argumentaire

Partie 2 - Évolution permanente de certains DMC ou algorithme(s)

Quel que soit le DMC et le type d'algorithme utilisé (déterministe ou fonctionnant par apprentissage supervisé), des mises à jour constantes peuvent être réalisées.

Question 2.1 – Existe-t-il des méthodologies d'études cliniques permettant d'intégrer l'évolution permanente de certains DMC ou algorithme(s) ? Si oui lesquelles ?

Réponse

OUI NON NE SE PRONONCE PAS

Argumentaire

Question 2.2 – Quels seraient les paramètres de l'étude clinique qui seraient impactés par ces particularités méthodologiques ?

Réponse

- Type
- Durée
- Critères de jugement
- Taille de l'échantillon
- Méthode de randomisation
- Méthode de recueil et d'analyse des résultats
- Autre

Argumentaire

Question 2.3 – Ces méthodologies d'études cliniques ont-elles déjà été mises en place pour certains DMC ? Si oui quelles sont les références de ces études ?

Réponse

OUI NON NE SE PRONONCE PAS

Argumentaire

Partie 3 - Évolutivité des algorithmes fonctionnant par apprentissage supervisé

Une des particularités de certains DMC est l'évolutivité des algorithmes intégrés fondés sur des processus automatiques apprenants (sont concernés uniquement les algorithmes apprenants fonctionnant par apprentissage supervisé). Ces algorithmes sont conçus de sorte que leur comportement évolue dans le temps, en fonction des données fournies. Leurs résultats sont ainsi mouvants et dépendants à chaque instant de la base d'apprentissage qui leur a été fournie et qui évolue elle-même au fur et à mesure de leur utilisation.

Question 3.1 – Existe-t-il des méthodes d'évaluation clinique permettant de prendre en considération l'évolutivité des algorithmes ?

Réponse

OUI NON NE SE PRONONCE PAS

Question 3.2 – Si oui, quelles sont-elles ?

Réponse

Question 3.3 – Ces méthodologies d'études cliniques ont-elles déjà été mises en place pour certains DMC ? Si oui quelles sont les références de ces études ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Argumentaire

Question 3.4 – Peut-on « figer » à un instant T l'algorithme pour permettre son évaluation clinique ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Question 3.5 – Peut-on imposer une garantie de non-régressivité de l'algorithme avant la nouvelle évaluation clinique ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Question 3.6 – Si oui, quelles en seraient les modalités de mise en œuvre (exigences, critères, etc.) ?

Réponse

Partie 4 - Recueil des données cliniques

Une des particularités de certains DMC est liée au fait qu'ils peuvent générer en permanence des données émanant du patient via des applications ou des capteurs, notamment.

Question 4.1 – Afin d'assurer l'exhaustivité et garantir la pertinence de l'analyse, peut-on exiger dans le protocole de l'essai clinique un cahier des charges de recueil des données collectées ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Argumentaire

Certains DMC peuvent générer une masse importante de données (big data).

Question 4.2 – L'approche « big data » va-t-elle modifier la robustesse des méthodes d'évaluation clinique conventionnelles des DMC ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Question 4.3 – Si oui, quelles sont les méthodologies permettant d'assurer la robustesse de leur évaluation clinique ? Quels seraient les paramètres de l'étude clinique qui seraient impactés par ces particularités méthodologiques ?

Réponse

- Type
- Durée
- Critères de jugement
- Taille de l'échantillon
- Méthode de randomisation
- Méthode de recueil et d'analyse des résultats
- Autre

Argumentaire

Question 4.4 – Ces méthodologies d'études cliniques ont-elles déjà été mises en place pour certains DMC ? Si oui quelles sont les références de ces études ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Argumentaire

Partie 5 - Mode d'action et interactions de certains DMC avec l'environnement d'utilisation

Question 5.1 – Faut-il isoler l'effet du DMC ou prendre en compte l'effet du DMC et de l'environnement d'utilisation ?

Réponse

Partie 6 - Aspects organisationnels engendrés par l'utilisation de certains DMC (impactant les patients, leur qualité de vie et/ou l'organisation des soins)

Certains DMC nécessitent une formation à l'utilisation, une appropriation du DMC par le patient et son adhésion au fil du temps.

Question 6.1 – Existe-t-il des méthodologies d'études cliniques permettant d'intégrer ces particularités ? Si oui lesquelles ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Argumentaire

Question 6.2 – Quels seraient les paramètres de l'étude clinique qui seraient impactés par ces particularités méthodologiques ?

Réponse

- Type
- Durée
- Critères de jugement
- Taille de l'échantillon
- Méthode de randomisation
- Méthode de recueil et d'analyse des résultats
- Autre

Argumentaire

Question 6.3 – Ces méthodologies d'études cliniques ont-elles déjà été mises en place pour certains DMC ? Si oui quelles sont les références de ces études ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Argumentaire

Question 6.4 – Outre les critères cliniques de morbi-mortalité, d'autres critères sont essentiels pour l'évaluation de l'efficacité de certains DMC, tels que la satisfaction, l'aptitude / la capacité du patient à utiliser le DMC « usability », l'observance et la qualité de vie des patients. Y a-t-il des critères / échelles spécifiques aux DMC ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Argumentaire

Question 6.5 – Parmi les critères cités ci-dessus, est-il envisageable de définir *a priori* certains seuils, tels que des seuils minimaux d'utilisation, nombre de connections aux applications, etc. ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Argumentaire

Question 6.6 – Est-il possible d'exiger une description stricte des modalités organisationnelles du suivi à distance du patient dans le protocole de l'essai clinique afin d'établir une stratégie de référence ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Partie 7 - Autre(s) spécificité(s) propre(s) aux DMC

Question 7.1 – Avez-vous identifié d'autres particularités propres à certains DMC pouvant impacter leur évaluation clinique (par exemple, liées au type d'algorithme utilisé) ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Question 7.2 – Si oui quelles sont-elles ?

Réponse

Question 7.3 – Quels seraient les paramètres d'étude clinique qui seraient impactés par ces particularités ?

Réponse

- Type
- Durée
- Critères de jugement
- Taille de l'échantillon
- Méthode de randomisation
- Méthode de recueil et d'analyse des résultats
- Autre

Argumentaire

Question 7.4 – Ces méthodologies d'études cliniques ont-elles déjà été mises en place pour certains DMC ? Si oui quelles sont les références de ces études ?

Réponse

OUI

NON

NE SE PRONONCE PAS

Argumentaire

Merci de nous indiquer vos coordonnées :

Nom Prénom :

Fonction :

Entité :

Téléphone :

E-mail :

Annexe 4. Grille de lecture

Projet de guide méthodologique sur les spécificités d'évaluation clinique d'un dispositif médical connecté (DMC) en vue de son accès au remboursement

Questionnaire adressé au groupe de lecture

Entité :

Nom, Prénom :

Fonction :

Téléphone :

E-mail :

COMMENTAIRES SUR LA FORME (LE DOCUMENT FERA L'OBJET D'UNE MISE EN PAGE PAR PAO)

Éléments (+)	Éléments (-)

COMMENTAIRES « INTRODUCTION » (PAGES 4&5)

Commentaires :

COMMENTAIRES « DMC EVALUES PAR LA CNEDIMTS ET CONCERNES PAR CE GUIDE METHODOLOGIQUE » (PAGE 6)

Commentaires :

COMMENTAIRES « LES CRITERES D'EVALUATION DE LA CNEDIMTS EN VUE DE L'ADMISSION AU REMBOURSEMENT D'UN DM » (PAGE 7)

Commentaires :

**COMMENTAIRES « LES CARACTERISTIQUES COMMUNES À CERTAINS DMC »
(PAGES 8 À 10)**

Commentaires :

**COMMENTAIRES « UNE ADAPTATION DU DEVELOPPEMENT CLINIQUE DES
DMC » (PAGES 11 A 13)**

Commentaires :

**COMMENTAIRES « LES ENJEUX DE L'ÉVALUATION DE LA HAS »
(PAGES 13 A 15)**

Commentaires :

COMMENTAIRES « CONCLUSION » (PAGE 16)

Commentaires :

Annexe 5. Liste des contributions à la relecture du guide

Le guide a été soumis à 64 personnes ou entités concernées par cette thématique. Nous remercions les 25 relecteurs listés ci-dessous pour leur contribution à ce guide.

- ▶ **Christine BALAGUÉ**, Institut Mines Telecom BS
- ▶ **Laure BELAYA**, Ingénieur
- ▶ **Pierre-Yves BENHAMOU**, Expert clinicien
- ▶ **Jean-Jacques CARRE**, CD Healthcare consulting
- ▶ **Mehdi CHERAITIA**, Neogia
- ▶ **David COULON**, @Health
- ▶ **Cécile DELVAL**, Air liquide santé international
- ▶ **Fabrice DENIS**, Expert clinicien
- ▶ **Claire DESFORGES**, Fédération française des diabétiques
- ▶ **Anne-Aurélié EPIS DE FLEURIAN**, Syndicat national de l'industrie des technologies médicales (SNITEM)
- ▶ **Michel GALINIER**, Expert clinicien
- ▶ **Laëtitia GAMBOTTI**, Institut national du cancer (INCa)
- ▶ **Olivier GOEAU-BRISSENIERE**, 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 clinicien
- ▶ **Laure MILLET**, Institut Montaigne
- ▶ **Gill MORISSE**, Quantmetry
- ▶ **Eric RENARD**, Expert clinicien
- ▶ **Nicolas ROCHE**, Société de pneumologie de langue française
- ▶ **Michel VICAIRE**, Association des insuffisants respiratoires
- ▶ **Mobin YASINI**, DMD Santé
- ▶ **Sarah ZOHAR**, Méthodologiste

Annexe 6. Liste des tableaux

Tableau 1. Données transmises par les agences du réseau INAHTA.....	12
Tableau 2. Catégories de DMC évaluées par la CNEDiMTS suite à des dépôts de dossiers par les industriels.....	14
Tableau 3. Classification fonctionnelle selon la finalité d'usage de la technologie	23
Tableau 4. Exigences pour les niveaux de classification fonctionnelle 3A et 3B.....	25
Tableau 5. Exemple de stratégie d'évaluation d'une application mobile.....	35
Tableau 6. Stratégie de recherche documentaire	49

Annexe 7. Liste des figures

Figure 1. Schémas d'études mentionnés dans le rapport du YHEC, 2016 (16)	29
Figure 2. Schémas d'études mentionnés dans le guide de l'OMS, 2016 (9)	36

Références

1. Commission nationale informatique et libertés. Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle. Synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion éthique confiée par la loi pour une république numérique. Paris : CNIL; 2017.
https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_garder_la_main_web.pdf
2. Pham Q, Wiljer D, Cafazzo JA. Beyond the randomized controlled trial: a review of alternatives in mHealth Clinical Methods. *JMIR Mhealth Uhealth* 2016;4(3):e107.
3. Food and Drug Administration. Multiple function device products: policy and considerations. Draft guidance for industry and Food and Drug Administration. Silver Spring: FDA; 2018.
https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM605683.pdf?utm_campaign=New%20Draft%20FDA%20Guidance%20for%20Devices%20with%20More%20than%20One%20Function&utm_medium=email&utm_source=Eloqua&elqTrackId=D8BD358D08D83E662DE37E56D5C76073&elq=b61855a317b748aab34e0c58bc26eb99&elqaid=3302&elqat=1&elqCampaignId=2494
4. U.S. Food and Drug Administration. Software as a Medical Device (SAMd): Clinical evaluation. Guidance for industry and Food and Drug Administration Staff. Washington: U.S. Department of Health and Human Services; 2017.
<https://www.federalregister.gov/documents/2017/12/08/2017-26441/software-as-a-medical-device-clinical-evaluation-international-medical-device-regulators-forum>
5. Food and Drug Administration, Center for Devices and Radiological Health, Center for Biologics Evaluation and Research. Mobile medical applications. Guidance for Industry and Food and Drug Administration Staff. Silver Spring: FDA; 2015.
<https://www.fda.gov/downloads/MedicalDevices/.../UCM263366.pdf>
6. York Health Economic Consortium, Craig J, Shore J. Evidence guide for Apps Developers: Study Designs Including Applied Examples for Apps. Evidence for Evaluation Purposes: Report. York: University of York; 2016.
<http://www.yhec.co.uk/yhec-content/uploads/2017/02/Study-Designs-Including-Applied-Examples-for-Apps.pdf>
7. Ministry of Health. Guidance on evaluating or developing a health app. Wellington: MH; 2017.
<https://www.health.govt.nz/system/files/documents/pages/guidance-evaluating-developing-health-app-oct17-v2.pdf>
8. National Institute for Health and Care Excellence. Evidence standards framework for digital health technologies. London: NICE; 2018.
<https://www.nice.org.uk/medicines/evidence-standards-framework-for-digital-health-technologies>
9. World Health Organization. Monitoring and evaluating digital health interventions. A practical guide to conducting research and assessment. Geneva: WHO; 2016.
<http://www.who.int/reproductivehealth/publications/mhealth/digital-health-interventions/en/>
10. DeVito Dabbs A, Song MK, Myers B, Hawkins RP, Aubrecht J, Begey A, *et al.* Clinical trials of health information technology interventions intended for patient use: unique issues and considerations. *Clin Trials* 2013;10(6):896-906.
11. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, *et al.* Mobile health technology evaluation: the mHealth evidence workshop. *Am J Prev Med* 2013;45(2):228-36.
12. Ben-Zeev D, Schueller SM, Begale M, Duffecy J, Kane JM, Mohr DC. Strategies for mHealth research: lessons from 3 mobile intervention studies. *Adm Policy Ment Health* 2015;42(2):157-67.
13. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, *et al.* Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016;51(5):843-51.
14. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. *J Med Internet Res* 2017;19(6):e232.
15. Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology* 2018;286(3):800-9.
16. York Health Economic Consortium, Craig J, Shore J. RX137: Evidence generation guide for apps and wearables developers: study designs including applied examples. York: University of York; 2016.
<http://www.yhec.co.uk/yhec-content/uploads/2017/03/YHEC-Study-Designs-28.03.17.pdf>
17. York Health Economic Consortium, Craig J, Shore J. RX137 Digital health evidence: case studies. Final report. York: University of York; 2016.
<http://www.yhec.co.uk/yhec-content/uploads/2017/02/Case-studies-17-Feb-2017.pdf>
18. Greaves F, Joshi I, Campbell M, Roberts S, Patel N, Powell J. What is an appropriate level of evidence for a digital health intervention? *Lancet* 2018;392(10165):2665-7.
19. Mohr DC, Cheung K, Schueller SM, Hendricks Brown C, Duan N. Continuous evaluation of evolving

behavioral intervention technologies. *Am J Prev Med* 2013;45(4):517-23.

20. Food and Drug Administration. Digital health innovation action plan. Silver Spring: FDA; 2018.

<https://www.fda.gov/downloads/MedicalDevices/DigitalHealth/UCM568735.pdf>

21. Faris O, Shuren J. An FDA viewpoint on unique considerations for medical-device clinical trials. *N Engl J Med* 2017;376(14):1350-7.

22. Avicenna Consortium, Viceconti M, Morley-Fletcher E, Henney A, Contin M, El-Arifi K, *et al.* In silico Clinical Trials: How computer simulation will transform the Biomedical Industry. Brussels: European Commission; 2016.

<http://avicenna-isct.org/wp-content/uploads/2016/01/AvicennaRoadmapPDF-27-01-16.pdf>

23. Viceconti M, Henney A, Morley-Fletcher E. In silico clinical trials : how computer simulation will transform the biomedical industry. *Int J Clin Trials* 2016;3(2):37-46.

24. Haute Autorité de santé. Travaux sur les spécificités méthodologiques d'évaluation clinique des Dispositifs Médicaux Connectés. Feuille de route. Saint-Denis La Plaine: HAS; 2018.

https://www.has-sante.fr/portail/upload/docs/application/pdf/2018-04/travaux_sur_les_specificites_methodologiques_devaluation_clinique_des_dispositifs_medicaux_connectes_feuille_de_route.pdf

Participants – Groupe de travail

Le groupe de travail est composé des professionnels suivants :

- ▶ **Jean Paul BEREGI**, Radiologue, Nîmes (30)
- ▶ **Alain BERNARD**, Chirurgien thoracique et cardio-vasculaire, Dijon (21)
- ▶ **Pierre-Yves BOELLE**, Méthodologiste, Paris (75)
- ▶ **Dominique COSTAGLIOLA**, Méthodologiste, Paris (75)
- ▶ **Marie-Cécile FOURNIER**, Biostatisticienne, Nantes (44)
- ▶ **Marie-Christine JAULENT**, Méthodologiste, Paris (75)
- ▶ **Pascal SELLIER**, Adhérent d'une association de malades et d'usagers du système de santé, Paris (75)
- ▶ **Franck SEMAH**, Neurologue, Lille (59)
- ▶ **Rodolphe THIEBAUT**, Méthodologiste, Bordeaux (33)
- ▶ **Valéry TROSINI DESERT**, Pneumologue, Paris (75)

Les membres du groupe de travail ont été nommés par le bureau de la CNEDiMTS.

Conformément au décret n° 2004-1139 du 26 octobre 2004 (art. R. 161-84 à R. 161-86 du Code de la sécurité sociale), tous les membres du groupe de travail ont rempli une déclaration d'intérêts, mentionnant les liens directs ou indirects avec toute entreprise ou organisme intervenant dans le champ des missions de la HAS. Ces déclarations d'intérêts ont été rendues publiques sur le site Internet de la HAS.

L'analyse des déclarations d'intérêts a été réalisée selon les critères du « *Guide des déclarations d'intérêts et de gestion des conflits d'intérêts* » de la HAS (adopté par le Collège de la HAS le 24 juillet 2013). Un tableau récapitulatif des intérêts déclarés a été examiné par le bureau de la CNEDiMTS, qui a arrêté la composition finale du groupe de travail. Les intérêts déclarés par les experts retenus ont tous été considérés comme « non majeurs » par le bureau de la CNEDiMTS.

Le tableau récapitulatif des intérêts déclarés a été exposé et, le cas échéant, mis à jour, à partir des déclarations d'intérêts actualisées des experts au début de chaque réunion du groupe de travail et lors de la présentation de la position du groupe de travail en CNEDiMTS.

Fiche descriptive

Intitulé	Spécificités méthodologiques d'évaluation clinique d'un dispositif médical connecté (DMC). Rapport d'élaboration du guide sur ses spécificités d'évaluation clinique en vue de son accès au remboursement
Méthode de travail	Analyse de la littérature, des données des autres agences d'évaluation, des dossiers ayant été soumis à la CNEDiMTS, des bases de données d'essais cliniques, consultation des diverses parties prenantes et position d'un groupe de travail multidisciplinaire.
Objectif(s)	Les objectifs de cette évaluation visent à identifier les spécificités méthodologiques d'évaluation clinique des dispositifs médicaux connectés (DMC) en vue de leur accès au remboursement.
Demandeur	Ministère des solidarités et de la santé.
Promoteur	Haute Autorité de santé.
Pilotage du projet	Estelle PIOTTO-PEYLAN, docteur en pharmacie, chef de projet au service évaluation des dispositifs (SED) (chef du service: Hubert GALMICHE ; adjointe : Corinne COLLIGNON). Secrétariat : Yakaré TOUNKARA, assistante.
Participants	Cf. liste des participants.
Recherche documentaire	Réalisée par Frédérique PAGES, chef du service documentation-information des publics, avec l'aide de Maud LEFEVRE, assistante documentaliste.
Auteurs de l'argumentaire	Estelle PIOTTO-PEYLAN, chef de projet au SED. L'analyse des bases de données d'essais cliniques (paragraphe 3.3) et une partie des données issues des dossiers ayant été soumis à la CNEDiMTS (paragraphe 3.2) ont été réalisées par Cyril OLIVIER, chef de projet au SED. Sous la responsabilité de Hubert GALMICHE (chef du service) et Corinne COLLIGNON (adjointe au chef de service).
Validation	Examen par la Commission nationale d'évaluation des dispositifs médicaux et des technologies de santé (CNEDiMTS) : janvier 2019. Présentation au Collège de la HAS, pour information.
Autres formats	Pas d'autre format.
Document d'accompagnement	Guide sur les spécificités d'évaluation clinique d'un DMC en vue de son accès au remboursement. L'objectif de ce guide est d'aider les entreprises qui fabriquent ou assurent l'exploitation d'un DMC à intégrer dans leur stratégie de développement les études cliniques qui permettront de déterminer l'intérêt du DMC, en vue de son remboursement par la solidarité nationale.



Toutes les publications de la HAS sont téléchargeables sur
www.has-sante.fr