



HAUTE AUTORITÉ DE SANTÉ

MESURER
& AMÉLIORER LA QUALITÉ

REFERENTIEL


**Évaluation des Applications
dans le champ de la santé
mobile (mHealth)**

- État des lieux et critères de
qualité du contenu médical pour
le référencement des services
numériques dans l'espace
numérique de santé et le
bouquet de services des
professionnels

Validé par le collège le 24 juin 2021

Descriptif de la publication

Titre	Évaluation des Applications dans le champ de la santé mobile (mHealth)- État des lieux et critères de qualité du contenu médical pour le référencement des services numériques dans l'espace numérique de santé et le bouquet de services des professionnels
Méthode de travail	Méthode par consensus simple
Objectif(s)	Élaborer un état des lieux de l'évaluation dans le secteur de la santé mobile et définir les critères de qualité du contenu médical en vue du référencement d'Applications dans l'espace numérique de santé
Cibles concernées	Les professionnels et associations professionnelles et de patients impliqués dans le secteur de l'évaluation des Applications mobiles. La délégation ministérielle au numérique en santé (DNS)
Demandeur	La délégation ministérielle au numérique en santé (DNS)
Promoteur(s)	Haute Autorité de santé (HAS)
Pilotage du projet	Coordination : Monsieur Pierre Trudelle, chef de projet Mission numérique en santé (MNS) de la HAS (chef de service par intérim : Marc Fumey) Secrétariat : Mesdames Chantal Hottin et Sophie Duthu
Recherche documentaire	Mesdames Marie Georget et Laurence Frigère Cheffe du service documentation-veille : Madame Frédérique Pages
Auteurs	Monsieur Pierre Trudelle, chef de projet
Conflits d'intérêts	Les membres du groupe de lecture ne déclarent pas de lien d'intérêt.
Validation	Version du 24 juin 2021
Autres formats	Fichier de synthèse des critères

Ce document ainsi que sa référence bibliographique sont téléchargeables sur www.has-sante.fr 

Haute Autorité de santé – Service communication et information

5 avenue du Stade de France – 93218 SAINT-DENIS LA PLAINE CEDEX. Tél. : +33 (0)1 55 93 70 00

© Haute Autorité de santé – juin 2021 – ISBN : 978-2-11-162648-5

Sommaire

Préambule	7
1. Contexte	9
1.1. Définition et trajectoire de développement de la santé mobile	9
1.2. Les défis à relever	9
1.3. Classifications	10
1.4. Développement récent	11
1.5. Les différents domaines d'évaluation de la santé mobile	12
1.5.1. Convergence de différents secteurs et différents acteurs	13
1.5.2. Les systèmes d'évaluation des développeurs	13
1.5.3. Les systèmes d'évaluation des professionnels de santé	14
1.5.4. Les systèmes d'évaluation des usagers	14
1.5.5. Les systèmes d'évaluation des évaluateurs	15
1.5.6. Les enjeux de l'évaluation	15
1.6. Évolution quantitative de la publication scientifique	16
1.7. Niveaux de preuve et évaluation de l'utilité clinique	17
2. Risques potentiels des fonctions de base dans le cadre de la santé mobile	19
2.1. La place du risque dans l'évaluation de la santé mobile	19
2.2. Les fonctions de base les plus fréquentes	20
2.2.1. Informer	21
2.2.2. Instruire	21
2.2.3. Enregistrer des données	21
2.2.4. Calculer/analyser	22
2.2.5. Alerter/rappeler	22
2.2.6. Afficher des données	23
2.2.7. Guider	23
2.2.8. Communiquer	23
2.2.9. La variation du risque selon les fonctions de base des Applications mobiles	23
3. Les méthodes d'évaluation de la santé mobile au niveau international	26
3.1. Standard d'édition scientifique pour les publications concernant la santé mobile	26
3.1.1. Le formulaire de soumission d'Application du JMIR	26
3.1.2. Standard CONSORT-EHEALTH et mERA	26
3.1.3. Méthode pour réaliser une revue de littérature sur des Applications spécifiques	27
3.2. Les différentes approches/outils d'évaluation de la santé mobile	27
3.3. Scores, check-lists et autres échelles d'évaluation	29

3.3.1.	Score de qualité du contenu médical	29
3.3.1.1.	Échelle ou score de Silberg	30
3.3.1.2.	Brief DISCERN	30
3.3.1.3.	Synthèse	31
3.3.2.	Score général de MARS	32
3.3.2.1.	MARS-G en Allemagne	32
3.3.2.2.	u-MARS	32
3.3.3.	Les autres scores de qualité généralistes des Applications	33
3.3.4.	Les scores de qualité spécifiques des Applications	33
3.3.4.1.	Scores spécifiques	33
3.3.4.2.	Scores spécifiques développés par des sociétés savantes	34
3.3.5.	Les référentiels de qualité spécifiques des Applications	35
3.3.6.	Synthèse	35
3.4.	Avis des utilisateurs, avis d'experts	37
3.4.1.	MyHealthApps-MHA (Grande-Bretagne)	37
3.4.2.	GGD Appstore (Pays-Bas)	38
3.4.3.	Health Navigator (Nouvelle-Zélande)	39
3.4.4.	Onemind (États-Unis)	39
3.4.5.	Observatoire : App sanitaire (Italie)	40
3.4.6.	Synthèse	40
3.5.	Labels/certifications avec une liste de critères	41
3.5.1.	ORCHA Review (Grande-Bretagne)	41
3.5.2.	TICSS (région de la Catalogne-Espagne)	42
3.5.3.	Certification Dekra (France)	44
3.5.4.	Sceau DiaDigital (Allemagne)	44
3.5.5.	HON code (Suisse)	44
3.5.6.	Synthèse	45
3.6.	Répertoire éditorialisé constitué d'une base de données proposant une annotation spécifique de chaque Application	46
3.6.1.	Les référentiels d'évaluation au niveau des pays européens	46
3.6.2.	Les domaines d'évaluation de eHealth Suisse	47
3.6.3.	Répertoire du NHS Health Apps Library (Grande-Bretagne)	48
3.6.4.	AppSaludable (région Andalousie, Espagne)	50
3.6.5.	Le Portugal	50
3.6.6.	Synthèse	51
3.7.	Répertoire éditorialisé constitué d'une base de données proposant une annotation spécifique de chaque Application sur un secteur spécialisé	51
3.7.1.	<i>Our Mobile Health</i> (OMH) (Grande-Bretagne)	51
3.7.2.	Santé mentale avec AppChecker (Danemark)	52

3.7.3. AppCheck (Allemagne)	53
3.7.4. Synthèse	53
3.8. Répertoire éditorialisé constitué d'une base de données proposant une sélection spécifique d'Applications définie comme des dispositifs médicaux	54
3.8.1. mHealthBelgium (Belgique)	54
3.8.2. DiGA (Allemagne)	54
3.8.3. Digi-HTA (Finlande)	55
3.8.4. Synthèse	56
4. Critères de qualité du contenu médical	58
4.1. Les différents types de contenu de santé	58
4.2. Utilisation des critères de qualité du contenu médical	58
4.3. Conception du contenu initial	60
4.3.1. Gestion de l'information	60
4.3.1.1. Organisation autour du service de l'information	60
4.3.1.2. Expertise des auteurs du contenu dans l'Application	60
4.3.1.3. Déclarations d'intérêts	60
4.3.1.4. Citation des sources clés et références bibliographiques	61
4.3.1.5. Actualisation des sources clés et références bibliographiques	61
4.3.1.6. Niveau de preuve	61
4.4. Contenu généré et standardisation	62
4.4.1. Pertinence, fiabilité des informations et support en vie réelle	62
4.4.1.1. Qualité de la mesure dans le contexte d'utilisation	62
4.4.1.2. Support utilisateur	62
4.5. Contenu interprété	63
4.5.1. Typologie d'interprétation	63
4.5.1.1. Interprétation humaine d'un contenu de santé	63
4.5.1.2. Interprétation automatisée d'un contenu de santé	63
4.6. Contenu affiché	65
4.6.1. Compréhension et implication des utilisateurs	65
4.6.1.1. Implication des utilisateurs (patients, professionnels, personnes spécifiques)	65
4.6.1.2. Description de la destination d'usage	65
4.6.1.3. Contre-indications, risques, limitations d'usage	65
4.6.1.4. Compréhension du contenu de santé	66
4.6.2. Performance de l'affichage du contenu	66
4.6.2.1. Aide à l'utilisation/instructions	66
4.6.2.2. Lisibilité texte et image et navigation	66
4.6.2.3. Prévention des erreurs de compréhension de l'information	67

Table des annexes	68
Références bibliographiques	87
Participants	92
Abréviations et acronymes	93

Préambule

Ce document a été élaboré en réponse à la saisine de la délégation ministérielle au numérique en santé (DNS) de juillet 2020, demandant à la HAS d'élaborer un état des lieux et une liste de critères de qualité du contenu médical utilisés dans le secteur de la santé mobile (*mHealth*).

Cette liste de critères, agrégée à d'autres critères (annexe 1), sera utilisée pour contribuer au référencement des services numériques qui seront recensés dans l'espace numérique de santé (ENS)¹ et, par extension, pour le bouquet de services numériques destinés aux professionnels (BSP).

Une mise à jour du référentiel sur la santé mobile publié en 2016

Cette contribution de la HAS est issue du référentiel publié en 2016 (1) et intitulée : « Référentiel de bonnes pratiques sur les applications et les objets connectés en santé (*Mobile Health* ou *mHealth*) ». Elle visait, à l'époque les industriels et les évaluateurs (structures d'évaluation, associations de consommateurs ou sociétés savantes médicales) qui pouvaient et peuvent toujours mettre en œuvre ce référentiel pour conduire leurs propres évaluations. Ils ne portaient que sur les applications et les objets connectés n'ayant pas de finalité médicale déclarée (souvent dénommée « zone grise » des applications ou des objets connectés ayant un effet potentiel sur la santé sans être un dispositif médical). À noter que pour aider à mieux délimiter les dispositifs médicaux des autres produits, l'Agence nationale de sécurité du médicament et des produits de santé (ANSM) a publié une mise au point² sur la typologie des applications pouvant relever ou non d'un marquage CE et être qualifiées de dispositif médical. Dans ce document, le terme « Application » couvre toutes les Applications qu'elles soient un dispositif médical (DM) ou pas.

Structure du document

Ce document comprend deux parties.

Une **première partie** (chapitres 1 à 3) reprenant un état des lieux de l'évaluation de la santé mobile au travers de quelques exemples concrets. Cette partie ne se veut pas exhaustive, mais permet de documenter les différentes stratégies et moyens d'évaluation utilisés au niveau international et de contribuer à développer l'information sur ce sujet. Les articles retenus décrivent des systèmes d'évaluation avec une description centrée sur les critères. Très souvent, le type d'Application (DM ou non-DM) évaluable avec ces outils n'est pas précisé. Cela permet également de comprendre les différents niveaux d'exigence de l'évaluation en fonction du type d'application et de son contexte d'utilisation.

La **seconde partie** (chapitre 4) propose une liste de critères de qualité du contenu médical en s'appuyant sur la publication de 2016 de la HAS (1). Les membres du groupe de travail et de lecture ayant participé au document de 2016 ont été sollicités pour la relecture de cette version, associés à de nouveaux relecteurs.

¹ <https://esante.gouv.fr/virage-numerique/feuille-de-route>

² <https://ansm.sante.fr/documents/referencement/exemples-de-logiciels-et-applications-mobiles-illustrant-le-positionnement-reglementaire>

Positionnement

Ce référentiel ne se substitue pas à la loi ou la réglementation concernant les dispositifs médicaux, la protection des données personnelles et la protection des consommateurs. L'application des bonnes pratiques définies dans le présent référentiel s'entend sans préjudice de la réglementation en vigueur.

Les critères d'évaluation du contenu médical proposé dans ce document s'appliquent à tous types d'Applications. Dans le cas d'Applications médicales qualifiées de DM, certains critères proposés peuvent déjà être couverts par l'obtention du marquage CE-DM.

La HAS rappelle que le marquage CE est obligatoire pour les services numériques qualifiables de DM. Cette obligation pourrait être prise en compte dans la procédure de référencement en identifiant les critères correspondant aux exigences du marquage CE et en considérant que, pour le référencement des services numériques qualifiés de DM, ces critères sont réputés acquis.

Ce référentiel de bonnes pratiques de la HAS ne constitue pas un outil d'évaluation en vue de l'admission au remboursement.

1. Contexte

1.1. Définition et trajectoire de développement de la santé mobile

La santé mobile est un domaine de la e-santé (2). Selon l’OMS, elle recouvre « les pratiques médicales et de santé publique reposant sur des dispositifs mobiles tels que téléphones portables, systèmes de surveillance des patients, assistants numériques personnels et autres appareils sans fil³ ».

Ce domaine est actuellement dominé par les « Applications mobiles de santé et de bien-être » qui sont censées améliorer notre qualité de vie, notre bien-être et notre santé par le biais, entre autres, de retour d’informations sur nos **comportements** (3, 4) et d’**informations ciblées**.

Ces Applications mobiles peuvent se connecter à des **dispositifs médicaux** (ex. : tensiomètre), des **biocapteurs** (ex. : bracelet connecté) ou des **systèmes d’analyses de données** (ex. : programme/algorithmes de gestion de données). Les données collectées pouvant, éventuellement, entrer directement dans le dossier informatique du patient.

Le terme « *Mobile Medical Application* » (MMA) est souvent utilisé dans la littérature pour spécifier les Applications qui sont des dispositifs médicaux. Cet usage n’est pas encore systématisé.

Chaque Application vise, le plus souvent, un objectif très spécifique et personnalisé, mais le secteur dans son ensemble revendique de faire évoluer la qualité des soins en général, à accroître son efficacité et à améliorer la recherche clinique.

De plus, la santé mobile vise à rendre plus autonome et responsable les utilisateurs (repris sous le terme *empowerment* au niveau international). Cette autonomie ne signifie pas le remplacement des professionnels de santé, mais une aide pour compléter la prise en charge. Ce qui devrait permettre d’optimiser le recours aux professionnels de santé.

À noter que concernant les maladies chroniques, l’intégration de la santé mobile est parfois appelée « *minimal disruptive medicine – MDM* », car elle vise à diminuer la « charge » des soins que le patient supporte (*MDM-based mHealth intervention*) (5, 6) via l’utilisation d’Applications ou d’objets connectés dédiés.

1.2. Les défis à relever

Pour la société ORCHA (*Organisation for the Review of Care and Health Applications*), spécialisée dans l’évaluation des Applications (avec son système dénommé *ORCHA Baseline Review – OBR*), la santé mobile doit relever trois défis⁴ :

- **Prendre conscience** – Les Applications et les solutions de santé numérique de manière plus générale ne sont pas encore reconnues comme faisant naturellement partie de la gestion quotidienne des problèmes de santé et de soins.
- **Accès** – Il est très difficile d’identifier les Applications dans les magasins d’Applications généraux actuels. Les termes de recherche courants comme le diabète ou la dépression renvoient vers un certain nombre d’Applications sur la plupart des sites, mais ce n’est généralement qu’une infime partie du nombre total d’Applications réellement disponibles pour ces problèmes.

³ https://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf?sequence=1&isAllowed=y (page 6)

⁴ <https://www.orcha.co.uk/the-challenge/> (consulté le 05/11/2020)

- **Confiance** – L'absence d'un indicateur de qualité approprié empêche l'adoption de la santé numérique par les utilisateurs, les patients et les professionnels de la santé. Les préoccupations les plus fréquentes qui affectent la confiance concernent : la confidentialité des données personnelles, la sécurité, et l'efficacité clinique du produit. Alors que de nombreuses Applications de santé englobent un éventail très complexe de réglementations et de normes, il n'existe pas de service unique et consolidé pour visualiser la conformité et le respect globaux d'une Application et une indication d'utilisation claire.

Le constat d'ORCHA résume assez bien le processus de changement et d'adoption de nouvelles technologies dans l'environnement et le contexte médical.

Les objets connectés sont des entités spécifiques et les dimensions de protection des données personnelles et de sécurité des échanges sont à appliquer systématiquement (7).

1.3. Classifications

De manière très spécifique et en ciblant uniquement le contexte de la prise en charge médicale, Kumar (8), en 2013, a exposé les retours d'un groupe de travail américain sur le développement de la santé mobile et la manière de l'aborder, de classer les Applications et de les évaluer. Il propose quatre sections intégrées en un continuum : mesurer, diagnostiquer, traiter-prévenir, et santé globale. Dans cette classification, les Applications peuvent être ou ne pas être des dispositifs médicaux.

Plus récemment, en 2017, l'Institut IQVIA (9) a classé l'offre de différentes solutions ne constituant pas des dispositifs médicaux proposés aux patients selon la classification suivante :

1. Catégorie – **Bien-être et prévention** :

- exercice et activité physique ;
- régime et alimentation ;
- hygiène de vie et stress ;
- gestion du stress ;
- sommeil/insomnie ;
- arrêt du tabac ;
- modulation de la consommation d'alcool.

2. Catégorie – **Apparition des symptômes et recherche de soins** (outils du vécu du patient) :

- information générale sur la santé ;
- vérification des symptômes ;
- trouver un praticien ;
- gérer les informations financières et cliniques ;
- médias sociaux.

3. Catégorie – **Diagnostic** :

- le professionnel de santé pourrait recommander une Application pour compléter le programme de prise en charge ;
- utiliser des capteurs connectés pour le suivi à distance ;
- utiliser des Applications pour des cas d'usage spécifique lié aux problèmes rencontrés par le patient.

4. Catégorie – **Suivi du problème** (prise en charge et suivi éducatif du problème) :
 - auto-suivi (*self-monitoring*) ;
 - suivi à distance du patient (*remote patient monitoring*) ;
 - programme de réhabilitation dédié *via* une Application.

5. Catégorie – **Traitement** (remplissage de prescription et observance) :
 - prise en charge financière ;
 - remplissage de la prescription ;
 - observance de la prise en charge thérapeutique.

Cette structure n'est pas exhaustive, mais complète les classifications répertoriées dans le référentiel HAS de 2016 (1). Elle montre l'étendue « concrète » du champ de la santé mobile actualisée en 2017.

Par ailleurs, la HAS a publié en 2021⁵, une **classification fonctionnelle** selon la finalité d'usage des solutions numériques utilisées dans le cadre de soins médicaux ou paramédicaux. Quatre niveaux sont proposés allant des services support aux patients jusqu'à la gestion autonome de la décision après analyse de données.

Pour conclure

Un des enjeux majeurs de l'évaluation est de trouver le moyen de garantir un niveau de qualité suffisant par rapport aux risques que présente l'utilisation de ces Applications mobiles ou de ces objets connectés.

1.4. Développement récent

Selon les données de 2017 de l'Institut IQVIA (9), la proportion des Applications orientées « bien-être » est passée de 73 à 60 % entre 2015 et 2017. Cela signifie que 40 % des Applications étaient orientées vers l'amélioration de la santé de l'utilisateur en 2017 et que le ratio bien-être/santé est en train de se déplacer vers le secteur « santé ».

Selon IQVIA, les domaines de santé où il existe des solutions numériques sont par ordre d'importance :

- la santé mentale et les troubles du comportement (28 %) ;
- le diabète (16 %) ;
- les problèmes cardio-circulatoire (11 %) ;
- les atteintes du système nerveux (7 %) ;
- les atteintes du système musculo-squelettique et du tissu conjonctif (7 %) ;
- le cancer (5 %) ;
- les atteintes du système respiratoire (5 %) ;
- les atteintes du système digestif (4 %) ;
- les atteintes des yeux et des oreilles (4 %) ;
- la douleur (4 %) ;

⁵ https://www.has-sante.fr/jcms/p_3238360/fr/classification-fonctionnelle-selon-leur-finalite-d-usage-des-solutions-numeriques-utilisees-dans-le-cadre-de-soins-medicaux-ou-paramedicaux

- les maladies métaboliques, endocriniennes et nutritionnelles (3 %) ;
- les atteintes des tissus cutanés et sous-cutanés (3 %).

L'organisme ORCHA⁶ cite des chiffres complémentaires :

- plus de **327 000 Applications sont répertoriées** ;
- 43 Applications représentent 83 % des téléchargements ;
- 65 % n'ont pas été mises à jour depuis plus de dix-huit mois ;
- plus de 80 % des Applications ont moins de 5 000 téléchargements ;
- moins de 7 % des diabétiques utilisent une Application pour gérer leur diabète, et moins de 2 % des patients présentant une bronchopneumopathie chronique obstructive (BPCO) ;
- plus de 240 pathologies sont couvertes ;
- seulement 15 % des Applications revues par ORCHA remplissent les critères minimums d'évaluation.

Pour aller plus loin, Fagherazzi (10), en 2019, propose le concept de « digitosome » qui agrège l'ensemble des données numériques produites par un individu et que le médecin doit intégrer dans le suivi.

Finalement, le secteur de la santé mobile revendique de contribuer à délivrer une médecine personnalisée en proposant le bon traitement (par analyse d'indicateurs en continu ou d'informations ciblées), pour le bon patient (par analyse de profils ou de *patterns* similaires), au bon moment (par analyse à distance et en temps réel), tout en permettant de développer la recherche clinique (par analyse et agrégation des données).

L'utilisation régulière d'Applications mobiles en santé sur cinq catégories de populations (prévention du diabète, suivi du diabète, asthme, réhabilitation cardiaque et pulmonaire) permettrait, selon l'institut IQVIA (9), d'économiser 7 milliards de dollars par an (principalement en réduisant les hospitalisations d'urgence).

Le rapport de PwC (PricewaterhouseCoopers) (11) commandé par le GSMA (*Global System for Mobile Communications*) en 2013 projetait une économie de 99 milliards d'euros pour l'Union européenne en 2017 (surtout par une meilleure prévention).

Enfin, des études ciblées tentent de modéliser les bénéfices économiques de l'utilisation d'une Application (ex. : améliorer la qualité de vie dans le secteur de la santé mentale (12)).

1.5. Les différents domaines d'évaluation de la santé mobile

La vitesse de développement du secteur de la santé mobile a entraîné un décalage des systèmes d'évaluation « classique » qui a été qualifié de « disruptif » par un groupe de travail de l'Union européenne⁷. Sur les cinq dernières années, nous avons assisté à une très grande variété d'approches évaluatives de ces solutions numériques (label, certification, répertoire en ligne, scores, etc.). Cela peut s'expliquer pour différentes raisons, et notamment, par des univers d'évaluation spécifiques aux différents acteurs du domaine (13-16).

⁶ <https://www.orchaco.uk/the-challenge/>

⁷ http://ec.europa.eu/newsroom/document.cfm?doc_id=45251

1.5.1. Convergence de différents secteurs et différents acteurs

L'essor de la santé mobile est catalysé par la conjonction de développement technologique dans le secteur des télécommunications (Bluetooth, Wi-Fi, 5G) de l'informatique et de l'électronique (smartphone, miniaturisation des composants électroniques, traitement et stockage des données) tout en s'appuyant, *a minima*, sur quatre groupes d'acteurs :

- les ingénieurs ;
- les informaticiens ;
- les citoyens et les patients ;
- les professionnels de santé.

Chaque groupe d'acteurs utilise des systèmes d'évaluation qui sont liés aux secteurs dans lequel il évolue. Ces systèmes d'évaluation visent des objectifs spécifiques (fiabilité technique, traitement de signal, santé publique, utilité clinique, cybersécurité, etc.) et sont aussi limités au champ évalué. Quelques exemples sont proposés ci-après pour illustrer cela.

1.5.2. Les systèmes d'évaluation des développeurs

Les ingénieurs ou les informaticiens évoluent classiquement dans des systèmes d'évaluation de types « processus » :

- les systèmes d'accès à des marchés économiques (ex. : marquage CE des dispositifs médicaux) ;
- la certification de processus de fabrication (ex. : la norme ISO) ;
- l'utilisation de standard de conception (ex. : standard HL7) ;
- les normes des technologies de connectivité (ex. : *Continua Design Guidelines* – CDG).

Concernant les normes, plusieurs auront probablement une place importante dans les prochaines années. Citons, entre autres :

- ISO 14971 : application de la gestion des risques aux dispositifs médicaux ;
- ISO 27001 : technologies de l'information. Techniques de sécurité. Systèmes de management de la sécurité de l'information. Exigences ;
- IEC 62304 : logiciels de dispositifs médicaux. Processus du cycle de vie du logiciel (contient ISO 13485).

En 2021, la norme ISO-82304-2 (logiciels de santé – Partie 2 : Applications de santé et de bien-être – Qualité et fiabilité)⁸ est en cours de validation au niveau européen, et vient compléter la norme ISO-82304-1 (logiciel de santé – Partie 1 : Exigences générales pour la sécurité des produits).

La norme ISO-82304-2 s'appuie en partie sur la norme britannique PAS-277 et elle vise spécifiquement les Applications mobiles avec 81 questions. **Cette norme pourrait contribuer à uniformiser le niveau d'exigence de la qualité des Applications** au niveau international et limiter la prolifération de différents systèmes d'évaluation dans différents pays.

Par ailleurs, concernant le marquage CE des dispositifs médicaux, les Applications médicales sont soumises au nouveau règlement européen 2017/745⁹ relatif aux dispositifs médicaux, entré en application le 26 mai 2021 en remplacement des directives européennes jusqu'alors en vigueur.

⁸ https://www.iso.org/news/isofocus_141-6.html

⁹ <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32017R0745>

1.5.3. Les systèmes d'évaluation des professionnels de santé

Les professionnels de santé utilisent des systèmes d'évaluation s'appuyant sur les méthodologies de l'épidémiologie (ex. : essai contrôlé randomisé pour l'efficacité d'un traitement) ou sur des approches qualitatives utilisées en santé publique (17).

En 2013, Kumar (8) a exposé les différentes méthodes d'évaluation qui pourraient être utilisées en fonction du type d'Application proposée.

En 2019, l'Organisation mondiale de la santé (OMS) (18) a édité un guide pour piloter et évaluer les solutions numériques et contribuer à harmoniser les pratiques avec des tableaux de synthèses compilant les différentes approches méthodologiques en fonction des objectifs de l'évaluation clinique de ces Applications.

Pham (19), en 2016, démontre que l'essai contrôlé randomisé (ECR) est l'approche la plus utilisée pour évaluer la santé mobile. Elle démontre également que cette approche touche une fraction des produits utilisés par les consommateurs. Des méthodes alternatives devraient être proposées pour s'adapter dans le temps, et pour répondre à des « systèmes socio-techniques ». Ce que confirme Byambasuren (20), en 2018, avec une revue de revues systématiques sur l'efficacité des produits de santé mobile. Elle met en avant le champ restreint des essais contrôlés randomisés.

Mohr (21), en 2013, a proposé de développer des méthodes d'évaluation pour évaluer l'impact des produits orientés sur le changement de comportement. Il propose un cadre appelé CEEBIT (*Continuous Evaluation of Evolving Behavioral Intervention Technologies*) pour évaluer l'impact sur le comportement des utilisateurs au cours du temps.

Grundy (22), en 2016, expose les aspects multidimensionnels à prendre en compte dans le secteur, ce qui expose à une évaluation « multiguichet » et la création d'approches innovantes pour s'adapter au secteur.

Ologeanu-Taddei (23), en 2020, synthétise la complexité des méthodologies utilisables dans le domaine de la santé mobile. Elle évoque le champ de l'évaluation des risques qui est encore sous-estimé.

Ainsi, orienter l'évaluation uniquement sur l'intérêt clinique peut entraîner des risques sur d'autres dimensions. C'est ce qui s'est produit avec le répertoire du NHS (*NHS Apps Library*) qui a répertorié en ligne des Applications dont les deux tiers transmettaient des données personnelles non cryptées sur le web (13, 24, 25). La mise en avant d'Applications par le NHS présentant ces failles a entraîné la fermeture du répertoire et une remise à plat du système d'évaluation pour être indexé.

Wierda (26), en 2020, a publié une analyse de risque sur les données personnelles selon les différentes phases de collecte et de transmission de ces données. Cette dimension de protection et de sécurité des données personnelles est citée régulièrement comme un enjeu de confiance pour l'utilisation des Applications et des objets connectés (7, 27).

1.5.4. Les systèmes d'évaluation des usagers

Les usagers ou les patients non-experts du sujet utilisent des ressources externes pour évaluer les Applications mobiles ou les objets connectés (28). Le plus souvent, les usagers s'appuient sur les avis de professionnels ou de leur entourage, ou les commentaires des utilisateurs sur le site web du fabricant.

À noter que ces dernières évaluations peuvent être sujettes à caution. Pustozarov (29), en 2016, retrouve une moyenne d'étoiles de 4 sur 5, quel que soit le type d'Applications évaluées (déviation

standard de 1,65). De plus, 90,7 % des évaluations portent sur seulement 4,1 % des Applications disponibles ce qui limite l'intérêt de ce type d'évaluation qu'aux Applications les plus populaires. Enfin, il convient de se méfier des faux avis qui sont parfois achetés par les fabricants et qui sont à différencier des avis certifiés. La direction générale de la concurrence, de la consommation et de la répression des fraudes (DGCCRF) a émis une fiche pratique sur le sujet¹⁰ avec une norme ISO.

Singh (30) a démontré la faible corrélation entre l'évaluation des utilisateurs et l'utilité clinique (rhô de Spearman à 0,21 ; $p = 0,02$) ou l'échelle de l'utilisation clinique (rhô de Spearman à 0,11 ; $p = 0,2$).

Quelques documents d'aide au choix des Applications ou des précautions à prendre sont publiés pour le grand public^{11,12,13,14}.

1.5.5. Les systèmes d'évaluation des évaluateurs

Les acteurs de l'évaluation sont nombreux. Les premiers acteurs du secteur privés commercialisent des systèmes d'évaluation propriétaires (le plus souvent sous forme de label) dont les critères d'évaluation ne sont pas publics (appelée « boîte noire » d'évaluation), et constituent ainsi leur secret industriel¹⁵.

Pour les acteurs institutionnels internationaux, impliqués dans la mise en place de systèmes d'évaluation des Applications, la sécurité des données des utilisateurs est une priorité majeure.

Les sociétés savantes et les associations de patients essayent de recommander des solutions spécifiques pour leurs membres et le plus souvent sur des domaines spécialisés en ayant le meilleur rapport bénéfice/risque pour des profils de patients spécifiques.

Actuellement, ces approches sont confrontées à deux écueils, soit le système d'évaluation est trop exigeant et le répertoire n'est pas utile et peu rempli, soit le système d'évaluation est trop généraliste et il n'est pas suffisamment précis pour être fiable et donner confiance aux utilisateurs.

1.5.6. Les enjeux de l'évaluation

Finalement, les points de vue des acteurs, des évaluateurs ou des utilisateurs couvrent de nombreux domaines d'évaluation et de nombreuses manières d'évaluer, ce qui rend complexes les systèmes d'évaluation qui tentent de synthétiser l'ensemble des approches de ce secteur.

Les enjeux de l'évaluation ciblent le plus souvent les domaines suivants :

- enjeu technique (ex. : conception fiable sur le plan logiciel et/ou matériel) ;
- enjeu de cybersécurité (ex. : processus d'évaluation du risque cyber¹⁶, gestion des failles) ;
- enjeu de protection des données personnelles (ex. : respect du règlement UE 2016/679 du Parlement européen et du Conseil du 27 avril 2016 sur la protection des données [règlement général sur la protection des données européen – RGPD] et du code européen « *Code of Conduct on privacy for mHealth apps* »^{17,18}) ;

¹⁰ <https://www.economie.gouv.fr/dgccrf/Publications/Vie-pratique/Fiches-pratiques/faux-avis-consommateurs-sur-internet>

¹¹ <https://www.economie.gouv.fr/dgccrf/Publications/Vie-pratique/Fiches-pratiques/objets-connectes>

¹² https://gskpro.com/content/dam/global/hcpportal/fr_FR/AiresThrapeutiques/pneumoneu/pdf/info-patients.pdf

¹³ https://www.mentalhealthcommission.ca/sites/default/files/2018-01/eMH_app_fr.pdf

¹⁴ <https://www.healthnavigator.org.nz/apps/h/how-to-choose-a-health-app-a-guide-for-clinicians/>

¹⁵ <https://dumas.ccsd.cnrs.fr/dumas-02971517/document> (pages 83-84)

¹⁶ <https://a51.nl/sites/default/files/pdf/Pacemaker%20Ecosystem%20Evaluation.pdf>

¹⁷ <https://ec.europa.eu/digital-single-market/en/news/code-conduct-privacy-mhealth-apps-has-been-finalised>

¹⁸ http://ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=16125

- enjeu de qualité médicale (ex. : fiabilité du contenu médical) ;
- enjeu de performance de l'interface avec l'utilisateur¹⁹ (ex. : littératie, navigation, ergonomie, tests de scénarios cliniques) ;
- enjeu économique (ex. : viabilité du modèle économique).

1.6. Évolution quantitative de la publication scientifique

Parallèlement, la recherche évolue de manière exponentielle pour étudier l'utilité de la santé mobile. Ali (31) a étudié l'évolution des publications entre 1995 et 2015. Il démontre la croissance des publications, et il regroupe en cinq domaines les secteurs étudiés :

- promotion de la santé et prévention ;
- diagnostic ;
- traitement ;
- suivi (*monitoring*) ;
- appui de service de santé.

Les articles concernant les Applications de **suivi de patients** (*monitoring*) et les Applications de **diagnostics** sont ceux qui montrent la plus forte croissance depuis 2013. Ces domaines sont évalués avec des études cliniques compilées dans des revues systématiques. À noter que Fiordelli (32), en 2013, avait déjà prédit l'évolution de la recherche dans le sens d'une plus grande mesure de l'impact des Applications sur la santé.

De plus, la recherche documentaire que nous avons effectuée sur les recommandations et les revues systématiques a démontré l'accélération des publications cliniques sur ce secteur. Cette **accélération s'accroît ces cinq dernières années** ce qui montre que le secteur cherche à développer l'évaluation clinique en recherchant des niveaux de preuve. Le niveau de qualité reste toutefois hétérogène (33, 34).

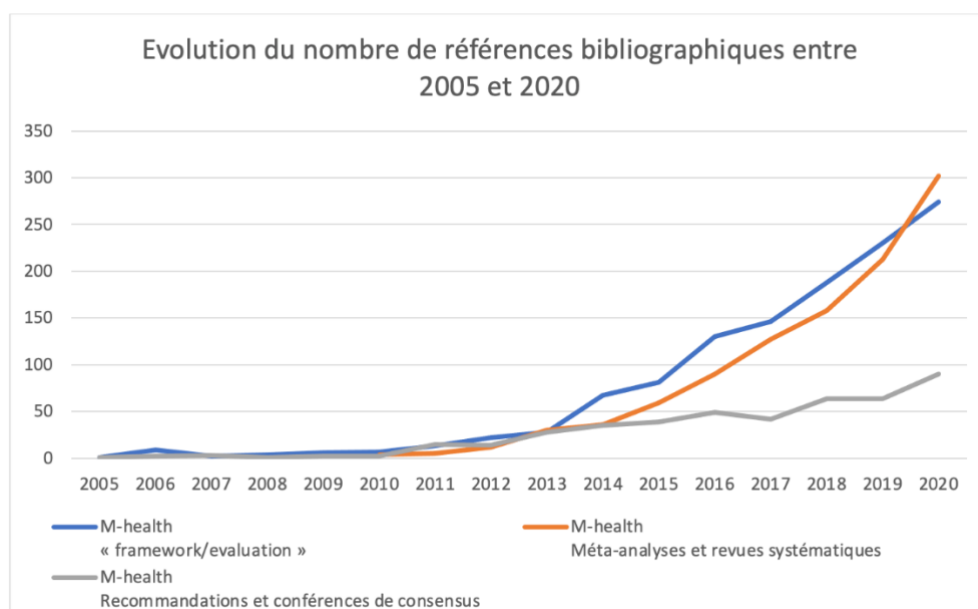


Figure 1. Évolution du nombre de publications dans le champ de la santé mobile de 2005 à 2020

¹⁹ <https://www.oracle.com/technetwork/topics/ux/applications/user-profile-template-1884987.pdf>

1.7. Niveaux de preuve et évaluation de l'utilité clinique

L'institut IQVIA (9) a synthétisé entre 2007 et 2017 les tendances du type de publications (234 essais contrôlés randomisés et 20 méta-analyses) et a regroupé les secteurs où des niveaux de preuve émergeaient.

Ainsi, pour l'institut IQVIA, les domaines du diabète, de la dépression et de l'anxiété sont des secteurs où des Applications devraient être citées dans les recommandations de pratique clinique.

IQVIA cite 24 domaines dont la qualité des publications rend éligible l'adoption des Applications :

- gestion du poids/manger sainement ;
- asthme ;
- maladie infectieuse & parasitaire ;
- bronchopneumopathie chronique obstructive (BPCO) ;
- perte d'audition & acouphènes ;
- insuffisance cardiaque ;
- accident vasculaire cérébral ;
- maladie d'Alzheimer ;
- gestion de la médication ;
- abus d'alcool et autres substances ;
- insomnie et qualité de sommeil ;
- cancer ;
- arrêt du tabac ;
- maladie de Parkinson ;
- hypertension ;
- réhabilitation cardiaque ;
- gestion du stress ;
- consommation modérée d'alcool ;
- trouble de stress post-traumatique ;
- problèmes de la sphère génito-urinaire ;
- prévention du diabète ;
- arthrite ;
- maladie rénale ;
- réhabilitation pulmonaire.

D'un autre côté, selon IQVIA en 2017, d'autres secteurs n'ont pas réussi à faire leurs preuves ou nécessitent d'autres études :

- exercice physique ;
- gestion de la douleur ;
- problèmes dermatologiques ;
- trouble bipolaire/schizophrénie ;
- sclérose en plaques ;
- autisme.

IQVIA évoque une chronologie progressive du type de publications : de l'étude observationnelle au premier essai clinique ou des essais cliniques aux méta-analyses avec un niveau de qualité qui s'améliore au cours du temps. Il s'agit d'une temporalité naturelle, car il existe un délai incompressible entre la mise en place d'une étude et sa publication et ce secteur n'a qu'une petite dizaine d'années de réel développement.

Concernant l'utilisation des Applications dans la recherche clinique, des sites dédiés sont en place comme le *Clinical Trials Transformation Initiative* – CTTI²⁰. Il propose des recommandations pour l'utilisation des Applications dans le cadre des essais cliniques^{21,22}. Fin 2020, la base de données de la CTTI recense 438 essais cliniques²³.

L'interrogation du registre américain des essais cliniques²⁴ recense en 2021 plus de mille essais cliniques interventionnels avec le mot-clé « mhealth ».

Enfin, des congrès internationaux entièrement dédiés à la santé mobile et leur impact dans les essais cliniques sont organisés²⁵.

Pour conclure

L'intérêt généré par les Applications entraîne une demande d'évaluation dans le double objectif d'améliorer la qualité des Applications et de renforcer la confiance des utilisateurs.

L'évaluation en santé mobile est multidimensionnelle. Elle s'appuie notamment sur la normalisation de la fabrication, le respect et la protection des données personnelles, la sécurité informatique et des télécommunications, la gestion et la prévention des risques induits par l'utilisation des Applications, l'impact sur la santé publique, l'adoption par les utilisateurs et le modèle économique pour maintenir et développer l'Application.

En conséquence de ce qui précède, après une phase de foisonnement dans la création et la production d'Applications couvrant le bien-être, la forme ou la santé, une phase d'accélération des publications scientifiques et d'études cliniques s'est enclenchée depuis 2013, avec une amélioration qualitative des publications. Cependant, la multitude des systèmes d'évaluation rend complexe la mise en place d'un système agrégé.

²⁰ <https://www.ctti-clinicaltrials.org/news/ctti-update-advancing-use-mobile-health-technology-transform-clinical-trials>

²¹ https://www.ctti-clinicaltrials.org/sites/www.ctti-clinicaltrials.org/files/ctti_recommendations_-_mct_engaging_patients_and_sites_final.pdf

²² <https://www.ctti-clinicaltrials.org/sites/www.ctti-clinicaltrials.org/files/mobile-technologies-executive-summary.pdf>

²³ <https://feasibility-studies.ctti-clinicaltrials.org>

²⁴ https://clinicaltrials.gov/ct2/results?term=mhealth&age_v=&gndr=&type=Intr&rslt=&Search=Apply

²⁵ <https://impacct-mhealth.com/about/what-to-expect-from-a-digital-event/>

2. Risques potentiels des fonctions de base dans le cadre de la santé mobile

L'utilisation d'Applications mobiles ou d'objets connectés expose à différents types de risque en fonction de l'usage et des fonctionnalités revendiquées, du profil de l'utilisateur et du contexte où elles sont utilisées.

2.1. La place du risque dans l'évaluation de la santé mobile

La notion de risque et de son évaluation peut être abordée sous forme de matrice qui permet de mettre en relation deux ou plusieurs dimensions après l'identification de menace ou de danger. Par exemple, avec une analyse « fréquence/gravité » suite à une analyse de fréquence des incidents/accidents et des conséquences.

Des analyses quantitatives (probabiliste) ou qualitatives existent également (35).

Lewis (36), en 2014, a recensé les différents niveaux de risque dans le domaine de la santé mobile. Pour lui, le risque est surtout lié à la **complexité du produit**, car il augmente le risque. Il définit donc deux dimensions à sa matrice de risque :

- de l'Application la plus simple (risque bas) à la plus complexe (risque élevé) ;
- du faible risque de préjudice (risque bas) au risque élevé de préjudice (risque élevé).

Les types de risques sont recensés et hiérarchisés dans cet article avec les scénarios pouvant provoquer ces risques.

Pour Bradway (37), en 2017, les dimensions de risque proposées dans le cadre de la santé connectée sont également dépendantes du **profil de l'utilisateur**. Il recense le :

- risque lié au côté spécifique de l'intervention (référence & guide, communiquer et coordonner, diagnostiquer et traiter, suivre et alerter) ;
- risque lié au niveau de fragilité de la personne (personne en bonne santé, patient atteint de problème chronique et modéré, patient atteint de problème sévère, patient fragile et à risque).

Cette approche cible l'utilisation individuelle des Applications et objets connectés dans sa dimension humaine. Les Applications sont développées pour répondre aux besoins spécifiques des utilisateurs. Cela peut entraîner des risques dans le **contexte d'utilisation** (38).

En Allemagne, le rapport IGES (39), en 2016, propose une matrice de risque pour les produits de santé pouvant être des dispositifs médicaux avec deux dimensions :

- en relation avec l'utilisateur ;
- en fonction du niveau d'autonomie de fonctionnement du produit (de l'information générale au remplacement du professionnel de santé par des algorithmes).

En Australie, la TGA (*Therapeutic Goods Administration*) (40), en 2020, a effectué une revue de littérature sur le risque des Applications. Ils ont été recensés de manière générale et aussi pour des problèmes de santé spécifiques (vérification des symptômes, suivi diabétique, mélanome/analyse de la peau, asthme, mesures cardio-vasculaires, calcul de dose). Ce rapport souligne la **faiblesse du niveau d'évaluation du risque et l'impact éventuel sur les prises de décisions**.

De son côté, la HAS (1), en 2016, a défini une matrice de risque sur deux dimensions pour les Applications qui ne sont pas des dispositifs médicaux :

- en relation avec l'usage principal du produit (de l'information à l'analyse de données ciblées) ;
- en relation avec l'utilisateur principal (du grand public au professionnel de santé).

Ce choix s'appuie sur le risque de mésusage du produit ou du détournement d'usage du produit surtout pour des Applications qui ne sont pas des dispositifs médicaux.

Le risque est élevé si les professionnels de santé prennent des décisions sur de mauvaises informations, car cela impactera de nombreux patients.

Pour les **Applications médicales qui sont des dispositifs médicaux, la classe de risque est évaluée et les exigences adéquates sont demandées**²⁶.

Pour conclure

L'évaluation devrait être modulé par le risque lié à l'utilisation de l'Application (voir le rapport bénéfice/risque de l'utilisation de l'Application).

Ce niveau de risque peut être appréhendé en considérant les trois dimensions suivantes :

- l'**usage proposé par l'Application** qui détermine le danger (intrinsèque) en cas de dommage lié à un fonctionnement normal ou de dysfonctionnements ;
- le **contexte d'utilisation de l'Application** qui détermine la capacité à agir sur le risque et peut, selon les cas, permettre d'éviter sa survenue (détection) ou en réduire les conséquences (atténuation, récupération) ;
- la **population cible** utilisant l'Application qui détermine les conséquences supportées et peut-être un facteur de risque éventuel (mésusage, mauvaise interprétation ou compréhension).

2.2. Les fonctions de base les plus fréquentes

Les Applications mobiles peuvent avoir un à plusieurs usages revendiqués par le fabricant. Elles vont de la diffusion d'informations générales aux recommandations spécifiquement élaborées en fonction des données fournies par l'utilisateur. Il existe des risques plus ou moins spécifiques des différents types de « fonction d'usage » des Applications. À ce risque, s'ajoute l'exposition des données personnelles à des tiers non autorisés ou par cyberattaque comme ce fut le cas pour le store du NHS en 2015 (41).

Kearney (42) a proposé une modélisation des usages en fonction des domaines de la santé et du type d'utilisateur. Il définit deux axes :

- un axe de **fonction d'usage** comprenant : information/communication, évaluation, intervention, *monitoring*, coordination/travail en équipe ;
- un axe **santé** comprenant : le bien-être, la prévention, le diagnostic, la thérapie, le suivi (contrôle).

L'intersection de ces deux axes permet de visualiser l'objectif de l'Application pour le grand public, les patients ou les professionnels de santé.

²⁶ <https://ansm.sante.fr/page/mise-sur-le-marche-des-dispositifs-medicaux-et-des-dispositifs-medicaux-de-diagnostic-in-vitro>

Schématiquement, les **fonctions de base** le plus souvent retrouvées²⁷ dans le domaine de la santé mobile sont les suivantes :

- informer ;
- instruire ;
- enregistrer des données ;
- calculer/analyser ;
- alerter/rappeler ;
- afficher les données ;
- guider ;
- communiquer.

Ces fonctions de base sont utilisées individuellement ou associées pour atteindre les objectifs de l'Application. Certaines produisent, diffusent et utilisent des données personnelles avec un risque d'exposition à des tiers non autorisés (30). Ce risque est complexe à évaluer (43).

L'ordre de présentation ci-dessous incrémente arbitrairement des fonctions de plus en plus avancées et vise tous les types d'utilisateurs (usagers, patients, aidants, professionnels de santé).

2.2.1. Informer

Il s'agit de la **diffusion et de l'accès à l'information en santé de manière générale et identique, quel que soit l'utilisateur**. Les informations concernent la promotion de l'hygiène de vie et de la santé, l'éducation à la santé, l'accessibilité aux informations selon des formats adaptés à l'utilisateur, aux patients ou aux professionnels de la santé. Cette information est parfois « poussée/servie » de manière automatique vers l'utilisateur, lorsque ce dernier a sélectionné des sujets spécifiques sur lesquels il souhaite être informé. Il peut également souhaiter être alerté ou « notifié » de nouvelles publications ou données sur des thématiques sélectionnées ou graduellement adressées (44).

2.2.2. Instruire

Il s'agit d'**éducation du patient ou de fournir des instructions plus spécifiques pour l'utilisateur ou le professionnel de santé**. L'utilisation de questionnaires, de jeux (conceptualisée par le terme *gamification* (45)), de situations réelles, d'échanges au travers d'interactions avec un environnement virtuel ou un groupe de personnes référentes est le moyen le plus souvent utilisé pour adapter les informations à fournir en fonction du niveau de connaissance ou de questionnement de l'utilisateur. Pour les professionnels de santé, il s'agit parfois de tutorat pour réaliser une procédure spécifique.

2.2.3. Enregistrer des données

Il s'agit de **suivre ou de capturer des données de santé** (ex. : suivi de photos de l'évolution de grain de beauté en mélanome, états du niveau d'anxiété au cours de la journée, et toute collecte de « biomarqueurs numériques » pour contribuer à prédire la survenue éventuelle de problèmes de santé ou de pathologies spécifiques, etc.).

Il peut s'agir également de **mesures renseignées par l'utilisateur** (ex. : calendrier mictionnel pour mieux gérer l'espacement entre les mictions lors de troubles mictionnels, remplissage de questionnaires réguliers, atteinte d'objectifs planifiés, etc.). Les domaines les plus connus concernent

²⁷ <https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/patient-adoption-of-mhealth.pdf> (pages 7-8)

l'activité physique, le poids et le régime alimentaire, ou des paramètres liés à la qualité du sommeil ou de l'humeur.

Dans le cadre de dispositifs médicaux, les paramètres mesurés permettent de suivre, en direct ou en différé, l'état de santé de l'utilisateur à distance. Les mesures les plus connues sont des paramètres cardio-vasculaires (ex. : pression artérielle, pulsation cardiaque, saturation en oxygène, débitmètre de pointe, etc.), la température corporelle lors de suivi post-opératoire ou d'infections ou le suivi d'affections chroniques (ex. : l'équilibre glycémique pour de patients diabétiques, etc.).

Les mesures peuvent être parfois **réalisées sous forme de questionnaires**. Par exemple, les PROMs (*patient-reported outcomes measures*) pour évaluer les résultats de soins ou les PREMs (*patient-reported experience measures*) pour évaluer le vécu ou la satisfaction des soins réalisés.

2.2.4. Calculer/analyser

Il s'agit de **réaliser des calculs à partir de données du patient** (ex. : opération mathématique pour le calcul de doses en fonction du poids, scores, etc.) ou des analyses de données (avec interprétation humaine ou *via* des algorithmes). La qualité du calcul est liée à la qualité de la collecte de données (46).

Concernant les algorithmes, au niveau européen, des recommandations ont été publiées en 2019 (47). En 2021, la Commission européenne travaille sur la mise en place d'une régulation de l'intelligence artificielle (IA)²⁸. Une proposition d'évaluation par niveaux de risques est attribuée aux différents systèmes d'IA. Pour le niveau de risque élevé, les systèmes d'IA pourraient être conformes aux obligations strictes suivantes :

- systèmes adéquats d'évaluation et d'atténuation des risques ;
- qualité élevée des ensembles de données alimentant le système afin de réduire au minimum les risques et les résultats ayant un effet discriminatoire ;
- enregistrement des activités afin de garantir la traçabilité des résultats ;
- documentation détaillée fournissant toutes les informations nécessaires sur le système et sur sa finalité pour permettre aux autorités d'évaluer sa conformité ;
- informations claires et adéquates à l'intention de l'utilisateur ; contrôle humain approprié pour réduire au minimum les risques ; niveau élevé de robustesse, de sécurité et d'exactitude.

Les systèmes d'IA autonomes pourraient être enregistrés dans une base de données de l'Union européenne. Pour être mise sur le marché, une déclaration de conformité pourrait être demandée et le système d'IA porterait le marquage CE. Ce domaine spécifique devra sans doute être évalué spécifiquement.

Il existe de nombreuses publications qui ont répertorié les risques autour du calcul et des analyses de données (24, 46, 48-50).

2.2.5. Alerter/rappeler

Il s'agit de **fournir des alertes ou des rappels** (cités le plus souvent sous le nom de *reminder*) pour l'utilisateur ou pour un professionnel de santé référent. Initialement réalisés *via* des SMS adressés à l'utilisateur pour la prise de médicaments, des rappels de vaccination ou de dépistages, les rappels et alertes pour les utilisateurs ont évolué vers des Applications spécialement dédiées utilisant différents formats de relance (notification, affichage et message dédié, etc.) et de suivi pour l'activité physique,

²⁸ https://eur-lex.europa.eu/procedure/FI/2021_106

la santé mentale, l'encouragement pour le sevrage tabagique, la géolocalisation pour le risque allergique (pollen, vaccination, etc.), etc.

Pour les professionnels de santé, les alertes sont paramétrées pour le suivi de patients, la prévention et surveillance de risque spécifique (ex. : télésurveillance médicale).

Dans le cadre de dispositifs médicaux, les paramètres de surveillance mesurés permettent d'alerter à distance le patient et le professionnel référent (fonction faisant partie de la télésurveillance, appelée parfois *remote patient monitoring* [RPM]). Un paramétrage de niveau d'alerte est réalisé pour situer les seuils et types d'alertes et les personnes concernées par la réception de l'alerte (51).

2.2.6. Afficher des données

Il s'agit d'**afficher par un ou différents modes de présentations les résultats des données collectées ou retravaillées par l'Application**. Les différents modes d'affichage visent à renforcer l'information renvoyée à l'utilisateur pour rendre plus performant le message adressé. Ces données auront pu être enregistrées par l'utilisateur, captées par des capteurs intégrés au smartphone, collectées par un objet connecté/dispositif médical ou produites par un algorithme informatique.

2.2.7. Guider

Il s'agit de **fournir des orientations** basées sur les informations collectées auprès de l'utilisateur, ou éventuellement recommander une action spécifique, une consultation avec un praticien particulier ou un traitement lorsque le système s'adresse à l'usager du système de santé.

Lorsque les outils sont développés pour les professionnels de la santé, ils sont le plus souvent orientés comme une aide à la prise de décision.

2.2.8. Communiquer

Il s'agit de **développer la communication entre les professionnels de santé et les usagers du système de santé**. Les outils peuvent être synchrones (ex. : communication en direct lors d'une consultation de télésoins) ou asynchrones (communication par messages distincts).

2.2.9. La variation du risque selon les fonctions de base des Applications mobiles

Si les fonctions de base d'une Application sont peu performantes, elles vont entraîner des risques. Ces risques se caractérisent par leurs probabilités de survenues et par leurs conséquences (selon le modèle fréquence/gravité).

Concernant le contenu médical ou de santé, les conséquences les plus répandues sont un **délai de prise en charge inadapté** en étant faussement rassuré ou mal informé ou, en sens inverse, une surmédicalisation ou une anxiété inutile de l'utilisateur. Le niveau de qualité de l'information et des données de santé utilisées influe sur l'optimisation de la prise en charge (définie comme étant l'application du bon soin, au bon moment, pour le bon patient).

Le tableau 1 reprend les fonctions de base des Applications dans le secteur de la santé avec des exemples de types de risques et leurs conséquences sur le contenu de santé. D'autres exemples de risques connexes sont reportés pour contextualiser et illustrer plus largement les conséquences potentielles.

Tableau 1 : Fonctions de base d'une Application et exemples de risques

Fonctions	Exemples de risques sur le contenu médical	Autres exemples de risques
Informier	Risques de désinformation. Dépend du niveau de qualité de l'information : source, qualifications des auteurs et liens d'intérêts, niveau de preuve, actualisation.	Risques de mauvaises interprétations. Dépend du niveau de compréhension et de pertinence de la sélection par l'utilisateur des canaux d'informations utilisés.
Instruire	Risques de mauvaise pratique. Dépend du niveau de qualité de l'information, mais aussi des méthodes pédagogiques ou éducatives utilisées au sein de l'Application.	Risques de biais de sélection. Dépend du niveau de performance de l'évaluation des besoins de l'utilisateur et de la pertinence des réponses proposées.
Enregistrer des données	Risques d'enregistrer des données de mauvaise qualité que ce soit auprès de l'utilisateur, ou que ce soit <i>via</i> un équipement externe. Dépend de la fiabilité des données renseignées, collectées, enregistrées et archivées.	Risques liés à l'accès ou aux modifications des informations et des données de santé collectées auprès de l'utilisateur (cybersécurité, protection des données personnelles, dysfonctionnements techniques).
Calculer/Analyser	Risques de générer des erreurs de calcul ou d'analyse. Dépend de la fiabilité des données enregistrées pour effectuer les calculs, mais aussi de la fiabilité des méthodes de calcul réalisées et du niveau de performance de l'interprétation.	Si l'interprétation est réalisée <i>via</i> un algorithme, les risques sont liés au maintien de la performance des algorithmes dans différentes situations cliniques et d'environnement d'utilisation. L'évaluation des systèmes dénommés globalement comme « intelligence artificielle (IA) » est à adapter en fonction du type d'IA utilisée.
Alerter/ Rappeler	Risques de ne pas générer d'alertes adaptées (absence, retard, trop d'alertes, etc.). Dépend des paramètres renseignés pour déclencher les alertes et à l'interprétation par les utilisateurs des notifications.	Risques liés à la collecte d'informations et à la fixation des seuils d'alerte ou du type d'alerte. Ces informations peuvent être renseignées ou modifiées par un tiers (risque de sécurité) ou par l'utilisateur (risque de mésusage) et ne pas être adaptées à la situation de l'utilisateur.
Afficher des données	Risques de générer un écart entre les données réelles et leur affichage. Dépend de l'importance de l'écart entre les données collectées brutes et la manière dont elles ont été transformées pour être affichées (lissage, granularité, calibration par rapport à des standards/cohortes, qualité de l'image, transformation en pattern de mouvements, etc.). Cet écart peut ne plus refléter la réalité de la mesure étudiée.	Risques liés à la compréhension. Dépend du niveau de compréhension de l'affichage des données (par l'utilisateur, par l'aidant ou par le professionnel) et des conséquences sur les prises de décisions qui en découlent. Risque lié à la « fonction de partage » des données de santé pour renforcer ou encourager l'utilisateur (protection des données personnelles).
Guider	Risques liés à la pertinence de l'orientation proposée. L'orientation proposée peut influencer la prise de décision et entraîner une prise en charge inadaptée en augmentant les coûts humains (ex. : iatrogénie, examen abusif, stress) ou économiques (ex. : dépenses non nécessaires).	Si l'interprétation est réalisée <i>via</i> un algorithme, les risques sont liés à la qualité de la justification proposée par le système (sur quels éléments reposent ces propositions, orientations, décisions ? Qui les garantit ?).
Communiquer	Risques liés à la pertinence des réponses apportées.	Risques liés aux délais de réponses d'origine organisationnelle ou de compétence technique.

Pour conclure

Une Application propose le plus souvent une ou plusieurs fonctions de base (informer, calculer, guider, communiquer, etc.). Ces fonctions peuvent être retrouvées à différents stades du parcours de l'utilisateur ou du patient (bien-être et prévention, suivi et monitoring, traitement, etc.).

La fiabilité des fonctions de base d'une Application détermine une partie de la qualité de l'information et des données de santé collectées. L'utilisateur devrait être averti des risques liés à l'utilisation de l'Application (description des conséquences potentielles) *via* des signalétiques adaptées ou des moyens de communication ciblés.

3. Les méthodes d'évaluation de la santé mobile au niveau international

Selon les pays, le ou les systèmes d'évaluation de la santé mobile s'appuient sur une ou plusieurs des approches répertoriées ci-dessous.

3.1. Standard d'édition scientifique pour les publications concernant la santé mobile

Le développement innovant des Applications de santé mobile a entraîné un impact sur le format de publications de la recherche sur le secteur. Le *Journal of Medical Internet Research* (JMIR) a proposé une grille d'évaluation des Applications de manière similaire à la soumission d'un article pour un journal (avec relecture par les pairs). Les standards d'édition ou les standards pour publier des revues de synthèses sur les Applications ont été mis en place. Cette structuration du format de publications s'est effectuée durant ces dix dernières années.

À noter qu'un consensus international pour généraliser l'utilisation du terme « App » a été publié par Lewis (52) en 2014 pour les articles scientifiques publiés en anglais abordant les Applications. L'acronyme MMA (*mobile medical application*) est souvent utilisé pour parler des Applications qui sont des dispositifs médicaux.

3.1.1. Le formulaire de soumission d'Application du JMIR

Lors de l'élaboration du référentiel de 2016 (1), la HAS s'est appuyée sur le travail de Riezebos (53). Il avait réalisé une synthèse de la littérature pour compiler les critères de qualité qui permettraient d'évaluer les Applications dans le secteur de la santé mobile.

Son travail a permis de créer une **grille de critères pour la soumission d'Application originale auprès de la revue JMIR** (2013). Ce formulaire, toujours disponible en ligne²⁹, permet d'effectuer une relecture par ses pairs au travers de plusieurs domaines (niveau de preuve, sécurité, etc.) selon la même procédure que pour la relecture d'un article.

3.1.2. Standard CONSORT-EHEALTH et mERA

Eysenbach (54), en 2011, a développé une check-list lors de la soumission pour publication d'un essai portant sur une Application ou d'autres types similaires d'interventions dans le secteur de la e-santé. La CONSORT-EHEALTH est référencée dans le réseau EQUATOR (*Enhancing the QUALity and Transparency Of health Research*)³⁰ qui regroupe les recommandations aux auteurs pour les chercheurs et les éditeurs scientifiques. C'est actuellement le seul standard disponible concernant la publication d'essai sur les Applications (au 21/11/2019).

À noter qu'Agarwal (55), en 2016, a développé une check-list de seulement 16 items pour permettre d'alléger la procédure de soumission pour certaines Applications (notamment issue de pays en voie de développement). Le projet a été soutenu par l'OMS. La check-list s'appelle mERA (*mHealth Evidence Reporting and Assessment*).

²⁹ tinyurl.com/appsform

³⁰ <https://www.equator-network.org/reporting-guidelines/consort-ehealth-improving-and-standardizing-evaluation-reports-of-web-based-and-mobile-health-interventions/>

3.1.3. Méthode pour réaliser une revue de littérature sur des Applications spécifiques

Plusieurs centaines d'essais contrôlés ont été publiés dans le domaine de la santé mobile. Boudreaux (56), en 2014, a proposé une stratégie pour effectuer une recherche documentaire spécifique pour évaluer le niveau de la littérature sur l'utilité d'Applications dans des domaines de santé spécifiques, mais aussi en explorant différentes autres sources d'information.

Il propose sept étapes :

1. **Effectuer une revue de littérature.** Rechercher les preuves scientifiques à partir de publications étudiant l'utilisation d'Application dans le domaine exploré ou des essais contrôlés randomisés robustes.
2. **Rechercher dans les répertoires et bases de données.** Ces répertoires regroupent et évaluent les Applications. Cette analyse permet d'identifier leurs forces et faiblesses.
3. **Rechercher dans les magasins en ligne d'Applications.** Même si la recherche est un défi dans ces magasins. Il est nécessaire d'affiner la sélection et d'utiliser les bons mots-clés en rapport avec le domaine exploré.
4. **Effectuer une revue de la description, de la notation et des commentaires des utilisateurs, des Applications.** Les publications de ces avis et commentaires permettent d'identifier des preuves concernant la facilité d'utilisation, les fonctionnalités, et l'efficacité ce qui permet de réduire le nombre d'Applications pertinentes dans le domaine exploré.
5. **Conduire un sondage sur les réseaux sociaux comprenant des professionnels et si possible des réseaux de patients.** Les réseaux sociaux pourraient faire apparaître les nouvelles Applications tendance, leur degré de sympathie par les utilisateurs, et d'autres données similaires.
6. **Effectuer un test pilote sur les Applications.** Le test mené par le professionnel de santé adapté devrait évaluer les fonctionnalités, la précision du contenu et la facilité d'utilisation.
7. **Solliciter les feed-back des patients.** Les patients devraient pouvoir donner un regard approfondi après avoir utilisé l'Application recommandée par leur prescripteur.

Cette stratégie pragmatique permettrait de faire un état des lieux de l'utilisation des Applications pour des domaines de santé précis ou des problèmes spécifiques identifiés, par exemple, par les sociétés savantes.

3.2. Les différentes approches/outils d'évaluation de la santé mobile

De nombreux auteurs ou organisations ont réalisé des revues de littérature pour agréger les approches évaluatives (hétérogènes) proposées au niveau international dans le domaine de la santé mobile. Certains auteurs citent le référentiel HAS de 2016 dans leur synthèse (57-64).

Parmi ces productions, la Suisse via l'organisme fédéral eHealth Suisse³¹ a publié une revue de littérature internationale des méthodes d'évaluation de la santé mobile (15) en 2019.

Cet organisme coordonne la doctrine de la e-santé dans le pays, et a produit différents documents³² (recommandations, critères de qualité, raccordement des Applications, normes et standards techniques, etc.).

³¹ <https://www.e-health-suisse.ch/fr/page-daccueil.html>

³² <https://www.e-health-suisse.ch/fr/mise-en-oeuvre-communautes/activites-ehealth/mhealth.html>

Il propose de regrouper les résultats de cette revue de littérature selon trois niveaux, avec la structure suivante, en reprenant et adaptant la classification de CHARISMHA (65).

→ Évaluations reliées à des outils/instruments généraux

Ils sont définis par eHealth Suisse (15) comme « divers outils ou diverses mesures, par exemple un guide avec une liste de contrôle et/ou des répertoires éditorialisés d'Applications mobiles ».

Deux classes d'outils sont incluses dans cette catégorie conceptuelle d'évaluation :

- la **règlementation** regroupant les attestations de conformité et les agréments (ex. : le marquage CE pour les dispositifs médicaux) ;
- les **codes** regroupant une liste de critères ou de règles de bonnes conduites ou de qualité proposés sur une base volontaire d'utilisation.

→ Évaluation par des tiers

Ils sont définis par eHealth Suisse (15) comme une évaluation externe réalisée par « des experts d'un domaine de spécialisation médical précis ou du personnel technique, et non pas par les fabricants ou les utilisateurs mêmes. Les experts peuvent cependant accéder également à des données auto-déclarées par les fabricants ».

Six classes d'outils sont incluses dans cette catégorie conceptuelle d'évaluation :

- les **codes avec attribution d'un label** regroupant les mêmes types de codes précédents, mais évalués par de tiers externes. Le respect des critères donnant la possibilité d'attribuer un label (les procédures de contrôle pouvant être plus ou moins strictes) ;
- les **labels de qualité** attribués le plus souvent par des entreprises ou des organisations principalement du secteur privé (concernent le plus souvent des Applications hors du marquage CE ou non soumises à réglementation) ;
- les **répertoires éditorialisés** ou **plateforme d'évaluation** organisée le plus souvent sous le format de bases de données qui recensent une sélection prédéfinie d'Applications. En fonction du répertoire, les Applications affichées sont adossées à une ou des évaluations réalisées par des experts médicaux, des experts techniques ou des utilisateurs ;
- les **avis d'experts** et les **avis individuels** laissés sur les magasins en ligne (store) sous forme de commentaires ou d'étoiles. Les avis sont parfois publiés sous forme de rapports ;
- les **tests méthodiques** sont des évaluations ciblées avec des critères connus ou en partie non divulgués réalisés par des entreprises ou des organisations proposant ces services ;
- les **études scientifiques** sont le plus souvent des revues systématiques ou des méta-analyses d'études publiées sur un type spécifique d'Applications qui proposent des objectifs similaires.

→ Évaluation individuelle par des personnes intéressées

Ils sont définis par eHealth Suisse (15) comme une façon indépendante d'évaluer, pour épauler les utilisateurs ou les responsables lors de la sélection et de la décision pour ou contre l'utilisation d'une solution « mHealth ».

- Les **présentations standardisées des informations produits** permettent de présenter de manière homogène et transparente les informations pertinentes de l'Application. Ces

informations sont publiées aux endroits appropriés par le fabricant (information produit par le magasin en ligne, le site web du produit, etc.).

- Les **instructions pour évaluation pratique** permettent de mettre à disposition une liste de contrôle sous le format de catalogue de critères pour aider les utilisateurs à évaluer l'Application. Ce format d'évaluation permet de calculer un score ou un classement en fonction des résultats calculés.

Cette classification sera utilisée ci-après pour structurer l'état des lieux des différentes approches évaluatives internationales retrouvées dans l'analyse documentaire.

Pour conclure

L'enjeu technique de l'évaluation se situe autour de deux axes :

- Quels **domaines** d'évaluation explorés de manière efficace et pragmatique ?
- Quelles **stratégies d'évaluation** et d'**organisation** mettre en place ?

Deux choix de stratégie sont possible : les modèles de dispositifs inclusifs (moins exigeant) ou sélectifs (plus exigeant).

3.3. Scores, check-lists et autres échelles d'évaluation

Les scores et les check-lists ont été conçus pour faciliter l'évaluation des Applications. À destination des usagers ou des professionnels, ces outils permettent d'évaluer rapidement, indépendamment et facilement les Applications en donnant un score ou d'identifier les composantes attendues de l'Application avec une check-list.

Azad-Khaneghah (66), en 2020, a effectué une revue de littérature sur les échelles de mesure dans le domaine de la santé mobile. Il recense 23 échelles évaluant les qualités d'usage et de facilité d'utilisation des Applications et 25 échelles évaluant le niveau de qualité des Applications (87 publications retenues et analysées en tout). Azad-Khaneghah souligne l'ambiguïté du terme qualité qui est interprété de manière différente selon les auteurs ou évalué sur des domaines différents ou restreints.

3.3.1. Score de qualité du contenu médical

La HAS (67), en 2007, a publié une revue de littérature des méthodes d'évaluation des sites e-santé et de la qualité de l'information diffusée sur internet. Cette revue montrait, à l'époque, la variété des échelles qui évaluaient le contenu des sites web. Le même constat peut être produit concernant l'évaluation de la qualité des Applications mobiles dans le secteur de la santé.

Certaines échelles citées dans le rapport HAS de 2007 (67), et utilisées pour les sites web ont été adaptées pour évaluer les Applications. Van Singer (68), en 2015, cite ainsi les échelles de Abbott, Brief DISCERN (version en six questions du DISCERN notées de 1 à 5), Health On the Net (HON), ou le score de Silberg.

Butcher (69), en 2015, a créé un score sur 100 points et trois domaines (contenu, sources et niveaux de preuve) pour évaluer les Applications qui produisent du contenu médical. Ce score est surtout adapté aux Applications recensant des publications scientifiques.

3.3.1.1. Échelle ou score de Silberg

Une grande partie des échelles d'évaluation du contenu médical disponible en e-santé font suite à l'éditorial publié en 1997 par Silberg (70), rédacteur en chef du JAMA, sur l'assurance qualité de l'information médicale sur internet. Dans cet éditorial, Silberg cible quatre paramètres pour mesurer la qualité de l'information :

- **auteurs** (*authorship*) : citer les auteurs et les contributeurs. Citer leurs contributions et leurs qualifications ;
- **attribution** : citer les références et les sources de l'information pour tout le contenu. Cela devrait être listé clairement ainsi que toutes les informations pertinentes concernant le copyright ;
- **transparence** (*disclosure*) : le (ou les propriétaires) devrai(en)t être visiblement et complètement identifié(s) ; de même que les sponsors, les parrainages éventuels, les publicités, les publi-rédactionnels, les arrangements de financement commerciaux ou des soutiens similaires, et ce qui engendre de potentiels conflits d'intérêts. Cela inclut aussi les arrangements qui lient et renvoient vers d'autres sites avec une rétribution financière. Ce type de standard devant s'étendre au forum de discussion ;
- **actualisation** (*currency*) : dates de publication du contenu et de mise à jour.

Cet éditorial a ensuite été repris et l'échelle de Silberg construite avec des scores dédiés et adaptés.

Par exemple, Zhang (71), en 2017, a évalué le contenu de 14 Applications dans le cadre de la dépression du post-partum. Le score moyen de Silberg était de 3 (+/- 1,52) sur 9 points, ce qui a permis à Zhang de souligner le manque de transparence des sources d'information. Le détail du score de Silberg utilisé dans cette étude est cité ci-dessous.

- **auteurs** (3 points) : si les auteurs sont identifiés, si les affiliations sont identifiées, si les qualifications sont identifiées ;
- **attribution des sources** (2 points) : si les sources sont citées, si les références et les liens hypertextes sont cités ;
- **transparence** (2 points) : s'il existe un affichage visible du propriétaire de l'Application, et affichage visible des sponsors ;
- **date de publication** (2 points) : si l'Application a été modifiée depuis le mois dernier, si l'Application donne une date de dernière révision.

En fonction des utilisateurs de l'échelle de Silberg des adaptations dans la répartition des points sont proposées (68, 72).

Cette échelle mesure principalement le niveau de transparence de l'information et les liens d'intérêts potentiels. Elle ne permet pas de juger de la pertinence des informations sélectionnées et diffusées.

3.3.1.2. Brief DISCERN

Le Brief DISCERN (73) est la version courte de DISCERN³³ qui comprend 16 questions. Une version 5 étoiles existe également (67).

Le Brief DISCERN (68) explore un peu plus profondément le contenu (par rapport au score de Silberg) avec 6 questions (notées de 1 à 5) :

- les sources d'information utilisées pour compiler la publication sont-elles claires ?

³³ http://www.discern.org.uk/discern_instrument.php

- est-il clair quand les informations utilisées ou rapportées dans la publication ont été produites ?
- la publication décrit-elle le fonctionnement de chaque traitement ?
- décrit-elle les bénéfices de chaque traitement ?
- décrit-elle les risques de chaque traitement ?
- décrit-elle comment les choix de traitement affectent la qualité de vie globale ?

Les réponses à ces questions sont plus descriptives, mais ne permettent pas de juger complètement de la fiabilité des informations sélectionnées.

3.3.1.3. Synthèse

Les scores uniquement dédiés à la qualité du contenu médical sont le plus souvent orientés sur la publication des informations selon les **quatre questions de Silberg** (auteur, source, propriétaire et lien d'intérêt, actualisation). Ces scores, trop limités pour la plupart des Applications, sont complétés par des scores couvrant plusieurs domaines supplémentaires (tableau 2).

Les approches spécifiques évaluant le contenu d'une thématique médicale spécifique sont les plus précises pour évaluer le contenu. Ils sont construits de manière similaire à des référentiels d'audit (74) avec parfois des évaluations métrologiques de validité de contenu et de reproductibilité ou des essais contrôlés randomisés. Ils seront abordés plus loin dans ce document.

Tableau 2 : Exemples de scores évaluant uniquement le contenu médical

Nom	Pays	Catégories évaluées	Score	Commentaires
Score de Silberg	États-Unis	Auteurs, sources, transparence, date de publication.	Variable selon les reprises, mais le plus souvent sur 9 points.	Pondération différente en fonction des auteurs.
Brief DISCERN	DISCERN (Grande-Bretagne) Brief DISCERN (Suisse)	Sources, date de production des informations, fonctionnement du traitement, bénéfice du traitement, risque du traitement, choix du traitement et impact sur la qualité de vie.	6 questions notées de 1 à 5.	Production de base de données regroupant les Applications évaluées par ce score.
Score de Butcher	Canada	Contenu, transparence, niveau de preuve.	Pourcentage sur 8 questions.	Dédié aux Applications recensant des ressources d'information.

À noter que le NHS, en Grande-Bretagne, a proposé une liste de six principes et 16 items pour évaluer le processus de production d'informations dénommé *Information Standard Principles*³⁴. Ces principes sont conçus pour montrer que le concepteur a mis en place un processus bien défini pour produire et maintenir des informations de bonne qualité, fondées sur des preuves, sur la santé et les soins. Ces informations devraient être disponibles pour l'utilisateur.

Ces items déclaratifs permettent d'approfondir le processus qualité de production d'information. Une traduction non validée en français de ce standard est proposée en annexe 2.

³⁴ <https://www.england.nhs.uk/tis/about/the-info-standard/#information-production>

3.3.2. Score général de MARS

Le score de MARS (*Mobile Application Rating Scale*) est un score composite composé de plusieurs domaines d'évaluation. Il a été construit à partir de 349 items regroupés en six catégories suite à une revue de littérature (75) en 2015. Le score a été réduit à **23 items et coté de 1 à 5 dans quatre domaines objectifs et un domaine subjectif**. Le score a été porté dans plusieurs langues, dont l'italien (76), l'espagnol (77), l'allemand (78) ou l'arabe (79). Une traduction française avait été proposée dans le référentiel de la HAS en 2016, cette traduction n'a pas été validée en français à l'heure actuelle (annexe 3).

Les qualités métrologiques de ce score sont bonnes (80), et évaluées pour différents types d'Applications (81). Plusieurs dizaines de publications ont utilisé (82) ou évalué l'utilisation de ce score général qui émerge au-dessus des nombreux autres scores ou check-lists, mais reste limité sur le champ de son évaluation et notamment les niveaux de preuve (83).

3.3.2.1. MARS-G en Allemagne

En Allemagne, le score de MARS a été utilisé pour mettre en place un répertoire d'Applications appelé : « *Mobile Health App Database (MHAD)*³⁵ ». Ce répertoire a été créé en 2018, et comprend fin février 2021, 1 112 Applications balayant neuf catégories : pleine conscience (192 Applis), anxiété (104 Applis), dépression (39 Applis), soutien aux enfants et adolescents (13 Applis), cancer (75 Applis), Stress post-traumatique (82 Applis), douleur (218 Applis), soutien aux personnes âgées (77 Applis), sport (312 Applis).

La méthodologie d'évaluation a été décrite par Stach (84), en 2020. Une exploration des stores en ligne (App Store et Google Play) est effectuée de manière automatisée. Les résultats sont ensuite filtrés et répartis dans les domaines adéquats par les relecteurs. Le score de MARS a été modifié avec une section supplémentaire abordant le « gain thérapeutique » (gain pour le patient, gain pour le thérapeute, les risques et les effets secondaires, facilité à intégrer en pratique courante). Cette version est dénommée par l'auteur MARS-G et a été validée en Allemand par Messner (78), en 2020. À noter qu'une vidéo d'entraînement à l'utilisation spécialement développée est disponible en ligne³⁶.

De manière similaire, l'Australie³⁷ utilise également le score de MARS pour mettre en ligne un répertoire dédié³⁸ avec plus de 300 Applications revendiquées en 2020.

NB : il n'a pas été effectué de recherche exhaustive sur les initiatives similaires dans d'autres pays.

3.3.2.2. u-MARS

Stonayov (85), en 2016, a développé la version usager de MARS. Composé de 20 questions notées sur 5 points (avec une partie qualité et une partie subjective), le score est complété par une partie « impact perçu ».

³⁵ <http://www.mhad.science>

³⁶ <https://www.youtube.com/watch?v=5vwMiCWC0Sc>

³⁷ <https://www.pulseitmagazine.com.au/australian-ehealth/2660-bad-apps-and-where-to-find-them>

³⁸ <https://www.vichealth.vic.gov.au/media-and-resources/vichealth-apps/healthy-living-apps>

3.3.3. Les autres scores de qualité généralistes des Applications

Différentes initiatives cherchent à étendre l'évaluation au-delà de ce que propose le score de MARS. Une liste non exhaustive présentée ci-dessous fait apparaître les catégories évaluées tout en détaillant un peu plus la partie concernant le contenu médical.

L'unité informatique médicale du *Royal College of Physicians* (RCP) a développé en 2015 (86), une check-list en **18 questions en trois thématiques** pour aider les médecins à évaluer les Applications :

- qui a développé l'Application et qu'est-ce qu'elle contient (9 questions) ?
- comment l'Application fonctionne-t-elle (4 questions) ?
- existe-t-il des preuves que l'Application atténue réellement le problème (5 questions) ?

Le ORCHA-24 (*Organisation for the Review of Care and Health Applications*) est une version courte (87), du questionnaire ORCHA. Il a été construit à partir de sources réglementaires, de la norme PAS-277 et de retour d'experts (selon la méthode Delphi). Un descriptif plus détaillé est proposé en annexe 4.

Jin (88), en 2015, a développé un score en s'appuyant sur une revue de littérature et la sollicitation d'un panel d'experts. **Cinq catégories ont été recensées pour 23 questions notées de 0 à 3 :**

- contenu (précision, compréhension) ;
- objectivité (auteurs, expertise) ;
- interface ;
- précision des mots ;
- sécurité technique.

3.3.4. Les scores de qualité spécifiques des Applications

Les scores précédents sont des **scores « généralistes »**, ils sont construits pour être applicables à tous les types d'Applications. Il est possible d'utiliser une autre approche pour évaluer le contenu médical avec un score « spécifique ». Dans ce cas, une échelle est construite pour rechercher les informations clés du thème étudié.

3.3.4.1. Scores spécifiques

Le plus souvent, les auteurs identifient dans des recommandations de bonnes pratiques les informations qui devraient être diffusées et un audit des Applications est réalisé sous forme d'évaluation systématique (*systematic assessment of Apps*) pour identifier les Applications qui reprennent ces informations.

Les critères sont conçus à partir de recommandations de bonnes pratiques, de données factuelles, ou de facteurs psychologiques ou comportementaux (89).

Ces évaluations systématiques sont réalisées le plus souvent lorsque le problème de santé est fréquent et qu'il existe de nombreuses Applications. La dimension éducative de l'information ou le changement comportemental est le plus souvent retrouvé.

Quelques exemples publiés :

- **à partir de recommandations** : conformité des recommandations sur l'activité physique (90, 91), conformité par rapport à une approche comportementale pour l'arrêt du tabac (92) ou donner son sang (93), conformité des recommandations sur l'eczéma (94), enquête auprès

d'utilisateurs pour les Applications pour gérer les problèmes douloureux (95), évaluation d'experts avec 10 questions notées de 1 à 5 pour les Applications concernant les problèmes d'énurésie (96) ;

- **à partir de changement de comportements** : promotion de l'activité physique (97), 62 critères du NICE pour les Applications proposant un changement de comportement (98).

Concernant la validation de ce type d'échelle, la sélection des critères est réalisée avec l'appui de spécialistes du domaine (validité de construit) et la reproductibilité de l'évaluation est mesurée en inter-examineur (concordance statistique qualitative ou quantitative).

Des études spécifiques sont parfois utilisées pour le suivi de recommandation *via* des Applications. Siebert (99), en 2020, a développé une Application (appelé « *Guiding Pad App* ») pour appliquer les recommandations pour la réanimation cardiaque pédiatrique. L'arbre décisionnel publié par une société savante a été découpé en plusieurs étapes, et un essai contrôlé randomisé a évalué l'efficacité de l'utilisation de cette Application.

Watson (100), en 2020, a développé une Application pour la prescription adaptée de la transfusion du sang. Un audit a comparé l'amélioration de la décision clinique avec ou sans l'Application.

DiFilippo (101), en 2017, a développé un score de qualité dénommé AQEL (*App Quality Evaluation*). Après une analyse de littérature et des avis des experts, il a proposé un score qui permet d'effectuer une évaluation globale (de la partie éducative et technique de l'Application), et il a ajouté un complément pour effectuer une évaluation de thématique spécifique. Ce score est dédié aux Applications du domaine de la nutrition.

L'ACDC (*App Chronic Disease Checklist*) est une check-list mise en place pour évaluer les Applications orientées sur les pathologies chroniques par les professionnels de santé (102). Elle est composée de quatre domaines (engagement, fonctionnalités, utilisation simple et gestion de l'information) et de 24 critères. L'objectif est d'inspecter et tester les Applications par des professionnels et rechercher la présence de fonctionnalités spécifiques.

L'intérêt de ce type de référentiel est double, il permet, d'une part, de cibler les contenus de santé que devraient contenir l'Application, et, d'autre part, de s'assurer que ces informations sont présentes, comprises et appliquées.

Un exemple du niveau de détail de l'évaluation du « contenu clinique » pour des Applications de gestion de la douleur est repris en annexe 5 à partir de l'article de Reynoldson (95) en 2014.

Cet exemple de critères de contenu clinique attendu permet d'évaluer le niveau d'exigence souhaité pour ce type d'Applications et la possibilité de comparer les Applications avec un référentiel identique.

3.3.4.2. Scores spécifiques développés par des sociétés savantes

En 2014, l'*American Society of Health-System Pharmacists* a rédigé un premier document de synthèse (103), pour proposer des check-lists afin d'évaluer les Applications concernant les médicaments (fiabilité, calcul de dose, etc.). Loy (104), en 2016, a détaillé davantage le score ciblant les fonctionnalités proposées pour les Applications concernant les médicaments (suivi/*monitoring*, vérification de l'interaction médicamenteuse, calcul de dose, information sur les médicaments, rappel et enregistrement des médicaments).

Dans le même esprit, Camacho (105), en 2020, présente une échelle (TEACH-Apps) soutenue par l'*American Psychiatric Association* composée de huit catégories (données personnelles, publication sur l'efficacité de l'Application, prix, évaluation, attributs, liste de fonctionnalités, intégration technique,

performance). L'évaluation cible les Applications dans le champ de la santé mentale et s'effectue par un comité de quelques experts qui évaluent les Applications en une trentaine de minutes.

Llorens-Vernet (63), en 2020, a réalisé une revue de littérature des différents scores et a compilé les critères identifiés. Un total de 503 critères a été recensé et une sélection de 36 critères considérés comme importants ont été retenus pour huit catégories d'évaluation.

3.3.5. Les référentiels de qualité spécifiques des Applications

Au-delà des scores qui sont souvent limités à une vingtaine de questions, les référentiels proposent une évaluation plus approfondie et couvrent systématiquement plusieurs domaines d'évaluation.

Quelques exemples sont cités ci-dessous, lorsque le contenu médical est abordé, il est davantage détaillé.

Nouri (60), en 2018, a publié une revue de littérature sur les critères pour évaluer la qualité des Applications. Il recense, entre autres, 15 études qui comprennent une évaluation du contenu médical. Un total de sept catégories a été synthétisé avec 37 sous-catégories.

Henson (106), en 2019, a rassemblé les critères de 45 référentiels de qualité et compilé les 604 questions pour en obtenir 357 qui ont été étudiées sur leurs caractéristiques de construction. Au total, 357 questions ont été retenues. Une pyramide en cinq niveaux a réparti les catégories d'évaluation :

- niveau 1 : informations générales ;
- niveau 2 : données personnelles et sécurité ;
- niveau 3 : données factuelles (premières impressions après usage, validité clinique, support des feed-back utilisateurs) ;
- niveau 4 : usage facile ;
- niveau 5 : intégration des données.

3.3.6. Synthèse

Pour résumer, les **scores proposent une analyse rapide des Applications avec une orientation généraliste ou plus approfondie pour les scores spécifiques**. Parmi ces échelles, le score de MARS (*Mobile Application Rating Scale*) est le plus étudié et le plus utilisé. Le tableau 3 compile les exemples cités précédemment en complétant avec d'autres exemples identifiés similaires.

Tableau 3 : Exemples de scores composites évaluant différents domaines de la santé mobile

Nom	Pays	Domaines évalués	Nombre de critères	Commentaires
MARS (75)	Australie	Engagement, Fonctionnalités, Esthétique, Qualité de l'information	23	Validation par différentes études
MARS-G (78)	Allemagne	Idem MARS, mais ajout d'un domaine gain thérapeutique	27	Base de données regroupant les Applications évaluées par ce score
Pas de nom (Jin) ³⁹	Corée du Sud	Contenu (précision, compréhension), objectivité, interface, précision des mots, technologie	23	
Pas de nom (Llorens-Vernet) (63)	Espagne	Facilité d'utilisation, données personnelles, sécurité, pertinence et adéquation, transparence du contenu, sécurité, support technique et mise à jour, technologie	36 questions	
QoE (Martinez) ⁴⁰	Espagne	Qualité du contenu, sécurité, facilité d'utilisation, disponibilité, performance, apparence, apprentissage, précision	21 questions	Évaluation subjective ressentie
Pas de nom (Robustillo) ⁴¹	Espagne	Design et pertinence, qualité et sécurité de l'information, fournitures des services, confidentialité et données personnelles	40 questions	Utilisé pour comparer les Applications
AppScript (IQVIA) ⁴²	États-Unis	Évaluation professionnelle, évaluation patient, évaluation en fonction de six fonctionnalités, soutien institutionnel, techniques de développement, évaluation clinique	ND	
APPLICATIONS (Chyjek) ⁴³	États-Unis	Compréhension de l'Application, prix, littérature utilisée, connectivité, publicité, champ de recherche, compatibilité entre dispositifs, autres composants médias, facilité de navigation, présentation subjective	16	Développé pour comparer les Applications
RCP (86)	Grande-Bretagne	Qui a développé, et qu'est-ce qu'elle contient ? Comment l'Application fonctionne-t-elle ? Existe-t-il des preuves que l'Application atténue le problème ?	18 questions	Aide au praticien pour sélectionner les Applis
ORCHA-24 (87)	Grande-Bretagne	Gouvernance des données, efficacité clinique, expérience utilisateur	24 questions	Extraction d'un questionnaire plus important

³⁹ <https://www.liebertpub.com/doi/pdf/10.1089/tmj.2014.0151>

⁴⁰ <https://link.springer.com/article/10.1007/s10916-013-9976-x>

⁴¹ <https://www.liebertpub.com/doi/pdf/10.1089/tmj.2013.0262>

⁴² <https://www.appscript.net/score-details>

⁴³ https://journals.lww.com/greenjournal/Fulltext/2015/06000/Rating_Pregnancy_Wheel_Applications_Using_the.29.aspx

MedAd-AppQ (Ali 2018) ⁴⁴	Singapour	Fiabilité du contenu, utilité de la fonctionnalité, facilité d'utilisation de la fonctionnalité	24
-------------------------------------	-----------	---	----

ND : non défini

Les référentiels polyvalents proposent plus de 300 critères en moyenne et visent une plus grande spécificité d'évaluation. Le tableau 4 compile les exemples cités précédemment en complétant avec d'autres exemples identifiés similaires.

Tableau 4 : Exemples de référentiels polyvalents pour évaluer la santé mobile

Nom	Pays	Domaines évalués	Nombre de critères	Commentaires
Nouri (60)	France	Design, contenu (crédibilité, précision, qualité de l'information, quantité d'information), facilité d'usage, fonctionnalité, éthique, sécurité et données personnelles, perception utilisateur	ND	
Henson (106)	États-Unis	Informations générales,	357	
Enlight ⁴⁵	États-Unis	Facilité d'utilisation, design visuel, engagement utilisateur, contenu, persuasion thérapeutique, alliance thérapeutique, évaluation générale subjective	476	Association de cinq check-lists : crédibilité, programme fondé sur les preuves, explication du respect des données personnelles, sécurité basique
ORCHA <i>Review</i> (OBR) ⁴⁶	Grande-Bretagne	Non détaillé	260 à 350	

ND : non défini

3.4. Avis des utilisateurs, avis d'experts

Cela a été le premier type d'évaluation rencontré lors de l'apparition des premières Applications.

Des enquêtes ou retours d'utilisateurs sont réalisés et étudiés. Van Haasteren (107), propose la check-list mHAT (pour sonder l'avis des utilisateurs sur la confiance dans l'Application). Ce type d'évaluation est décrit ci-après avec quelques exemples représentatifs.

3.4.1. MyHealthApps-MHA (Grande-Bretagne)

MyHealthApps-MHA est une des plus importantes bases de données d'Applications ouvertes au public et en accès libre (1 160 Applications recensées avec un classement par spécialités médicales et par

⁴⁴ <https://www.sciencedirect.com/science/article/abs/pii/S1551741117306654?via%3Dihub>

⁴⁵ <https://www.jmir.org/2017/3/e82/>

⁴⁶ <https://nhsprocurement.org.uk/wp-content/uploads/2020/11/Attachment-1-ORCHA-Baseline-Review-OBR-Process-.pdf>

pathologies)⁴⁷. Les Applications sont soumises à évaluation en ligne⁴⁸ via un formulaire disponible en neuf langues. La sélection des Applications s'appuie sur deux facteurs principaux⁴⁹ :

- l'Application de santé a été désignée comme favorite par des groupes de patients/handicapés/soignants/familles/consommateurs, ou par des consommateurs autonomes (ex. : des défenseurs des consommateurs, des membres actifs/blogueurs de forums modérés sur la santé des consommateurs) ;
- le développeur de l'Application est transparent sur la nature de l'Application. Les vérifications des éléments sous-jacents à l'Application sont effectuées par PatientView et comprennent différents éléments (prix, l'authenticité de l'évaluation par les utilisateurs, les références des concepteurs et propriétaires de l'Application).

Chaque Application possède sa fiche descriptive et son éventuelle recommandation par une association, une société savante ou une structure officielle.

3.4.2. GGD Appstore (Pays-Bas)

GGD Appstore est une base de données dont l'évaluation est réalisée par des professionnels et des patients (229 Applications recensées, avec possibilité pour une Application de figurer dans plusieurs catégories). Les Applications sont soumises à une évaluation en ligne.

GGD Appstore propose six catégories d'Applications (avec possibilité pour une Application de figurer dans plusieurs catégories)⁵⁰ :

- fonctions corporelles ;
- bien-être mental ;
- sens à sa vie ;
- qualité de vie ;
- participation ;
- fonctionnement quotidien.

Chaque Application possède sa fiche descriptive et deux onglets et un nombre d'étoiles de 0 à 5 en fonction des réponses à une liste de question. Il est possible de télécharger un document de compilation avec le descriptif et les réponses aux questions posées. La méthodologie d'évaluation générale est publiée⁵¹ :

- description ;
- présentation de l'Application, de son utilisation et de son prix ;
- à qui s'adresse l'Application ?
- à quoi peut-on utiliser l'Application ?
- que peut faire l'Application ?
- où la télécharger ?
- évaluation (40 questions) ;
- convivialité (15 questions) ;
- fiabilité (5 questions) ;

⁴⁷ <http://myhealthapps.net>

⁴⁸ <http://myhealthapps.net/submit>

⁴⁹ <http://myhealthapps.net/methodology>

⁵⁰ <https://www.ggdappstore.nl/Appstore/Homepage>

⁵¹ <https://www.ggdappstore.nl/Appstore/Testmethode>

- justification (5 questions) ;
- confidentialité et sécurité (15 questions).

3.4.3. Health Navigator (Nouvelle-Zélande)

Health Navigator est **une des premières bases de données avec des critères structurés**, mises en ligne. Cette base de données est construite à partir d'une recherche d'Applications (219 Applications recensées dans 68 catégories différentes de problèmes de santé ou en rapport avec un aspect d'un mode de vie sain). Les catégories sont déterminées par l'équipe éditoriale de Health Navigator⁵², ou *via* les demandes externes de cliniciens et de consommateurs.

Health Navigator effectue les recherches pour localiser les Applications. Les nouvelles Applications sont identifiées notamment par⁵³ :

- une revue de la littérature des articles de recherche publiés ;
- une recherche sur d'autres sites web d'évaluation d'Applications indépendants ;
- une recherche sur les magasins d'Applications ;
- une recherche sur les Applications tendance sur les réseaux sociaux et les actualités populaires ;
- des alertes auprès de développeurs d'Applications qui peuvent soumettre l'Application *via* un formulaire en ligne⁵⁴ ;
- *via* les utilisateurs du site web, y compris les consommateurs utilisant les Applications.

Une autre recherche documentaire est ensuite effectuée pour identifier les fonctionnalités et critères utiles pour évaluer les Applications dans cette catégorie (ex. : les fonctionnalités souhaitables dans les Applications de rappel de médicaments).

Le processus d'évaluation des Applications comprend quatre étapes :

- évaluation interne (fonctionnalités, qualité des informations produites, public cible) ;
- évaluation de la pertinence pour la Nouvelle-Zélande (unité de mesure, produit alimentaire non disponible en Nouvelle-Zélande) ;
- évaluation clinique (évaluation par un professionnel de santé travaillant dans le domaine concerné, valeur clinique, pertinence, sécurité. Note de 1 à 5 attribuées) ;
- évaluation par les utilisateurs (vise à savoir si l'Application fait ce qu'un utilisateur attend d'elle, ce que l'utilisateur apprécie ou pas dans l'Appli).

Les Applications sont exclues, si elles sont jugées cliniquement dangereuses ou potentiellement dangereuses pour les utilisateurs. Les autres raisons d'exclusion des Applications sont : un contenu incomplet, des problèmes de fonctionnalité et des problèmes de sécurité ou de confidentialité. Le nom de l'Application et le motif de l'exclusion sont documentés sur la page de présentation de la catégorie d'Application.

3.4.4. Onemind (États-Unis)

Onemind est une base de données spécialisée sur une thématique spécifique : la santé mentale⁵⁵ (193 Applications recensées et 55 retirées, car plus téléchargeables). Elle est mise en place par une

⁵² <https://www.healthnavigator.org.nz/apps/a/app-library/>

⁵³ <https://www.healthnavigator.org.nz/apps/p/people-process/#App%20selection%20process>

⁵⁴ <https://docs.google.com/forms/d/e/1FAIpQLSdTd3amG2qAMj-JO5Rn5e2OtQzQnnQIFBKCI1wHO5JKgPJ9PQ/viewform>

⁵⁵ <https://onemindpsyberguide.org/apps/>

association professionnelle qui regroupe un groupe d'experts pilote et un comité de relecteurs de spécialistes. L'évaluation s'effectue sur trois axes notés sur 5 points :

- crédibilité ;
- expérience utilisateur ;
- transparence.

Le nom des relecteurs est cité pour chaque Application évaluée. Des informations techniques et l'état de la recherche sur l'Application sont également documentés dans la description de l'Application évaluée.

3.4.5. Observatoire : App sanitaire (Italie)

Il s'agit ici d'un répertoire/observatoire mis en place au niveau national en Italie en 2016 (App sanitaire⁵⁶) pour recenser les Applications couvrant vingt spécialités médicales. Un total de 640 Applications et 12 ayant obtenu le marquage CE sont répertoriées et décrites.

Le document est une compilation des Applications. Les informations produites sont principalement descriptives.

3.4.6. Synthèse

Le tableau 5 compile les exemples cités précédemment en complétant avec d'autres exemples identifiés similaires.

Tableau 5 : Exemples de sites proposant des évaluations d'Applications par des utilisateurs et/ou des experts

Nom	Pays	Système d'évaluation	Nombre d'Applications	Commentaires
MyHealthApps ⁵⁷	Grande-Bretagne	Évaluation par utilisateurs	1 160	
GGD Appstore ⁵⁸	Pays-Bas	Évaluation par experts et utilisateurs	229 pour les 6 catégories	Certaines Applications appartiennent à plusieurs catégories
Health Navigator ⁵⁹	Nouvelle-Zélande	Évaluation par experts et utilisateurs	219	Certaines Applications appartiennent à plusieurs catégories
Onemind ⁶⁰	États-Unis	Évaluation par experts	193	55 retirées, car plus téléchargeables
Osservatorio APP sanitaire ⁶¹	Italie	Évaluation par experts	640 (en 2016)	12 Applications avec le marquage CE en 2016
HealthOn ⁶²	Allemagne	Évaluation par experts	474	

⁵⁶ <https://www.appsanitarie.it>

⁵⁷ <http://myhealthapps.net>

⁵⁸ <https://www.ggdappstore.nl>

⁵⁹ <https://www.healthnavigator.org.nz/apps/a/app-library/>

⁶⁰ <https://onemindpsyberguide.org/apps/>

⁶¹ <https://www.appsanitarie.it>

⁶² <https://www.healthon.de>

Digimeda ⁶³	Allemagne	Évaluation par experts	Accès payant	Évaluation ⁶⁴ gradée de 1 à 7 (d'essai contrôlé à pas d'évaluation). Site web, chatbots évalués
Groupe Pasteur Mutualité ⁶⁵	France	Évaluation par experts	Revendique plus de 800 Applications évaluées ⁶⁶	Accès restreint

Le nombre d'Applications a été compté fin février 2021

3.5. Labels/certifications avec une liste de critères

Certains organismes publics ou privés mettent en place des labels ou des certifications spécifiques pour permettre aux Applications de se différencier ou de valoriser un processus d'évaluation spécifique (adéquation à une liste de critères, tests techniques, expertise juridique, etc.). Le label est éventuellement mis en avant par le fabricant sur différents supports.

Le schéma suit le modèle d'assurance qualité et de contrôle qualité externe classique. La qualité du service est définie le plus souvent comme la capacité à satisfaire les besoins de l'utilisateur en toute sécurité.

3.5.1. ORCHA Review (Grande-Bretagne)

Système d'évaluation privé⁶⁷, ORCHA (*Organisation for the Review of Care and Health Applications*) a été **une des premières structures en Europe à effectuer un processus d'évaluation automatisé et avec une revue d'experts** (2015).

Conçu par des cliniciens, ORCHA Review propose une accréditation pour les services de santé numérique dans le monde entier. Son examen de base évalue 260 critères (conformité dans les domaines de la sécurité clinique, de la confidentialité des données et de l'expérience utilisateurs, de la réglementation et des normes internationales), et couvre plus de 350 catégories de problèmes de santé.

Chaque évaluation est adaptée et personnalisée en fonction de l'orientation et des capacités fonctionnelles de l'Application. La plateforme procède à des examens automatisés hebdomadaires de toutes les mises à jour de l'Application.

Les Applications sont classées en cinq niveaux (de 0 à 4) en fonction de leur domaine d'intérêt et de leurs capacités fonctionnelles. Plus une Application est axée sur la « santé » et riche en fonctionnalités, plus son niveau est élevé et plus de domaines d'évaluation sont activés pour être évalués. Les Applications très simples axées sur le bien-être (de niveau 0 ou 1) ne sont, par exemple, pas évaluées en fonction de leur performance clinique, car elles ne fournissent pas vraiment de solution clinique contrairement aux Applications de niveau 4. L'évaluation fonctionnelle modifie également de manière dynamique les champs d'évaluation suivis par les examinateurs lors de l'examen proprement dit pour s'assurer que l'évaluation est aussi adaptée que possible au type d'Application.

Un score ORCHA global est construit à partir des réponses à chacune des questions des domaines d'évaluation. Certaines questions rapportent des points positifs et d'autres des points négatifs. Des

⁶³ <https://digimeda.de>

⁶⁴ <https://digimeda.de/ueber-digimeda#ranking>

⁶⁵ <https://www.gpm.fr/actualite.html?id=10093>

⁶⁶ <https://www.ticsante.com/story/2142/e-sante-une-application-recense-et-evalue-plus-de-800-dispositifs.html>

⁶⁷ <https://www.orchacore.co.uk/our-solution/the-orchacore-review/#0>

seuils à 65 % (Application avec des problèmes) et inférieurs à 45 % (Application potentiellement inutile ou dangereuse).

Si une nouvelle version n'est pas produite dans les 18 mois suivant la version actuelle, le lien relatif à l'Application est indiqué comme « obsolète » et le score de l'Application ORCHA commencera à se dégrader au taux de 5 % par mois.

Le cadre d'évaluation propose également une évaluation approfondie couvrant plus de 500 critères dans cinq domaines (y compris la stabilité financière et commerciale, et des adaptations sur mesure en fonction de pays spécifiques).

ORCHA revendique de pouvoir mener des centaines d'examen chaque semaine, dans plus de 180 domaines de conditions et de catégories. À ce jour, ORCHA aurait évalué plus de 6 000 Applications.

ORCHA revoit chaque trimestre ses critères avec des experts internationaux. Il procède actuellement à un examen de l'échelle du système avec NeLL pour préparer 2030⁶⁸.

Les analyses critiques des Applications sont hébergées sur une plateforme qui permet une recherche intelligente pour trouver des Applications par rapport à une gamme de critères précis. La plateforme permet de créer rapidement et facilement des bibliothèques d'Applications sur mesure pour les demandes spécifiques des clients.

ORCHA effectue des revues pour les organisations gouvernementales à travers l'Europe, le Moyen-Orient et l'Australasie.

Au Royaume-Uni, ORCHA mène des revues pour le NHS Digital et le NHS dans 50 % des régions (comme programme national d'accélérateur d'innovation).

3.5.2. TICSS (région de la Catalogne-Espagne)

TICSS est une fondation dépendante du ministère de la santé catalan qui propose différents services. Dans le domaine de la santé mobile, elle propose une évaluation qu'elle nomme accréditation. Le processus fait intervenir des associations professionnelles et un comité d'experts⁶⁹.

Elle couvre quatre domaines :

- **Convivialité et accessibilité.** L'application doit avoir une interface intuitive, avec un design adapté à sa fonction prévue et elle doit assurer un accès universel et inclusif aux personnes avec une diversité fonctionnelle pour maximiser les avantages offerts par la technologie.
- **Technologie.** L'application fonctionne de manière efficace et fiable d'un point de vue technologique. L'application doit s'adapter à un minimum de fonctionnalités acceptable pour l'utilisateur final en garantissant robustesse et cohérence.
- **Sécurité.** Des mécanismes robustes sont en place pour préserver la confidentialité des données générées par les utilisateurs et la plus grande confidentialité dans la transmission de ces informations. Il est nécessaire de garantir un stockage adéquat des informations et d'établir des mécanismes de cryptage lors de l'enregistrement des mots de passe.
- **Fonctionnalité et contenu.** Un comité d'experts, composé de professionnels de différentes associations professionnelles du secteur telles que les associations professionnelles, évalue la

⁶⁸ <https://www.nell.eu/upload/images/news/ORCHA%20mHealth%202030.pdf>

⁶⁹ <https://ticsalutsocial.cat/en/serveis/mhealth-en/accreditation-service-and-ticss-guarantee-certification/>

qualité du contenu et l'utilité des fonctions proposées. Leur examen comprend la convivialité et la conception de l'application, et si elle informe l'utilisateur de toute mise à jour logicielle.

3.5.3. Certification Dekra (France)

MedAppCare a été la première structure à obtenir un niveau de certification accrédité (accréditation n° 5-0598) par le Comité français d'accréditation (Cofrac). Elle a été reprise par Dekra⁷⁰ en 2020.

Elle évalue et certifie la qualité des applications mobiles, plateformes web et services connectés en santé, bien-être, perte d'autonomie et handicap.

Quatre domaines sont couverts :

- la protection des données ;
- la sécurité numérique ;
- la pertinence du contenu ;
- l'ergonomie et l'usage.

La certification, une fois obtenue, est valable trois ans.

3.5.4. Sceau DiaDigital (Allemagne)

Le fabricant de l'Application demande le sceau^{71,72} et effectue une auto-évaluation de son Application (108). Le Centre de télématique et de télémédecine de Bochum (ZTG) effectue un contrôle technique et établit un rapport. Les testeurs d'Applications de DiaDigital effectuent leur évaluation individuelle. Lors d'une conférence téléphonique, à laquelle tous les testeurs peuvent participer, il est vérifié si l'Application répond à tous les critères importants. Les résultats des testeurs sont résumés dans un rapport de conclusion.

L'Application est publiée si le choix de l'auto-divulgaration est décidé. Le résultat de l'examen technique et la conclusion sont disponibles dans la section « Applications certifiées ».

Si une Application ne répond pas aux critères, le fabricant reçoit des commentaires sur le potentiel d'amélioration et peut ensuite postuler à nouveau.

3.5.5. HON code (Suisse)

La Fondation HON qui proposait une évaluation pour les sites web a développé une évaluation pour les Applications : le HON code⁷³. Présenté comme une certification, ce système d'évaluation semble plutôt être un label. Il est composé de huit domaines :

- autorité (équipe éditoriale) ;
- complémentarité (limite de l'Application) ;
- confidentialité (exigence légale) ;
- validité (date de mise à jour) ;
- justifiabilité/objectivité (références complètes et objectives) ;
- expérience utilisateur (facilité d'utilisation et d'accès au support) ;
- financement (sources de financement et transparence de service payant) ;
- politique publicitaire (identification et séparation de la publicité).

⁷⁰ <https://www.dekra-certification.fr/certification-de-services/certification-applications-mobiles-et-sites-web-dekra-certification.html>

⁷¹ <https://appcheck.de/bewertung-durch-diadigital-und-pneumodigital/>

⁷² <https://www.diabetesde.org/diadigital>

⁷³ <https://www.hon.ch/fr/certification/app-certification.html>

L'évaluation est effectuée par des experts médicaux ou juridiques pour la protection des données. HON code propose également des tests automatisés de vulnérabilités ou des tests de confidentialité ou de cryptage.

3.5.6. Synthèse

Le tableau 6 compile les exemples cités précédemment en complétant avec d'autres exemples identifiés similaires.

Tableau 6 : Exemples de structures proposant des labels, certification ou accréditation

Nom	Pays	Système d'évaluation	Nombre d'Apps
ORCHA	Grande-Bretagne	Label	Plus de 6 000
TICSS	Espagne (Catalogne)	Accréditation	4 (01/21)
Dekra (anciennement MedAppCare)	France	Certification	ND
OMH (Our Mobile Health) ⁷⁴	Grande-Bretagne	Label	ND
DiaDigital	Allemagne	Label	ND
HON code	Suisse	Label	ND
OMH (Our Mobile Health)	Grande-Bretagne	Label	ND

Le nombre d'Applications a été compté fin février 2021. ND : non défini

⁷⁴ <https://www.ourmobilehealth.com>

3.6. Répertoire éditorialisé constitué d'une base de données proposant une annotation spécifique de chaque Application

Certains organismes publics mettent en place des bases de données d'Applications (répertoire) en ajoutant une évaluation très approfondie et spécifique sur plusieurs domaines.

3.6.1. Les référentiels d'évaluation au niveau des pays européens

Les domaines d'évaluation de la qualité des Applications mobiles sont variés. Un **réseau européen est en place pour recenser les initiatives européennes et communiquer les référentiels utilisés.**

En 2016, de manière simultanée à l'élaboration du référentiel HAS (1), un groupe de travail organisé par la Commission européenne était constitué pour élaborer des recommandations européennes sur les critères d'évaluation des Applications mobiles et objets connectés. Ce travail faisait suite au livre vert publié sur la santé mobile en 2014⁷⁵. L'objectif était de déterminer les données qui pouvaient être entrées dans le dossier du patient électronique *via* ces outils numériques spécifiques⁷⁶.

Le groupe de travail n'a pas pu aboutir à un consensus, car, pour simplifier, deux tendances s'opposaient :

- une tendance qui souhaitait une évaluation multidomaine ;
- une tendance qui souhaitait une évaluation uniquement sur la fiabilité de la donnée qui pouvait entrer dans le dossier du patient électronique.

Au final, le référentiel de la HAS a été cité parmi les six réalisations (deux britanniques, deux espagnoles, une allemande) de recommandation au niveau européen⁷⁷. Il est également cité par le rapport européen sur la sécurité des logiciels embarqués publié en 2019 (SMART 2016/071)⁷⁸.

Le travail de recommandation européenne pour la santé mobile s'est poursuivi avec un appui de l'Organisation mondiale de la santé (OMS), l'union internationale des communications (ITU), le ministère de la Santé de la région Andalouse, et une bourse de la Commission européenne avec le « *European mHealth Innovation and Knowledge Hub*⁷⁹ ». Le référentiel de la HAS est cité dans le rapport (en cours de finalisation lors de la rédaction de ce document) comme cadre d'évaluation parmi 24 référentiels couvrant neuf pays.

Ce rapport⁸⁰ propose une chronologie récapitulative des différents référentiels ou systèmes d'évaluation utilisés en Europe sur les douze dernières années.

2008 : Continua Design Guidelines ;
2012 : AppCheck ; MedAppCare ; région andalouse ;
2013 : My Health Apps ; Our Mobile Health ; NICE initiative ;
2015 : ORCHA ; PAS 277 ; TicSalutSocial ;
2016 : DiaDigital ; GGD Appstore ; recommandations européennes ; HAS ;
2017 : mindapps.dk ; NHS digital
2018 : APPKri ; cMHAFF, PneumoDigital ; MySNS Seleção ; mHealthBelgium ;
2019 : AppQ ; eHealth Suisse ;
2020 : BfArM ; initiative ISO.

⁷⁵ <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:52014DC0219&from=EN>

⁷⁶ <https://ec.europa.eu/digital-single-market/en/news/report-working-group-mhealth-assessment-guidelines>

⁷⁷ <https://ec.europa.eu/digital-single-market/en/news/report-working-group-mhealth-assessment-guidelines>

⁷⁸ <https://op.europa.eu/fr/publication-detail/-/publication/aad6a287-5523-11e9-a8ed-01aa75ed71a1/language-en>

⁷⁹ <https://mhealth-hub.org/assessment-frameworks>

⁸⁰ <https://mhealth-hub.org/download/d2-1-knowledge-tool-1-health-apps-assessment-frameworks-pending-ec-approval> (page 53)

mHealth Hub a passé en revue ces 24 référentiels d'évaluation et synthétisé les critères concordants ou pas entre eux et proposé 28 recommandations d'évaluation pour les domaines analysés. L'objectif pourrait permettre de développer une reconnaissance mutuelle en comparant les critères utilisés par les différents pays.

Les **12 domaines d'évaluation** analysés étaient les suivants :

- respect des données personnelles ;
- transparence ;
- sûreté ;
- fiabilité ;
- validité ;
- interopérabilité ;
- stabilité technique ;
- efficacité ;
- accessibilité ;
- évolutivité ;
- expérience utilisateur et convivialité ;
- sécurité.

À noter que l'Union européenne commence à proposer des outils spécifiques de type check-list pour les développeurs d'Applications (ex. : FI-STAR⁸¹).

Une autre voie « orientée ISO » et standardisation de l'évaluation, citée précédemment, est aussi étudiée au niveau européen avec plusieurs normes ISO, dont une orientée spécifiquement sur les Applications (ISO-82304-2⁸²) qui sera publiée en septembre 2021.

Des financements européens (programme « horizon Europe ») sont proposés en 2021 pour le développement d'un label européen⁸³.

3.6.2. Les domaines d'évaluation de eHealth Suisse

eHealth Suisse après une revue de littérature (15) proposent **neuf domaines** d'évaluation déclinés en 18 caractéristiques et 25 exigences (109) :

- transparence ;
- adéquation au but ;
- proportionnalité du risque ;
- acceptabilité éthique ;
- conformité au droit ;
- validité du contenu ;
- adéquation technique ;
- facilité d'utilisation ;
- efficience des ressources.

⁸¹ <https://www.fi-star.eu/publications/checklist-for-app-developers.html>

⁸² https://www.iso.org/news/isofocus_141-6.html

⁸³ <https://digitalhealtheurope.eu/resources/funding-opportunities/promoting-a-trusted-mhealth-label-in-europe-uptake-of-technical-specifications-for-quality-and-reliability-of-health-and-wellness-apps/>

Ci-dessous un exemple du domaine « validité du contenu » avec les caractéristiques à auto-déclarer et les exigences citées :

- **validité du contenu** (les contenus et les fonctions font appel à des sources valides. Ils reposent sur des faits probants qui respectent les éventuelles directives applicables et qui reflètent les connaissances [scientifiques] récentes) :
 - **validité des sources** : il est demandé d'indiquer comment la qualité des contenus est garantie (ex. : recours à des experts du domaine concerné) et quelles sources valides sont utilisées (notamment prise en compte de nouvelles connaissances, directives, études, etc. scientifiques en précisant le niveau de preuve) ;
 - **actualité des contenus** : il est demandé d'indiquer comment l'App est régulièrement adaptée de manière visible aux nouvelles exigences en matière de contenu.

L'intérêt de ce modèle est de couvrir de nombreuses dimensions et de limiter la quantité des exigences demandées. Le système n'a pas encore évalué d'Application en 2020.

3.6.3. Répertoire du NHS Health Apps Library (Grande-Bretagne)

Parmi les **premiers répertoires mis en place au niveau national**, le *NHS Health Apps Library*⁸⁴ a subi différentes évolutions.

La sélection des domaines et des critères d'évaluation à utiliser s'est appuyée sur le rapport du Digital Health & Care Institute (58), en 2018, supporté par le programme européen INTERREG. Ce rapport a synthétisé les différentes approches d'évaluation de la santé. Il a permis de définir les domaines devant être évalués :

- protection des données personnelles ;
- sources fiables – information fondée sur des preuves (evidence-based information) ;
- expérience utilisateur et facilité d'utilisation ;
- fonctionnalité ;
- authentification et sécurité ;
- impact – efficacité ;
- interopérabilité.

Ce rapport a permis de mettre en place le *Digital Assessment Questions* (DAQs) et le *Digital Assessment Portal* (DAP) qu'utilisait le NHS⁸⁵ pour évaluer les produits de santé jusqu'en octobre 2020.

En février 2021, une nouvelle version des modalités d'évaluation est proposée. Elle s'appuie sur le *Digital Technical Assessment Criteria* (DTAC)⁸⁶ et repose sur quatre domaines techniques et un domaine couvrant les principes clés du succès.

Les quatre domaines techniques :

- sécurité clinique ;
- protection des données personnelles ;
- sécurité technique ;
- critères d'interopérabilité.

⁸⁴ <https://www.nhs.uk/apps-library/>

⁸⁵ <https://digital.nhs.uk/services/nhs-apps-library>

⁸⁶ <https://www.nhs.uk/key-tools-and-info/digital-technology-assessment-criteria-dtac/>

Le cinquième domaine propose des principes clés du succès sur l'accessibilité et les modalités d'utilisation de l'Application.

Le tableau 7 reprend un exemple de présentation des deux premiers critères sur la sécurité clinique⁸⁷.

Tableau 7 : Exemples de critères décrits dans le référentiel du NHS (Grande-Bretagne)

Code	Question	Options	Renseignement complémentaire	Score du critère
C1.1	Avez-vous entrepris une évaluation de gestion des risques cliniques en accord avec le standard DCB0129 ?	Oui/Non	Le standard DCB019 ⁸⁸ s'applique aux organisations qui sont responsables du développement et du maintien de système informatique en santé. Un système informatique en santé est défini comme « un produit utilisé pour fournir une information électronique dans un but de santé ou de soins sociaux ».	Pour passer ce critère, le développeur doit confirmer qu'il a entrepris des activités de gestion des risques cliniques conforme au standard DCB019.
C1.1.1	Merci de détailler votre système de gestion des risques cliniques	Document joint/pas de preuve jointe	Le standard DCB019 décrit les activités qui doivent et devraient être mises en œuvre pour des systèmes informatiques en santé. Un modèle de système de gestion des risques en santé est téléchargeable ⁸⁹ sur le site du NHS digital.	Pour passer, le développeur doit montrer des preuves qu'un système de gestion des risques cliniques en santé est mis en place et qu'il est conforme au standard DCB019. Cela inclut : <ul style="list-style-type: none"> – les modalités de gouvernance de la gestion des risques cliniques qui sont en place ; – les activités de gestion des risques cliniques ; – les compétences et les formations en sécurité clinique ; – les audits.

Cet exemple de critères montre le niveau de détail et des pièces à fournir ainsi que la documentation de soutien mise en place pour aider les développeurs à répondre à chaque critère et sous-critère.

Parallèlement à cette liste de critères, le NHS s'appuie sur un cadre de référence produit par le *National Institute for Health and Care Excellence (NICE)* appelé « *Evidence standards framework for digital health technologies* ». Une actualisation en avril 2021 a été publiée et concerne les dispositifs médicaux⁹⁰.

Ce cadre de référence définit le niveau des preuves demandées en fonction d'une classification fonctionnelle des dispositifs médicaux numériques. À noter que Nwe (110), en 2020, a mis en évidence une mauvaise reproductibilité de cette classification (Kappa de 0,32, avec un IC95% compris entre 0,16 et 0,47).

⁸⁷ https://www.nhsx.nhs.uk/documents/60/DTAC_version_1.0_FINAL_updated_16.04.odt

⁸⁸ <https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0129-clinical-risk-management-its-application-in-the-manufacture>

⁸⁹ <https://digital.nhs.uk/services/clinical-safety/documentation#clinical-risk-management>

⁹⁰ <https://www.nice.org.uk/corporate/ecd7/chapter/section-a-evidence-for-effectiveness-standards#functional-classification-of-dhts>

Concernant le niveau demandé pour la qualité du contenu médical, le minimum demandé est :

- **valide** (en accord avec les meilleures sources disponibles, telles que les recommandations du NICE, les organisations professionnelles pertinentes ou les organisations de patients Britanniques reconnues, et appropriées pour la population cible) ;
- **exact** ;
- **à jour** ;
- **révisé et mis à jour par des experts compétents** à des intervalles définis, par exemple chaque année ;
- **suffisamment complet**.

En 2020, Rowland (111) cite les fonctionnalités de l'approche NHS pour les Applications pour les patients. Elles visent :

- un appui du diagnostic clinique et/ou de la prise de décision ;
- une amélioration des résultats cliniques à partir de chemins thérapeutiques préétablis et au travers de changements comportementaux, et d'amélioration de l'adhérence et d'observance des traitements ;
- à agir en tant que solution thérapeutique numérique autonome ;
- à dispenser une éducation par rapport à la maladie.

En 2020, une petite centaine d'Applications ont été indexées dans le répertoire du NHS.

3.6.4. AppSaludable (région Andalousie, Espagne)

Développé par la région Andalousie en 2012, AppSaludable⁹¹ est une des premières plateformes mises en place en Europe pour recenser les Applications mobiles (112).

L'évaluation couvre quatre domaines et 31 recommandations et s'effectue en quatre phases (formulaire de soumission, auto-évaluation, évaluation externe et certification de la région).

Le détail des recommandations est cité en annexe 6.

3.6.5. Le Portugal

Développé au niveau national au Portugal, MySNS Seleção suit le même modèle d'évaluation que la région Andalousie⁹², les recommandations proposées au niveau européen⁹³, et le code de respect des données personnelles européen⁹⁴.

Chaque Application est soumise à un processus d'évaluation⁹⁵ composé de quatre critères :

1. **performance** : assurer une interface intuitive, une conception adéquate permettant l'accès au plus grand nombre de personnes, afin de maximiser les avantages offerts par la technologie ;
2. **sécurité** : respecter les exigences légales en matière de protection des données, en tenant compte de la nature sensible des informations enregistrées et en garantissant la vie privée et la confidentialité de l'utilisateur ;

⁹¹ <http://www.calidadappsalud.com/distintivo/catalogue>

⁹² <https://digitalhealtheurope.eu/twinnings/dhe-twinning-results/appsaludable/>

⁹³ https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20160607_co06_04_en.pdf

⁹⁴ <https://ec.europa.eu/digital-single-market/en/news/code-conduct-privacy-mhealth-apps-has-been-finalised>

⁹⁵ <https://mysns.min-saude.pt/criterios-de-avaliacao/>

3. **utilité publique** : s'assurer que l'Application ajoute de la valeur et contribue à une amélioration significative de la vie des citoyens et/ou des professionnels de la santé ;
4. **qualité et sécurité de l'information** : garantir la crédibilité et l'actualité des informations, dans une perspective de transparence, en indiquant les sources d'information, les éventuels conflits éthiques et les contenus publicitaires.

Le processus d'évaluation est décrit en cinq phases avec une phase de plus que la région Andalousie qui est préparatoire avant la soumission de l'Application.

3.6.6. Synthèse

Le tableau 8 est une compilation des exemples cités précédemment.

Tableau 8 : Exemples de structures proposant des répertoires éditorialisés régionaux ou nationaux

Nom	Pays	Système d'évaluation	Nombre d'Applications	Commentaires
NHS Health Apps Library (premier lancement en 2013-2015 ⁹⁶ et relance 2017 [beta] ⁹⁷ et 2019-) ⁹⁸ mise à jour en 2021	Grande-Bretagne	DAQS (de 55 à 243 questions) et DAP DAQS : 7 domaines explorés	92 (06/20)	Nouvelle version d'évaluation en février 2021
NHS Health Apps Library (février 2021) ⁹⁹	Grande-Bretagne	DTAC 5 domaines explorés		Nouvelle version d'évaluation (02/21)
AppSaludable ¹⁰⁰	Espagne (Andalousie)	4 domaines	39 et 81 en cours (01/21)	
MySNS Seleção ¹⁰¹	Portugal	Jumelé avec Andalousie	ND	

Le nombre d'Applications a été compté fin février 2021. ND : non défini

3.7. Répertoire éditorialisé constitué d'une base de données proposant une annotation spécifique de chaque Application sur un secteur spécialisé

Certains organismes publics ou privés mettent en place des bases de données d'Applications (répertoire) en ajoutant une évaluation très approfondie et spécifique sur des domaines spécifiques.

3.7.1. *Our Mobile Health (OMH) (Grande-Bretagne)*

OMH a mis en place une base de données d'Applications généraliste¹⁰² ou ciblée, par exemple sur la maladie de Parkinson¹⁰³.

⁹⁶ <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-015-0451-z>

⁹⁷ <https://mhealthintelligence.com/news/uk-tries-again-with-a-library-of-certified-mobile-health-apps>

⁹⁸ <https://mhealthintelligence.com/news/uks-nhs-moves-to-fast-track-mhealth-app-validation-process>

⁹⁹ <https://www.nhs.uk/apps-library/>

¹⁰⁰ <http://www.calidadappsalud.com>

¹⁰¹ <https://www.mysns.min-saude.pt/mysns-selecao/>

¹⁰² <https://www.ourmobilehealth.com>

¹⁰³ <https://www.parkinsons.org.uk/information-and-support/apps-and-devices-parkinsons>

Le processus d'évaluation couvre les domaines suivants : la sécurité des données, la conformité aux réglementations, la stabilité technique, la convivialité et les éléments d'efficacité. Un panel de plus de 150 examinateurs experts indépendants issus de milieux cliniques et non cliniques fournit des évaluations détaillées de chaque Application.

OMH dispose d'un vaste réseau de développeurs d'Applications de santé proposant du contenu dans divers domaines, notamment le diabète, la santé des femmes, la remise en forme, la gestion du cancer, la nutrition, le suivi cardiaque et la santé mentale.

Le répertoire d'Applications contient plus de 200 Applications qui ont été soumises au processus d'évaluation.

3.7.2. Santé mentale avec AppChecker (Danemark)

Le système AppChecker¹⁰⁴ cible les Applications dans le champ de la santé mentale.

Le processus d'évaluation comprend plusieurs étapes :

→ Étape 1 – Informations sur l'Application.

Dans cette partie, l'auto-évaluation documente les informations couvrant les spécifications techniques, qui est le développeur, comment l'Application a été développée, le groupe cible, la langue, le système d'exploitation, le prix, etc.

→ Étape 2 – Évaluation de la sécurité et de la confidentialité.

L'évaluation de la sécurité et de la confidentialité se compose de trois étapes commençant par une évaluation des risques. L'évaluation des risques comporte huit niveaux (noté de R1 à R8). Le niveau de risque de l'Application détermine le nombre de paramètres à évaluer. Si le niveau de risque est supérieur à R1, un arbre décisionnel indiquant si les données personnelles sensibles sont collectées ou non est proposé. Si le niveau de risque est supérieur à R5, un arbre décisionnel est proposé pour décider si l'Application doit porter le marquage CE.

Les trois étapes pour évaluer la sécurité et la confidentialité :

- L'évaluation des risques. Dans l'évaluation des risques, il est vérifié si l'Application collecte ou non des données, et si oui, quelles données et comment. Un niveau de risque est évalué dans un tableau d'évaluation qui correspond à l'Application. Le niveau de risque détermine si une vérification de sécurité et de confidentialité des Applications est nécessaire et si elle doit être étiquetée comme dispositif médical (marquage CE).
- Évaluation de la sécurité et de la confidentialité des données. Si l'Application est placée au-dessus de R1 dans l'évaluation des risques, la sécurité et la confidentialité sont évaluées. Dans cette partie, un arbre décisionnel est proposé concernant la sécurité et la confidentialité, qui aide à déterminer si l'Application est conforme aux exigences liées à la sécurité et à la confidentialité. Les Applications et la collecte de données peuvent varier en complexité, ainsi certaines Applications réussiront le test après une ou deux questions, tandis que d'autres devront passer par huit questions.
- Évaluation de la nécessité du marquage CE. Peu d'Applications auront besoin d'obtenir la conformité avec le marquage CE, mais si l'évaluation des risques à l'étape 1 place l'Application dans la catégorie R5 ou plus, il est nécessaire de vérifier si un marquage CE est nécessaire.

¹⁰⁴ <https://mindapps.dk/en/guidance-to-use-the-app-checker/>

→ Étape 3 – Évaluation de la qualité.

Il s'agit de l'évaluation finale de la qualité de l'Application au travers de 4 domaines, chacune contenant trois questions :

- informations et transparence ;
- qualité clinique ;
- fonctionnalité ;
- convivialité.

Pour chaque question, un score de 1 à 3 est appliqué ce qui permet d'établir un score global moyen. Il faut qu'une Application obtienne un score supérieur à 1,50 pour être recommandée.

3.7.3. AppCheck (Allemagne)

Financé par le ministère de la Santé pour la région Rhénanie-du-Nord–Westphalie, AppCheck effectue une évaluation interne sur les Applications concernant des pathologies chroniques pulmonaires (asthme, etc.) et le diabète¹⁰⁵. Il vise à aider les patients à utiliser des Applications recommandées pour faciliter la gestion de la maladie.

3.7.4. Synthèse

Le tableau 9 compile les exemples cités précédemment.

Tableau 9 : Exemples de structures proposant des répertoires éditorialisés spécialisés

Nom	Pays	Spécificités d'évaluation	Nombre d'Apps	Commentaires
AppCheck	Allemagne	Pneumologie et diabète	17	
AppCheck	Danemark	Santé mentale	ND	Évaluation du risque
OMH (<i>Our Mobile Health</i>)	Grande-Bretagne	Parkinson	ND	Flexibilité sur d'autres thématiques
ORCHA	Grande-Bretagne	Sur demande	ND	

Le nombre d'Applications a été compté fin février 2021. ND : non défini

¹⁰⁵ <https://appcheck.de/zertifizierte-apps-2/>

3.8. Répertoire éditorialisé constitué d'une base de données proposant une sélection spécifique d'Applications définie comme des dispositifs médicaux

3.8.1. mHealthBelgium (Belgique)

En Belgique, un portail¹⁰⁶ est proposé depuis 2019 pour les Applications mobiles qui sont des dispositifs médicaux. Trois niveaux sont proposés.

Un **niveau 1 qui correspond au marquage CE**, un niveau 2 avec une analyse de risque indépendante¹⁰⁷, et un niveau 3 demandant la publication d'évidence clinique au travers d'essais contrôlés randomisés. Ce niveau permet un financement de l'assurance maladie belge.

La fiche technique détaille les niveaux d'exigence en fonction de ces trois paliers¹⁰⁸.

3.8.2. DiGA (Allemagne)

L'Allemagne a mis en place un processus d'évaluation des dispositifs médicaux orientés en partie vers les Applications mobiles et dénommés DiGA (*Digital Health Applications*)¹⁰⁹. **L'évaluation proposée permet une admission préliminaire selon certains critères.** Une période probatoire est proposée durant un an pour faire ses preuves cliniques (effets positifs sur les soins) entraînant une indexation dans le répertoire et la détermination du service médical.

Le processus suit les étapes suivantes :

- le fabricant soumet son Application ;
- le BfArM (*Bundesinstitut für Arzneimittel und Medizinprodukte*) qui est l'Institut fédéral des médicaments et des dispositifs médicaux (équivalent de l'Agence nationale de sécurité du médicament et des produits de santé – ANSM – en France) donne un avis et examine le dossier ;
- évaluation de trois mois portant sur les prérequis (sécurité, fonctionnalités, qualité, protection des données personnelles, sécurité des données, interopérabilité) et sur les effets positifs de la prise en charge (bénéfice médical, amélioration structurelle ou des procédures) ;
- trois décisions sont possibles :
 - rejet ;
 - si les prérequis sont remplis, mais pas de données suffisantes pour montrer des effets positifs, une admission préliminaire est proposée pour une durée de douze mois. Une justification des résultats plausibles est demandée. À la fin de cette période probatoire, le rejet ou l'admission est décidé ;
 - admission et publication dans le répertoire (détermination du service médical rendu) et éventuelle négociation de prix et de remboursement.

De manière similaire au répertoire belge, le processus de sélection allemand de **DiGA vise également à déterminer les Applications éligibles à un remboursement en fonction du service médical rendu.**

¹⁰⁶ <https://mhealthbelgium.be/fr/toutes-les-apps>

¹⁰⁷ <https://mhealthbelgium.be/wp-content/uploads/2019/09/Criteria-mHealth-apps-ENv4.pdf>

¹⁰⁸ <https://mhealthbelgium.be/fr/pyramide-de-validation>

¹⁰⁹ <https://diga.bfarm.de/de/verzeichnis>

3.8.3. Digi-HTA (Finlande)

En Finlande, un répertoire cible également uniquement des dispositifs médicaux. Digi-HTA¹¹⁰ **propose des recommandations Digi-HTA spécifiques au produit qui sont basées sur les informations du produit fournies par l'entreprise**. Ces informations sont complétées par une revue de la littérature, une évaluation d'experts et d'éventuelles questions supplémentaires pour l'entreprise. Les informations sur les produits sont collectées à l'aide d'un questionnaire Digi-HTA¹¹¹.

En ce qui concerne la sécurité des informations et la protection des données, les informations sont collectées sur deux documents. Ces documents sont basés sur les exigences en matière de sécurité des informations et de protection des données du Centre national de cybersécurité de Finlande (NCSC-FI)¹¹². Par défaut, la recommandation est valable trois ans. Le produit sera réévalué si de nouvelles informations importantes deviennent disponibles et que l'entreprise demande une réévaluation.

L'évaluation couvre les dispositifs médicaux numériques (santé mobile, intelligence artificielle, robotique) (113). Un total de 11 domaines et composé d'une centaine de questions est proposé selon la répartition suivante

- information de la société (3 questions) ;
- information du produit (21 questions) ;
- stabilité technique (6 questions) ;
- coût (6 questions) ;
- efficacité (6 questions) ;
- sécurité clinique (9 questions) ;
- protection et sécurité des données personnelles (déporté sur 2 documents dédiés) ;
- utilisation et accessibilité (11 questions et déportés sur 5 documents dédiés) ;
- interopérabilité (10 questions, dont 2 conditionnelles et 2 documents dédiés) ;
- intelligence artificielle (22 questions) ;
- robotique (4 questions).

Un score sur 10 points est attribué après évaluation de cinq domaines notés 2 points, 1 point ou - 4 points :

- efficacité (suffisante, prometteuse, faible ou inconnu) ;
- sécurité (suffisante, probablement à un niveau suffisant, mais pas assez connu, faible ou inconnu) ;
- coût (raisonnable, élevé, irraisonnablement élevé) ;
- protection et sécurité des données (suffisante, manquement mineur, manquement) ;
- utilisation et accessibilité (suffisante, manquement mineur, manquement) ;

Le score est découpé en cinq niveaux : 10 (vert foncé : recommandé) ; 9 (vert clair : il y a une chose à considérer en utilisant le produit) ; 7-8 (jaune : il y a différentes choses à considérer en utilisant le produit) ; 5-6 (orange : il y a beaucoup de choses à considérer en utilisant le produit) ; ≥ 4 (rouge : il y

¹¹⁰ <https://www.ppsphp.fi/Tutkimus-ja-opetus/FinCCHTA/Sivut/Digi-HTA.aspx>

¹¹¹ <https://www.oulu.fi/cht/digihealthhub/digi-hta>

¹¹² <https://www.kyberturvallisuuskeskus.fi/en/ncsc-news/instructions-and-guides/information-security-and-data-protection-requirements-social>

a des choses critiques à considérer en utilisant le produit). Un document exportable pour chaque produit évalué est généré avec le score/couleur de chaque domaine.

3.8.4. Synthèse

Le tableau 10 les exemples compile les exemples cités précédemment.

Tableau 10 : Exemples de structures proposant des répertoires éditorialisés pour les dispositifs médicaux

Nom	Pays	Système d'évaluation	Nombre d'Apps	Commentaires
mHealthBelgium	Belgique	3 niveaux du DM à la preuve socio-économique 6 domaines	25	Dispositif médical uniquement
DiGA	Allemagne	2 phases : P1 : prérequis technique + effets médicaux ou amélioration des procédures. P2 : justification possible.	11	Période probatoire de 12 mois pour faire ses preuves cliniques
Digi-HTA	Finlande	11 domaines et des documents dédiés	4	Porte sur les Applis et la robotique
<i>Evidence standard framework (ESF) for digital health technology</i> ¹¹³	Grande-Bretagne	Classification fonctionnelle et questions spécifiques	ND	S'appuie sur une classification fonctionnelle. En cours de réactualisation en avril 2021

Le nombre d'Applications a été compté fin février 2021. ND : non défini

Pour démarquer le secteur de la e-santé de l'HTA (*Health Technology Assessment*), les produits et technologies de santé numérique et les processus d'évaluation sont regroupés sous l'**appellation eHTA** pour « *electronic Health Technology Assessment* ».

Des revues systématiques abordent les problématiques liées à l'évaluation des technologies numériques dans le monde de la santé (HTA).

Par exemple sur :

- les approches évaluatives (114) ;
- la concordance des critères au niveau européen de *European Network for Health Technology Assessment* (EUnetHTA) (115) ;
- les processus de décisions pour évaluer ces solutions numériques (116) ;
- les cadres d'évaluation communs et spécifiques pour des produits particuliers (santé mobile, intelligence artificielle, robotique) (113) ;
- un cadre d'évaluation spécifique (technique et matériel, social et personnel, politique et organisationnel) (117) ;
- le développement d'un module spécifique pour les Applications (118).

La nature et la durée de l'évaluation sont un challenge à relever pour des solutions numériques innovantes et agiles (régulièrement mises à jour) qui ne peuvent pas subir des délais de traitement ou

¹¹³ <https://www.nice.org.uk/corporate/ecd7>

de processus d'évaluation trop long. Par exemple, une Application dans le domaine de la santé mentale disparaît tous les 2,9 jours des magasins en ligne (119).

La FDA (*Food and Drug Administration*) propose une voie rapide pour les solutions qui sont sur des cycles courts¹¹⁴. Le réseau EUnetHTA propose également des « *Rapid Relative Effectiveness Assessment* » (REAs) (114).

Cependant, Alon (120), en 2020, montre que l'évaluation du risque par le biais de catégories de la FDA (précertification) n'a pas permis d'identifier les Applications à risque parmi le TOP10 des meilleures Applications couvrant six problèmes de santé (addiction, anxiété, dépression, diabète, hypertension et schizophrénie).

La santé mobile est une partie spécifique de l'ensemble du secteur de la santé numérique et évolue sur des cycles encore plus courts et avec des modèles économiques souvent limités (Application gratuite, etc.).

Pour conclure

Les stratégies d'évaluations du **contenu médical** des Applications sont orientées le plus souvent sur la qualité de l'édition scientifique (source, expertise des auteurs, liens d'intérêts, actualisation, niveaux de preuves). Les outils utilisés sont des scores, des auto-évaluations à partir de critères généralistes ou des référentiels spécifiques pour des Applications ciblées.

Lorsque l'Application **utilise ou traite des données de santé**, les évaluations deviennent plus complexes et les outils d'évaluations utilisés vont des scores ou des labels à des référentiels d'évaluation de plusieurs centaines de critères ou des tests externes.

Des organisations (non lucratives, commerciales ou gouvernementales) compilent ces différentes évaluations et leurs résultats dans des bases de données sous forme de « répertoire éditorialisé ».

En 2020, cet état des lieux montre une grande variété des systèmes d'évaluation à l'échelle d'un même pays ou au niveau international. Même si les dimensions de l'évaluation sont globalement les mêmes, le niveau de précision des critères et leur découpage rendent difficile la mise en place d'équivalence.

Des initiatives européennes ont eu lieu en 2016 et se poursuivent actuellement avec le « *European mHealth Innovation and Knowledge Hub* » ou la norme ISO-82304-2. L'objectif étant de normaliser les systèmes d'évaluation pour que les utilisateurs et les fabricants puissent mieux s'y retrouver ou mettre en place des systèmes de reconnaissance mutuelle.

En attendant une éventuelle harmonisation, les concepteurs doivent porter une attention particulière sur des informations clés qui devraient être publiques et qui permettraient l'évaluation externe des Applications. Ces informations devraient pouvoir répondre aux critères d'évaluation le plus souvent retrouvés dans les divers outils d'évaluation cités précédemment.

¹¹⁴ <https://www.fda.gov/medical-devices/device-software-functions-including-mobile-medical-applications/examples-premarket-submissions-include-mmas-cleared-or-approved-fda>

4. Critères de qualité du contenu médical

Pour définir les critères de qualité du contenu médical, ce document s'appuie sur le référentiel HAS de 2016 (1) publié pour évaluer la qualité des Applications en santé mobile. Il recensait à l'époque cinq domaines et 14 sous-domaines qui sont listés ci-dessous :

- informations utilisateurs (description, consentement) ;
- **contenu de santé** (conception du contenu initial, standardisation, contenu généré, contenu interprété) ;
- contenant technique (conception technique, flux des données) ;
- sécurité/fiabilité (cybersécurité, fiabilité, confidentialité) ;
- utilisation/usage (utilisation/design, acceptabilité, intégration/import).

4.1. Les différents types de contenu de santé

Concernant le contenu de santé, le référentiel faisait la distinction entre le contenu publié pour diffuser des informations générales (appelé contenu initial) et le contenu généré par l'utilisateur *via* des données collectées par des objets connectés, des questionnaires, des biocapteurs incorporés dans le smartphone, etc.

Pour ce document, qui ne traite que du contenu médical, cette distinction est reprise, avec **quatre sections différentes** correspondant à des types de contenu spécifiques :

- le **contenu initial** : en relation avec la fonction d'informer ou d'instruire de manière générale ;
- le **contenu généré** : en relation avec des informations générées par l'utilisateur ;
- le **contenu interprété** : en relation avec des informations produites par l'utilisateur et interprétées par un professionnel ou un algorithme ;
- le **contenu affiché** : en relation avec la manière dont les informations sont éditées par le concepteur et comprises par l'utilisateur.

4.2. Utilisation des critères de qualité du contenu médical

Dans les sous-chapitres suivants, une sélection de critères a été effectuée à partir du référentiel HAS de 2016 (1) et après une relecture externe (processus de sélection décrit en annexe 7). La sélection finale des 17 critères retenus est présentée ci-après.

Chaque critère est justifié et accompagné d'exemples de documents ou d'éléments à produire.

Le nombre de critères adapté à l'évaluation de l'Application est lié à la destination d'usage de l'Application. Par exemple, pour une Application ne collectant pas de données, il faudra renseigner uniquement les critères liés au contenu initial et au contenu affiché. Les autres critères liés à la collecte de données étant inadaptés.

Cette sélection peut être mise en relation avec les **fonctions de base** d'une Application qui ont été décrites précédemment : informer, instruire, enregistrer des données, calculer/analyser, alerter/rappeler, afficher les données, guider et communiquer (cf. chapitre 2).

De plus, le niveau de risque concernant l'utilisation de l'Application pourrait être relié aux trois dimensions décrites précédemment :

- la **destination d'usage** de l'Application revendiquée par le fabricant (risques liés aux objectifs revendiqués de l'Application, risques liés à un retard de prise en charge, etc.) ;
- le **profil de l'utilisateur** (risques pour une population fragile, utilisateur novice, etc.) ;
- le **contexte d'utilisation** (risques dans le processus de collecte de données, l'environnement d'utilisation inadapté, etc.).

L'évaluation devrait être documentée de manière proportionnelle à l'importance du risque identifié.

4.3. Conception du contenu initial

Ces critères sont à appliquer pour toute production d'informations concernant la santé.

4.3.1. Gestion de l'information

Le niveau d'exigence attendu dépend du risque de proposer une information incomplète, erronée, ancienne, ou biaisée pour différentes raisons ou causes.

La qualité de la gestion de l'information se traduit par la fiabilité des informations publiées.

- Est-ce que l'information publiée est une information de référence clairement établie ou fait-elle l'objet de polémique ou de désaccord ?
- Quels sont les risques éventuels et leur niveau de gravité en cas de mauvaise gestion de l'information sur le sujet traité par l'Application ?

4.3.1.1. Organisation autour du service de l'information

Critère : présence d'une organisation ou d'un comité éditorial d'experts qui sélectionne et valide les informations publiées dans l'Application.

Justification : la rédaction et la gestion du contenu disponible dans l'Application s'appuient sur une organisation qui valide et garantit la qualité de l'information publiée.

Type d'éléments à produire : préciser la composition du comité qui doit être composé d'experts couvrant le domaine avec soit le CV correspondant soit la qualification de l'expert. Lister les liens vers les informations publiées. S'il s'agit de la reprise de recommandations de bonnes pratiques, lister les liens vers les documents sources.

4.3.1.2. Expertise des auteurs du contenu dans l'Application

Critère : des experts (professionnels de santé, ingénieurs, experts en algorithmes, organismes professionnels, associations de patients ou consommateurs, etc.) sont impliqués dans la production du contenu disponible dans l'Application.

Justification : le niveau d'expertise des auteurs du contenu de l'Application est un gage de qualité. La reconnaissance par ses pairs ou l'adossement à des organismes ou associations professionnelles améliore la crédibilité du contenu édité.

Type d'éléments à produire : lister où sont consultables en ligne, ou à un endroit accessible à tous, le nom des auteurs rédacteurs du contenu ainsi que leurs références ou qualifications. S'il s'agit de la reprise de recommandations de bonnes pratiques, lister uniquement les liens vers les documents sources.

4.3.1.3. Déclarations d'intérêts

Critère : les déclarations des liens d'intérêts des différents contributeurs sont consultables par tous.

Justification : les déclarations des liens d'intérêts éventuels sont un gage de transparence pour les utilisateurs et évaluateurs externes. Les liens d'intérêts peuvent entraîner des biais qui pourraient remettre en cause la fiabilité du produit.

Type d'éléments à produire : lister où sont consultables en ligne, ou à un endroit accessible à tous, les liens d'intérêts. S'il s'agit de la reprise de recommandations de bonnes pratiques, lister uniquement les liens vers les documents sources.

4.3.1.4. Citation des sources clés et références bibliographiques

Critère : les sources clés et références relatives à des publications argumentant le contenu de l'Application sont documentées et peuvent être consultables par tous.

Justification : dans le domaine de la santé, la citation des sources bibliographiques et d'une sélection objective des meilleures données disponibles est un gage de qualité requis. L'accès à la liste de références doit être consultable facilement.

Type d'éléments à produire : décrire comment sont présentées ces sources et références. La citation peut s'effectuer soit en intra-Application, soit sur un site web ressources, soit par une documentation externe, soit en référence en fin de contenu, etc. S'il s'agit de recommandations de bonnes pratiques, les liens vers les documents sources sont listés.

4.3.1.5. Actualisation des sources clés et références bibliographiques

Critère : le processus de veille et de mise à jour des sources clés et des références relatives à des publications est documenté.

Justification : la veille bibliographique permet de mettre à jour et d'adapter l'état des connaissances traitées par l'Application. La date de la mise à jour de l'information est à citer.

Type d'éléments à produire : décrire où la date de mise à jour de l'information est publiée et la manière dont l'actualisation est réalisée auprès de l'utilisateur.

4.3.1.6. Niveau de preuve

Critère : s'il existe une évaluation spécifique du produit et des niveaux de preuves, ces références appropriées sont consultables par tous.

Justification : la HAS a produit des guides d'analyse critique de la littérature, de gradation de niveaux de preuves ou de méthodologie d'évaluation^{115,116,117,118,119}.

Certaines Applications ont fait l'objet d'essai contrôlé randomisé (ECR) ou certains types d'Applications ont fait l'objet de revue systématique. Ces références sont majeures et doivent être accessibles pour justifier l'intérêt du produit.

Type d'éléments à produire : indiquer où sont présentés ces niveaux de preuve.

¹¹⁵ <http://www.has-sante.fr/portail/upload/docs/application/pdf/analiterat.pdf>

¹¹⁶ http://www.has-sante.fr/portail/upload/docs/application/pdf/2013-06/etat_des_lieux_niveau_preuve_gradation.pdf

¹¹⁷ http://www.has-sante.fr/portail/upload/docs/application/forcedownload/2016-03/guide_methodologique_analyse_critique.pdf

¹¹⁸ http://www.has-sante.fr/portail/upload/docs/application/pdf/eval_interventions_ameliorer_pratiques_guide.pdf

¹¹⁹ http://www.has-sante.fr/portail/upload/docs/application/pdf/2011-11/guide_methodo_vf.pdf

4.4. Contenu généré et standardisation

Si l'Application ne génère pas de données, les critères suivants ne sont pas applicables.

4.4.1. Pertinence, fiabilité des informations et support en vie réelle

Le niveau d'exigence attendu dépend du risque lié à une mauvaise collecte d'informations en situation réelle.

La qualité du suivi de l'utilisation de l'Application se traduit par la gestion des retours d'expériences et la maintenance du produit.

- Est-ce que la nature des données est suffisante pour remplir le ou les objectif(s) fixé(s) par l'Application ?
- La nature des données renvoie à des caractéristiques de précision, reproductibilité, granularité et fidélité.
- Quelle est la politique de suivi de l'Application en situation réelle ?
- Quels sont les risques éventuels et leur niveau de gravité en cas d'insuffisance de suivi ?

4.4.1.1. Qualité de la mesure dans le contexte d'utilisation

Critère : la qualité de la mesure (robustesse contextuelle) dans le milieu ou le contexte d'utilisation est documentée et justifiée en regard de la destination d'usage du produit.

Justification : la mesure effectuée en situation réelle peut différer par rapport aux mesures effectuées en laboratoire.

La mesure de données de santé ou de bien-être dans l'environnement de l'utilisateur doit être conforme à l'objectif de l'Application.

Type d'éléments à produire : fournir la documentation précisant quelle est la performance de la mesure en situation réelle. Des exemples concrets peuvent illustrer ce suivi.

4.4.1.2. Support utilisateur

Critère : un support est mis à disposition et permet de solliciter une demande d'assistance auprès des utilisateurs pour les demandes relatives à l'utilisation du produit (compréhension des contenus et utilisation des fonctionnalités). Les questions fréquemment posées sont documentées et actualisées.

Justification : un soutien à l'utilisation du produit permet d'améliorer la qualité d'utilisation. Ce soutien peut prendre différentes formes en fonction des objectifs du produit et de différents formats d'utilisation.

Type d'éléments à produire : lister et décrire les systèmes de support à l'utilisation proposés et adaptés pour l'utilisateur (hot-line, foire aux questions, fils de discussion modérés, forums utilisateurs avec charte d'utilisation, etc.).

4.5. Contenu interprété

Si l'Application ne génère pas de données, les critères suivants ne sont pas applicables.

Le niveau d'exigence attendu dépend du risque de proposer une information interprétée de manière erronée pour différentes causes.

La qualité de l'interprétation se traduit par la fiabilité des informations analysées et les capacités de les analyser.

- Est-ce que la politique d'interprétation est clairement définie ?
- Quels sont les risques éventuels et leur niveau de gravité en cas de mauvaise gestion de l'interprétation de l'information ciblée par l'Application ?

4.5.1. Typologie d'interprétation

Certaines Applications proposent une interprétation des contenus. L'interprétation peut être réalisée par un professionnel expert du sujet ou avec l'aide d'algorithme. Concernant l'utilisation d'intelligence artificielle (IA), Haverinen (113) a proposé en 2019 une liste de 20 questions à se poser lors de l'utilisation d'IA. En 2021, une régulation européenne est en développement sur le sujet de l'IA¹²⁰. Une évaluation spécifiquement adaptée sera prochainement en place en fonction de cette nouvelle régulation.

Dans l'état actuel des connaissances, il est nécessaire de s'assurer que ces outils puissent rediriger vers un professionnel de santé en cas de doute.

4.5.1.1. Interprétation humaine d'un contenu de santé

Critère : en cas d'interprétation humaine (non automatisée) de contenus à visée de santé (données de santé, contenu scientifique, etc.), celle-ci est assurée par des professionnels de santé adaptés à la thématique étudiée ou de personnes compétentes spécifiquement formées.

Justification : l'interprétation de contenu scientifique ou de données de santé nécessite l'implication de personnes qualifiées et compétentes.

Type d'éléments à produire : décrire qui, quand (en première ligne, en deuxième ligne, etc.) et comment est mise en place l'interprétation.

4.5.1.2. Interprétation automatisée d'un contenu de santé

Ce critère est amené à évoluer dans les prochaines années, car le secteur de la santé mobile et le niveau technologique actuel contribuent au développement des algorithmes et de l'IA (cf. supra).

Critère : en cas d'interprétation par des algorithmes ayant pour objet, par exemple, d'interpréter des contenus à portée médicale ou à visée de conseil, l'utilisateur doit être informé de l'intégration ou non d'IA dans le ou les algorithmes utilisés. L'exhaustivité des études évaluant le niveau de performance des algorithmes doit être à disposition de l'utilisateur.

Justification : l'interprétation de contenu scientifique ou de données de santé réalisée de manière automatique nécessite d'évaluer la fiabilité de l'interprétation. La crédibilité des tests des algorithmes est un élément critique à évaluer pour garantir cette fiabilité.

¹²⁰ https://eur-lex.europa.eu/procedure/FI/2021_106

Type d'éléments à produire : lister l'ensemble des publications justifiant l'utilisation des algorithmes au sein de l'Application. En absence de publications spécifiques, expliquer comment sont utilisés les algorithmes et comment sont réalisés les tests pour garantir la performance de l'interprétation.

4.6. Contenu affiché

Le niveau d'exigence attendu dépend du risque de proposer une information affichée qui pourrait être difficilement comprise par l'utilisateur.

4.6.1. Compréhension et implication des utilisateurs

Le niveau d'exigence attendu dépend du risque de ne pas répondre aux attentes des utilisateurs futurs de l'Application.

Les besoins, les tests utilisateurs ou l'évaluation du niveau de compréhension de l'usage de l'Application par les différents profils d'utilisateurs sont des éléments de qualité qui valorise le produit.

- Quelle a été l'implication des utilisateurs dans la conception et la validation de l'Application ?
- Quels sont les risques éventuels et leur niveau de gravité liés à une insuffisance d'implications des utilisateurs ?

4.6.1.1. Implication des utilisateurs (patients, professionnels, personnes spécifiques)

Critère : les principaux utilisateurs sont impliqués dans les différentes phases de développement de l'Application.

Justification : la conception avec les différentes parties prenantes, spécifiée de manière transparente est un gage de qualité.

Type d'éléments à produire : décrire la stratégie de sollicitation des utilisateurs dans le développement de l'Application. Des exemples concrets peuvent illustrer l'impact et l'intérêt de cette implication (grille d'analyse, écriture, *Living Lab*, etc.).

4.6.1.2. Description de la destination d'usage

Critère : la destination d'usage principale (objectifs ou finalités) du produit fait l'objet d'une description précise et consultable par tous.

Justification : cette déclaration est un élément important pour définir l'usage qui sera fait de l'Application.

Si l'usage déclaré par le fabricant est un instrument, appareil, équipement ou encore un logiciel destiné à être utilisé chez l'homme à des fins, notamment, de diagnostic, de prévention, de contrôle, de traitement, d'atténuation d'une maladie ou d'une blessure (Règlement [UE] 2017/745 du Parlement européen et du Conseil du 5 avril 2017 relatif aux dispositifs médicaux), il est éligible pour être un dispositif médical. Le fabricant devra se rapprocher de l'Agence nationale de sécurité du médicament et des produits de santé (ANSM) et se conformer aux dispositions légales et réglementaires applicables¹²¹.

Type d'éléments à produire : indiquer où est décrit l'objectif principal du produit. Préciser comment le concepteur explique sans ambiguïté ce que fait et ce que ne fait pas l'Application.

4.6.1.3. Contre-indications, risques, limitations d'usage

Critère : les contre-indications, les risques de mauvaise interprétation et les limitations d'usage sont évalués et documentés. Ces informations sont accessibles et consultables par les utilisateurs.

¹²¹ <http://ansm.sante.fr/Produits-de-sante/Dispositifs-medicaux>

Justification : l'Application peut présenter des limites d'utilisation ou de fiabilité.

Type d'éléments à produire : indiquer les limites d'utilisation ou de fiabilité que l'Application peut présenter. Préciser où et comment le concepteur publie cette information de manière transparente, compréhensible et adaptée à l'utilisateur.

4.6.1.4. Compréhension du contenu de santé

Critère : l'Application utilise une terminologie compréhensible par l'utilisateur.

Justification : des termes ambigus, une mauvaise traduction ou l'utilisation de jargon peuvent entraîner une mauvaise compréhension ou interprétation du contenu diffusé.

Type d'éléments à produire : indiquer la stratégie du concepteur pour rendre accessibles et compréhensibles les informations diffusées (relecture externe, test utilisateur, thésaurus, info-bulle, etc.).

4.6.2. Performance de l'affichage du contenu

L'utilisateur étant le plus souvent autonome lors de la lecture du contenu affiché, il est nécessaire de s'assurer de la bonne compréhension de l'information produite.

- Est-ce que l'affichage est suffisamment clair, compréhensible et adapté ?
- L'affichage des contenus renvoie à des questions délicates autour de l'ergonomie (ex. : l'information doit être visible entièrement sans être tronquée à l'écran, etc.), de l'utilisation de pictogrammes lorsqu'il s'agit par exemple d'alerter sur un niveau de sécurité d'une donnée, de la lecture à l'écran des blocs de texte, etc.
- Il s'agit aussi de prendre en compte le contexte d'utilisation et les utilisateurs afin d'anticiper leur perception de l'information et leur capacité à la comprendre avec ou sans aide à l'utilisation.
- Quels sont les risques éventuels et leurs niveaux de gravité en cas de mauvaise gestion de l'affichage de l'information par l'Application ?

4.6.2.1. Aide à l'utilisation/instructions

Critère : un système d'aide à l'utilisation du produit est mis à la disposition des utilisateurs (aide contextuelle, aide en ligne, manuel utilisateur, tutoriel, didacticiel, e-learning, etc.). Ce système favorise les capacités d'apprentissage de l'utilisateur.

Justification : le niveau de soutien didactique proposé par le concepteur à l'utilisateur permet d'optimiser l'utilisation du produit.

Type d'éléments à produire : lister les ressources utilisées pour aider à l'utilisation de l'Application.

4.6.2.2. Lisibilité texte et image et navigation

Critère : la navigation et la lisibilité des différents médias utilisés (texte, image, vidéos) ont été testées. L'interface permet le changement de la lisibilité de l'Application (modification de la taille/de la police de caractères, de la couleur des alertes, etc.).

Justification : la navigation et la lisibilité des informations par des utilisateurs ayant des capacités différentes sont un facteur d'accessibilité du produit.

Type d'éléments à produire : décrire la stratégie utilisée pour permettre d'optimiser la navigation et la lisibilité du texte et des images auprès de l'utilisateur.

4.6.2.3. Prévention des erreurs de compréhension de l'information

Critère : un système d'alerte adapté lors de décisions critiques suite à d'éventuelles erreurs de compréhension des informations par l'utilisateur est mis en place pour prévenir les risques ou réadresser vers un professionnel de santé.

Justification : certaines erreurs d'interprétations peuvent entraîner des erreurs de prise de décision de la part de l'utilisateur. Des éléments d'alertes sont à garantir par le concepteur ainsi que des explications ou descriptions des raisons de l'alerte produite.

Type d'éléments à produire : décrire la stratégie de prévention des erreurs d'interprétation de l'information utilisée auprès de l'utilisateur.

Table des annexes

Annexe 1.	Référencement dans l'Espace numérique de santé (ENS) et le Bouquet de services professionnels (BSP)	69
Annexe 2.	Norme d'information pour la qualité de la production d'informations sur la santé et les soins (NHS – UK)	71
Annexe 3.	Le <i>Mobile App Rating Scale</i> (MARS)	73
Annexe 4.	ORCHA-24 (<i>Organisation for the Review of Care and Health Applications</i>)	76
Annexe 5.	Exemple de « contenu clinique » détaillé pour évaluer les Applications de gestion de la douleur (Reynoldson, 2014 (95))	77
Annexe 6.	Détail des recommandations d'évaluation régionale AppSaludable (Espagne, région andalouse)	78
Annexe 7.	Processus de sélection des critères	80
Annexe 8.	Recherche documentaire	85

Table des figures

Figure 1. Évolution du nombre de publications dans le champ de la santé mobile de 2005 à 2020	16
---	----

Table des tableaux

Tableau 1 : Fonctions de base d'une Application et exemples de risques.....	24
Tableau 2 : Exemples de scores évaluant uniquement le contenu médical	31
Tableau 3 : Exemples de scores composites évaluant différents domaines de la santé mobile	36
Tableau 4 : Exemples de référentiels polyvalents pour évaluer la santé mobile	37
Tableau 5 : Exemples de sites proposant des évaluations d'Applications par des utilisateurs et/ou des experts.....	40
Tableau 6 : Exemples de structures proposant des labels, certification ou accréditation.....	45
Tableau 7 : Exemples de critères décrits dans le référentiel du NHS (Grande-Bretagne)	49
Tableau 8 : Exemples de structures proposant des répertoires éditorialisés régionaux ou nationaux.....	51
Tableau 9 : Exemples de structures proposant des répertoires éditorialisés spécialisés.....	53
Tableau 10 : Exemples de structures proposant des répertoires éditorialisés pour les dispositifs médicaux.....	56
Tableau 11 : Comparaison des domaines et sous-domaines repris entre 2016 et 2021 (avant relecture).....	80
Tableau 12 : Comparaison des critères repris entre 2016 et 2021 (avant relecture).....	81
Tableau 13 : Comparaison des critères repris après avis du groupe de lecture	83

Annexe 1. Référencement dans l'Espace numérique de santé (ENS)

L'espace numérique de santé, appelé « Mon espace santé », est prévu par la loi sur l'organisation et la transformation du système de santé de juillet 2019¹²².

Mon espace santé correspond à l'action 16 de la feuille de route pour le numérique en santé¹²³.

La Délégation ministérielle au numérique en santé souhaite que la HAS contribue à l'élaboration du cadre de référencement en définissant le socle minimal de critères de qualité du contenu médical. Un décret précisant les modalités de référencement ainsi que la commission de référencement sera publié d'ici la fin de l'année 2021.

Ces travaux sont menés en collaboration avec la Délégation du numérique en santé (DNS), la Caisse nationale de l'assurance maladie (CNAM) l'Agence du numérique en santé (ANS) et le GIE Sesam-Vitale avec une version pilote mise en œuvre dès l'été 2021 et un déploiement national à partir de janvier 2022.

Les services numériques éligibles sont les suivants :

- applications mobiles ;
- logiciels ;
- plateformes web.

Les services numériques éligibles s'adressent aux usagers (patients/citoyens) et sont édités par des acteurs publics ou privés, dans les secteurs sanitaires (santé et bien-être), médico-social et social. Ils doivent entrer dans le périmètre défini par l'article L 1111-13-1 du code de la santé publique. Ils peuvent être :

- reliés ou non à un ou plusieurs objets connectés ;
- gratuits ou payants ;
- des dispositifs médicaux ou non.

Le service candidat au référencement doit satisfaire à toutes les dispositions en vigueur, légales et réglementaires, nationales et communautaires. Cela concerne notamment les services relevant du règlement européen 2017/745 relatif aux dispositifs médicaux qui doivent être conformes aux exigences de sécurité et de performances qui leurs sont applicables. La conformité à cette réglementation est matérialisée par l'apposition du marquage CE dispositif médical.

Chaque service candidat au référencement est évalué par rapport :

- à la doctrine technique du numérique en santé (référentiels éthique, sécurité, interopérabilité) ;
- aux exigences liées aux critères de qualité du contenu médical ;
- aux critères relatifs à la protection des données personnelles ;
- aux exigences liées à l'interaction avec l'ENS.

La conformité à la doctrine technique du numérique en santé ainsi que la conformité aux exigences liées aux critères de la qualité du contenu et certaines exigences en matière de protection des données personnelles sont documentées *via* la plateforme Convergence¹²⁴.

¹²² <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000038821260/>

¹²³ <https://esante.gouv.fr/virage-numerique/feuille-de-route>

¹²⁴ <https://convergence.esante.gouv.fr>

Une commission de référencement est co-présidée par la DNS et la CNAM pour rendre un avis pour le référencement.

Une synthèse de l'évaluation de chaque service candidat au référencement est présentée à la commission.

Les avis de la commission peuvent revêtir deux formes :

- un avis défavorable au référencement, qui doit être motivé (exposé des motifs). L'éditeur est alors invité à corriger son service selon une procédure de traitement des non-conformités ;
- un avis favorable au référencement.

L'avis favorable peut être accompagné de commentaires à des fins d'amélioration du service.

La décision est ensuite prise par le ministre chargé de la santé. En cas de décision favorable au référencement une convention tripartite est établie entre l'éditeur, le ministre chargé de la santé et la CNAM, pour chaque service référencé.

La décision est communiquée à l'éditeur sur son espace privé au sein du portail éditeur.

Annexe 2. Norme d'information pour la qualité de la production d'informations sur la santé et les soins (NHS – UK)

Le NHS a proposé une liste de six principes et 16 items pour évaluer le processus de production d'informations. Ces principes sont conçus pour montrer que le fabricant a mis en place un processus pour produire et maintenir des informations de bonne qualité, fondées sur des preuves, sur la santé et les soins. Ces informations devraient être disponibles pour l'utilisateur.

Voilà une traduction réalisée à partir de la page web¹²⁵ (le questionnaire est rédigé à l'attention de l'organisation qui produit l'information).

Production d'informations

Ce principe est conçu pour démontrer que l'organisation a mis en place un processus défini pour produire des informations de bonne qualité sur la santé et les soins et de manière cohérente.

- Déclarations de qualité :
 - il existe un processus défini pour produire des informations (y compris l'identification du besoin de produire l'information, les étapes de vérification, l'approbation finale, la relecture, le contrôle de version et l'archivage) ;
 - toutes les personnes impliquées dans le processus de production d'information ont la formation/l'expérience à jour pertinente et suivent le processus défini pour tous les produits d'information.

Sources de preuves

Ce principe vise à garantir que lorsque des preuves sont utilisées, elles sont pertinentes et proviennent d'une source reconnue.

- Déclarations de qualité :
 - les informations sont créées à l'aide de preuves de haute qualité (là où elles existent) et sont présentées de manière objective et équilibrée. Lorsqu'il n'y a aucune preuve pour étayer les allégations faites pour produire l'information, cela est clairement indiqué à l'utilisateur final. Les preuves sont examinées et sont actualisées chaque fois que la ressource est mise à jour ;
 - les informations sont examinées par des professionnels/pairs concernés avant d'être approuvées pour utilisation.

Compréhension et implication des utilisateurs

Ce principe est conçu pour garantir, grâce à la participation des utilisateurs, que vous comprenez à qui les informations sont destinées, pourquoi elles sont nécessaires et quels sont les besoins des utilisateurs. Et pour s'assurer, par le biais de tests utilisateurs, qu'il reflète ces besoins et les opinions de ceux qui l'utilisent. La participation des utilisateurs, y compris les tests, doit être représentative du public cible et impliquer un nombre approprié de ces personnes.

- Déclarations de qualité :

¹²⁵ <https://www.england.nhs.uk/tis/about/the-info-standard/#information-production>

- l'information est créée en tenant compte des besoins en littératie en santé et/ou en accessibilité de la population à laquelle elle s'adresse ;
- le jargon n'est pas utilisé et les termes médicaux (lorsqu'ils sont utilisés) sont expliqués ;
- les utilisateurs finaux sont impliqués dès le début et tout au long de la production et leur contribution est activement utilisée.

Produit fini

Ce principe vise à garantir que le produit a été développé selon votre processus et qu'il est de haute qualité.

- Déclarations de qualité :
 - un ou plusieurs approbateurs autorisés vérifient que votre processus a été suffisamment suivi avant qu'un produit ne soit approuvé pour publication ;
 - la date et la date de révision de chaque produit d'information sont clairement indiquées ;
 - les informations sont rédigées dans un langage simple, exempt d'orthographe et de grammaire – et les termes médicaux sont expliqués si nécessaire ;
 - les références aux preuves utilisées dans les informations sont conservées et mises à disposition sur demande ;
 - des signaux informatifs indiquent à l'utilisateur final d'autres sources d'information ;
 - le produit d'information donne à l'utilisateur final des détails sur la manière dont il peut donner son avis.

Retour d'information

Ce principe vise à garantir que tous les retours (en dehors du processus de développement) sont traités de manière appropriée, en particulier en ce qui concerne les erreurs, les omissions ou les points à clarifier. Ces commentaires sont enregistrés, traités et résolus le cas échéant, en particulier si une modification ou un retrait d'un produit d'information est nécessaire.

- Déclarations de qualité :
 - les utilisateurs sont encouragés à donner des commentaires en continu après la publication du produit, et cela est pris en compte le cas échéant.

Mise à jour et révision

Ce principe vise à garantir que vos produits d'information sont examinés de façon planifiée et régulière, dans un délai approprié au type d'information, normalement pas plus de trois ans. Tous les produits qui ne sont pas examinés au cours de vos périodes d'examen définies ne doivent plus être distribués.

- Déclarations de qualité :
 - il existe un processus défini pour réviser des informations publiées/approuvées ;
 - tout le personnel impliqué dans le processus d'information suit le processus d'examen défini pour tous les produits d'information.

Annexe 3. Le *Mobile App Rating Scale (MARS)*

Le score de MARS a été produit par le *Queensland University of Technology* (121) avec comme auteur Stoyanov (75).

La traduction ci-dessous n'est pas validée en français.

Section A – Engageant (amusant, intéressant, personnalisable, interactif (envoi des alertes, messages, *reminder*, *feed-back*), permet le partage), bien ciblé sur l'audience.

Score sur 25 points

- Divertissant : est-ce que l'Application est divertissante à utiliser ? Est-ce qu'elle utilise des stratégies pour améliorer l'implication à travers l'aspect divertissant (par exemple : un côté ludique) (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Intérêt : est-ce que l'Application est intéressante à utiliser ? Est-ce qu'elle use de différentes stratégies pour améliorer l'implication au travers d'une présentation intéressante du contenu ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Personnalisable : est-ce qu'elle fournit/conservent tous les réglages nécessaires/les réglages de préférences des Applications (ex. : son, contenu, notifications, etc.) ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Interactivité : est-ce qu'elle permet à l'utilisateur d'entrer des données, fournir un *feed-back*, contenu rapide (*reminder*, options de partage, notifications, etc.) ? Note : ces fonctions ont besoin d'être personnalisables et de ne pas s'écraser les unes les autres. (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Groupe cible : est-ce que le contenu de l'Application (information visuelle, langage, design) est approprié pour l'audience ciblée ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)

Section B – Fonctionnalité – fonctionnement de l'Application, facile à apprendre, navigation, flux/parcours logique, design gestuel de l'Application

Score sur 20 points

- Performance : comment fonctionnent avec précision/vitesse les éléments de l'Application (fonctions) ainsi que ces composantes (boutons/menus) ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Facilité d'utilisation : avec quelle facilité est-il possible d'apprendre à utiliser l'Application ; quel est le niveau de clarté des icônes/étiquettes de menu et des instructions ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Navigation : est-ce que le déplacement entre les écrans est logique/précis/non interrompu ; est-ce que tous les liens vers les écrans sont présents ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Design gestuel : est-ce que les interactions (toucher/glisser/pincer/défiler) sont conformes et intuitives avec l'ensemble des composantes/écrans ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)

Section C – Esthétique – design des graphismes, attractivité visuelle, cohérence des couleurs, et style uniforme

Score sur 15 points

- Mise en page : est-ce que la disposition et la taille des boutons/icônes/menus/contenu de l'écran sont appropriées ou peut-on zoomer si nécessaire ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Graphisme : à quel niveau est la qualité de la résolution des graphismes utilisés pour les boutons/icônes/menus/contenu ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Attractivité visuelle : à quel niveau l'Application est-elle visuellement belle ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)

Section D – Information – contenu de haute qualité d'informations (ex. : texte, feed-back, mesures, références) provenant de source fiable. Sélectionner non adapté si la composante de l'Application ne convient pas.

Score sur 35 points

- Précision de la description de l'Application (sur le store) : est-ce que le contenu de l'Application est décrit ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Buts : est-ce que l'Application a des buts spécifiques, mesurables et atteignables ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Qualité de l'information : est-ce que l'Application présente un contenu correct, bien écrit, et adapté à l'objectif/au sujet visé par l'Application ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Quantité d'information : est-ce que l'étendue du domaine couvert est comprise dans le cadre de l'Application ; et compréhensible en restant concis ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Information visuelle : est-ce que l'explication visuelle des concepts – au travers de schéma/graphique/image/vidéos, etc. – est claire, logique, et correcte ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Crédibilité : est-ce que l'Application provient d'une source légitime (spécifiée dans la description sur le store ou dans l'Application elle-même) ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Fondée sur des preuves : est-ce que l'Application a été testée/évaluée par un essai contrôlé ; a dû être vérifiée par une étude fondée sur les preuves ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)

Score total de qualité : A+B+C+D

Partie subjective

Section E

Score sur 20 points

- Est-ce que vous recommanderiez cette Application à des personnes qui pourraient en tirer un bénéfice ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Combien de fois pensez-vous que vous pourriez utiliser cette Application dans les douze prochains mois, si elle vous semblait pertinente ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Paieriez-vous pour cette Application ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)
- Quelle est de manière globale la note que vous attribueriez à cette Application ? (5 niveaux de réponses cotés de 1 à 5 avec descriptif)

Annexe 4. ORCHA-24 (*Organisation for the Review of Care and Health Applications*)

Version courte du questionnaire ORCHA publiée par Leigh en 2018 (87)

Trois catégories sont recensées représentant 24 questions et un score maximum de 24. La partie concernant l'assurance qualité est détaillée ci-dessous.

- Gouvernance des données (8 questions)
- Efficacité clinique et assurance qualité (7 questions) :
 - Y a-t-il une déclaration dans l'Application elle-même ou dans le magasin en ligne concernant les commentaires des utilisateurs lors de la conception, du développement ou des tests ?
 - Y a-t-il une déclaration dans l'Application ou dans le magasin en ligne concernant l'implication des utilisateurs dans les tests ?
 - Y a-t-il une déclaration dans l'Application selon laquelle elle a été testée et s'est avérée bénéfique pour une personne atteinte de la maladie concernée ?
 - Y a-t-il une déclaration dans l'Application ou dans le magasin en ligne indiquant que l'Application a fait l'objet d'un essai clinique ou d'une autre forme de test pour montrer son efficacité dans le monde réel et a reçu des retours positifs ?
 - Existe-t-il une déclaration sur la fréquence à laquelle les conseils, directives ou contenus seront examinés pour garantir leur exactitude et leur pertinence clinique ?
 - Y a-t-il une déclaration dans l'Application selon laquelle elle a été positivement évaluée ou validée par un expert clinique ou autre expert pertinent ?
 - Y a-t-il des preuves dans l'Application que le développeur a tenté de valider des conseils ou des recommandations avec une expertise académique ?
 - Y a-t-il une déclaration dans l'Application identifiant une liste d'analyse externe, un établissement accrédité d'évaluation ou par des individus ?
- Expérience utilisateur et engagement (8 questions).

Annexe 5. Exemple de « contenu clinique » détaillé pour évaluer les Applications de gestion de la douleur (Reynoldson, 2014 (95))

Une bonne application de suivi de la douleur devrait inclure la possibilité de saisir les éléments suivants lorsque cela est approprié : (O/N)

- site : où est la douleur ?
- apparition des symptômes :
 - à quelle date la douleur est-elle apparue ?
 - à quelle heure la douleur est-elle apparue ?
- type de douleur : à quoi ressemble la douleur ? Dans quels schémas apparaît-elle ?
- localisation-rayonnement : est-ce qu'elle se déplace ailleurs ?
- symptômes associés : y a-t-il d'autres symptômes (ex. : nausées, vision trouble et engourdissement) ?
- durée : quelle a été la durée de l'épisode douloureux ?
- facteurs aggravants ou atténuants :
 - y a-t-il quelque chose qui augmente la douleur ?
 - y a-t-il quelque chose qui améliore la douleur ?
- environnement : lieu dans lequel la douleur est survenue (ex. : à la maison, au travail ou à l'extérieur) ;
- intensité : une échelle d'évaluation (ex. : échelle numérique de 1 à 10, 10 étant la pire douleur) ;
- médicaments ou traitements : une section de traitement distincte permettant aux utilisateurs de saisir des détails sur les médicaments pris ou les traitements subis :
 - l'Application comprend-elle une telle section ?
 - quel médicament a été pris/quel traitement a été suivi ?
 - quand est-ce arrivé ?
 - quelle posologie et quelle quantité ? (dose de médicament ; détails du traitement)
 - comment cela a-t-il affecté la douleur ? Cela a-t-il amélioré ou aggravé la situation, ou n'a-t-il eu aucun effet ?
- notes : possibilité d'ajouter des notes supplémentaires sur l'épisode (ex. : si quelque chose était important, mais non couvert par les questions).
- les options peuvent-elles être modifiées (ex. : des options peuvent être ajoutées ou supprimées pour rendre l'Application plus personnelle/pertinente) ? (O/N)
- l'Application reste-t-elle fidèle aux affirmations du développeur dans la description commerciale ?
- résultats :
 - y a-t-il une section résultats ? (O/N)
 - les résultats peuvent-ils être filtrés ? (O/N)
 - génère-t-il des rapports ? (O/N).

Annexe 6. Détail des recommandations d'évaluation régionale AppSaludable (Espagne, région andalouse)

Développé par la région Andalousie en 2012, AppSaludable¹²⁶ est une des premières plateformes mises en place en Europe pour recenser les Applications mobiles (112).

L'évaluation couvre quatre domaines et 31 recommandations et s'effectue en quatre phases (formulaire de soumission, auto-évaluation, évaluation externe et certification de la région). Le détail est cité en annexe 5.

Le détail des domaines et des recommandations est listé ci-dessous.

Conception et pertinence

- **Caractère approprié.** Recommandation 1. L'Application définit clairement sa portée fonctionnelle et son objectif, identifiant les groupes cibles d'information et les objectifs poursuivis à l'égard de ces groupes.
- **Accessibilité.** Recommandation 2. L'Application suit les principes de conception universelle, ainsi que les normes et recommandations d'accessibilité de référence.
- **Conception.** Recommandation 3. L'Application suit les recommandations, modèles et directives inclus dans les manuels officiels des différentes plateformes.
- **Tests d'utilisation.** Recommandation 4. L'Application a été testée par des utilisateurs potentiels avant d'être mise à la disposition du public.

Qualité et sécurité des informations

- **Adéquation au public.** Recommandation 5. L'Application s'adapte à son public cible.
- **Transparence.** Recommandation 6. L'Application offre des informations transparentes sur l'identité et la localisation de ses propriétaires. Recommandation 7. L'Application offre des informations sur ses sources de financement, sa promotion et son parrainage, ainsi que sur les conflits d'intérêts possibles.
- **Auteurs.** Recommandation 8. L'Application identifie les auteurs de son contenu et leurs qualifications professionnelles.
- **Mise à jour/révisions des informations.** Recommandation 9. L'Application comprend la date de la dernière révision effectuée dans le matériel publié. Recommandation 10. L'Application met en garde contre les mises à jour qui modifient ou influencent le fonctionnement du contenu lié à la santé, ainsi que d'autres données sensibles.
- **Contenu et sources d'information.** Recommandation 11. L'Application est basée sur une ou plusieurs sources d'information fiables et prend en compte les preuves scientifiques disponibles. Recommandation 12. L'Application offre des informations concises sur la procédure utilisée pour sélectionner son contenu. Recommandation 13. L'Application est basée sur des principes et des valeurs éthiques.
- **Gestion des risques.** Recommandation 14. Les risques pour la sécurité des patients liés à l'utilisation de l'Application sont identifiés. Recommandation 15. Les risques connus et les événements indésirables (quasi-accidents) sont analysés, et les actions appropriées commencent à être élaborées.

¹²⁶ <http://www.calidadappsalud.com/distintivo/catalogue>

Fourniture de services

- **Assistance technique/Demandes de renseignements.** Recommandation 16. L'Application dispose d'un système d'aide à son utilisation. Recommandation 17. L'Application offre un mécanisme de contact pour le support technique avec un temps de réponse garanti et fixe.
- **Commerce électronique.** Recommandation 18. L'Application informe sur les termes et conditions de commercialisation de ses produits et services.
- **Bande passante.** Recommandation 19. L'Application fait une utilisation efficace de la bande passante des communications.
- **Publicité.** Recommandation 20. L'Application met en garde contre l'utilisation des mécanismes de publicité et permet de les désactiver ou de les ignorer.

Confidentialité et vie privée

- **Confidentialité et protection des données.** Recommandation 21. Avant de télécharger et d'installer, l'Application informe sur le type de données de l'utilisateur à collecter et la raison, sur les politiques d'accès et le traitement des données, et sur d'éventuels accords commerciaux avec des tiers. Recommandation 22. L'Application décrit clairement les conditions générales relatives aux données personnelles enregistrées. Recommandation 23. Le fonctionnement de l'Application préserve la confidentialité des informations enregistrées, recueille les consentements exprès accordés par les utilisateurs et met en garde contre les risques liés à l'utilisation des Applications de santé mobiles en ligne. Recommandation 24. L'Application garantit des mesures de sécurité pertinentes lorsque des informations sur la santé des utilisateurs ou des données sensibles doivent être collectées ou échangées. Recommandation 25. L'Application informe les utilisateurs lorsqu'elle a accès à d'autres ressources de l'appareil, aux comptes des utilisateurs et aux profils sur les réseaux sociaux. Recommandation 26. L'Application garantit le droit d'accès aux informations enregistrées et aux mises à jour concernant les modifications de sa politique de confidentialité. Recommandation 27. L'Application a des mesures concernant la protection des mineurs conformément à la législation en vigueur.
- **Sécurité logique.** Recommandation 28. L'Application ne présente aucune sorte de vulnérabilité connue ni aucun type de code malveillant. Recommandation 29. L'Application décrit les procédures de sécurité mises en place pour éviter tout accès non autorisé aux données personnelles collectées, ainsi que pour limiter l'accès par des tiers. Recommandation 30. L'Application dispose de mécanismes de cryptage pour le stockage et l'échange d'informations, ainsi que de mécanismes de gestion des mots de passe. Recommandation 31. Lorsque l'Application utilise des services du Cloud (cloud computing), les termes et conditions de ces services sont déclarés et les mesures de sécurité pertinentes sont assurées.

Annexe 7. Processus de sélection des critères

Le référentiel HAS de 2016 (1) comprenait 101 critères répartis dans cinq domaines différents.

L'objet de ce travail était de ne sélectionner que les **critères de qualité pour évaluer le contenu de santé**.

Les domaines « informations utilisateurs », « contenant technique » et « sécurité/fiabilité » ne sont pas dans le champ de la demande de 2021, excepté un seul critère « contre-indications, risques, limitations d'usages » qui figurait dans le sous-domaine « fiabilité » et qui a été repris dans le domaine « contenu initial ».

Le domaine « contenu de santé » et une partie des critères du domaine « utilisation/usage » ont été repris avant l'envoi en relecture (tableau 11).

Tableau 11 : Comparaison des domaines et sous-domaines repris entre 2016 et 2021 (avant relecture)

Domaines et sous-domaines de 2016	Domaines de 2021 avant relecture
Informations utilisateurs : – description – consentement	Non repris par rapport au thème demandé. Ils font l'objet de critères techniques dédiés auprès de l'ENS.
Contenu de santé : – conception de contenu initial – standardisation – contenu généré – contenu interprété	– conception du contenu initial – contenu généré et standardisation – contenu interprété Les critères du sous-domaine « standardisation » ont été en partie repris. Les critères concernant globalement le format « d'interopérabilité » font l'objet de critères techniques dédiés auprès de l'ENS et ne sont pas repris dans ce travail.
Contenant technique : – conception technique – flux des données	Non repris par rapport au thème demandé. Ils font l'objet de critères techniques dédiés auprès de l'ENS.
Sécurité/Fiabilité : – cybersécurité – fiabilité – confidentialité	Non repris par rapport au thème demandé, excepté le critère « Contre-indications, risques, limitations d'usage » du sous-domaine fiabilité qui est conservé et intégré aux autres critères du contenu initial. Ils font l'objet de critères techniques dédiés auprès de l'ENS.
Utilisation/usage : – utilisation/design – acceptabilité – intégration/import	– contenu affiché (issu de « utilisation/design ») Pas de reprise des deux sous-domaines restants par rapport au thème demandé. Ils font l'objet de critères techniques dédiés auprès de l'ENS.

De manière plus détaillée, la sélection du domaine « contenu de santé » publié en 2016 a permis de reprendre l'ensemble des critères publiés sur les trois types de contenus (initial, généré, interprété). Le sous-domaine « standardisation » a été repris partiellement, car une grande partie des critères techniques (notamment d'interopérabilité) sont traités spécifiquement par ailleurs¹²⁷.

Le contenu peut également être mal interprété lors de son affichage ce qui a entraîné la création d'un domaine « contenu affiché » qui est une synthèse du sous-domaine de 2016 « utilisation/design » (tableau 12). Il regroupe les critères concernant la forme pour l'interprétation des contenus affichés.

¹²⁷ <https://convergence.esante.gouv.fr>

Au total, 28 critères sur les 101 critères ont été retenus avant envoi au groupe de lecture.

Tableau 12 : Comparaison des critères repris entre 2016 et 2021 (avant relecture)

Sous-domaines et critères de 2016	Domaines/sous-domaines et critères de 2021 (avant relecture)
<p>Conception du contenu initial :</p> <ul style="list-style-type: none"> - implications des utilisateurs (patients, professionnels, personne spécifique) - méthodologie d'ingénierie des besoins utilisateurs - organisation du service de l'information - expertise des auteurs du contenu - déclarations d'intérêts - citations des sources clés et références bibliographiques - actualisation des sources clés et références bibliographiques - niveau de preuve - description de la destination d'usage - langue du produit - thésaurus-glossaire 	<p>Contenu initial de santé/gestion de l'information :</p> <ul style="list-style-type: none"> - organisation du service de l'information - expertise des auteurs du contenu - déclarations d'intérêts - citations des sources clés et références bibliographiques - actualisation des sources clés et références bibliographiques - niveau de preuve - Contenu initial de santé/compréhension et implications des utilisateurs : - implications des utilisateurs (patients, professionnels, personne spécifique) - méthodologie d'ingénierie des besoins utilisateurs - description de la destination d'usage - contre-indications, risques, limitations d'usage - langue du produit - thésaurus-glossaire
<p>Standardisation :</p> <ul style="list-style-type: none"> - interopérabilité : standards de sémantique, terminologies de références - précision et reproductibilité des données - granularité des données - perte d'informations (par agrégation, par compression, etc.) - performance de la mesure dans le contexte d'utilisation - possibilité de synchronisation des données 	<p>Contenu généré et standardisation/fiabilité des données :</p> <ul style="list-style-type: none"> - précision et reproductibilité des données - perte d'informations (par agrégation, par compression, etc.) <p>Les autres critères ne font pas partie du thème traité. Ils font l'objet de critères techniques dédiés auprès de l'ENS.</p>
<p>Contenu généré :</p> <ul style="list-style-type: none"> - pertinence des données collectées - minimisation des données collectées - nombre d'interfaces/périphériques/applications - pertinence des informations dans le contexte - fils de discussions électroniques - assistances fonctionnelles, « hot-line » 	<p>Contenu généré et standardisation/pertinence des informations et support en vie réelle :</p> <ul style="list-style-type: none"> - performance de la mesure dans le contexte d'utilisation - pertinence des informations dans le contexte - fils de discussions électroniques - assistances fonctionnelles, « hot-line » <p>Les autres critères ne font pas partie du thème traité. Ils font l'objet de critères techniques dédiés auprès de l'ENS.</p>
<p>Contenu interprété :</p> <ul style="list-style-type: none"> - types d'algorithmes - interprétation humaine d'un contenu de santé - interprétation automatisée d'un contenu de santé 	<ul style="list-style-type: none"> - Contenu interprété/typologie d'interprétation : - types d'algorithmes - interprétation humaine d'un contenu de santé - interprétation automatisée d'un contenu de santé
<p>Utilisation/design :</p> <ul style="list-style-type: none"> - ergonomie - processus d'installation et configuration - aide à l'utilisation/instructions - convivialité et intuitivité - lisibilité texte et image - niveau d'utilisation 	<p>Contenu affiché/performance de l'affichage du contenu :</p> <ul style="list-style-type: none"> - ergonomie - aide à l'utilisation/instructions - convivialité et intuitivité - lisibilité texte et image - Contenu affiché/adaptation du contenu à l'utilisateur : - niveau d'utilisation - prévention des erreurs

- accessibilité du contenu pour des personnes en situation de handicap
- facilité d'emploi
- prévention des erreurs
- cas d'usage, scénarios métier
- flexibilité/customisation
- délais de réponse, temps d'affichage

- cas d'usage, scénarios métier

Les autres critères ne font pas partie du thème traité. Ils font l'objet de critères techniques dédiés auprès de l'ENS.

Le groupe de lecture a reçu les 28 critères et le document reprenant l'état des lieux de la littérature. Une évaluation sur le fond, la forme et les éléments de justification a été demandée (tableau 13).

Concernant le contenu initial, les commentaires de fond ont porté sur les problèmes liés autour du secret industriel pour diffuser les informations des tests réalisés auprès des utilisateurs (ingénierie des besoins) et des difficultés à avoir des éléments objectifs d'évaluation sur ce champ.

Concernant la langue utilisée pour le produit ou les outils comme le thésaurus, les retours ont précisé que le niveau d'exigence et d'utilité de ces critères était dépendant du type d'Application. Ils ont été fusionnés et remplacés par « compréhension du contenu de santé ».

Concernant le contenu généré, les critères concernant la fidélité et la reproductibilité des mesures ont été jugés utiles, mais, pour certaines Applications complexes lorsqu'il n'y avait pas d'étalon-or par exemple ou suivant les profils de patients. Ces critères ont été regroupés en un seul critère « performance de la mesure dans le contexte d'utilisation ».

Concernant le critère « hot-line » ou les fils de discussion, les retours ont souligné le coût de ces services et le fait que ces services puissent ne pas être adaptés pour certaines Applications. Un critère « support utilisateur » permet de proposer des solutions adaptées sans en citer de spécifiques comme critère.

Concernant le contenu interprété, le critère concernant l'information sur le « type » d'algorithme utilisé a été jugé peu pertinent, car trop complexe à comprendre par l'utilisateur.

Concernant le contenu affiché, les retours ont souligné que les critères proposés étaient soit éloignés de la thématique « contenu de santé », soit déjà redondants avec le critère « implication des utilisateurs » qui évalue l'interface, soit complexe à objectiver de manière pragmatique. Ils ont été regroupés en trois critères : l'aide à l'utilisation, la lisibilité/compréhension de tous les contenus et un système d'alerte pour des affichages entraînant des erreurs potentielles.

Tableau 13 : Comparaison des critères repris après avis du groupe de lecture

Avant relecture (28 critères)	Après avis du groupe de lecture (17 critères)
<p>Contenu initial de santé/gestion de l'information :</p> <ul style="list-style-type: none"> - organisation du service de l'information - expertise des auteurs du contenu - déclarations d'intérêts - citations des sources clés et références bibliographiques - actualisation des sources clés et références bibliographiques - niveau de preuve <p>Contenu initial de santé/compréhension et implications des utilisateurs :</p> <ul style="list-style-type: none"> - implications des utilisateurs (patients, professionnels, personne spécifique) - méthodologie d'ingénierie des besoins utilisateurs - description de la destination d'usage - contre-indications, risques potentiels, limitations d'usage - langue du produit - thésaurus-glossaire 	<ul style="list-style-type: none"> - Contenu initial de santé/gestion de l'information : - organisation autour du service de l'information - expertise des auteurs du contenu dans l'Application - déclarations d'intérêts - citations des sources clés et références bibliographiques - actualisation des sources clés et références bibliographiques - niveau de preuve <p>La partie « contenu initial de santé/compréhension et implications des utilisateurs » est déplacée au niveau du contenu affiché. Car l'implication des utilisateurs concerne tous les types de contenus</p>
<p>Contenu généré et standardisation/fiabilité des données :</p> <ul style="list-style-type: none"> - précision et reproductibilité des données - perte d'informations (par agrégation, par compression, etc.) 	
<p>Contenu généré et standardisation/pertinence des informations et support en vie réelle :</p> <ul style="list-style-type: none"> - performance de la mesure dans le contexte d'utilisation - pertinence des informations dans le contexte - fils de discussions électroniques - assistances fonctionnelles, « hot-line » 	<p>Contenu généré et standardisation/pertinence, fiabilité des informations et support en vie réelle :</p> <ul style="list-style-type: none"> - qualité de la mesure dans le contexte d'utilisation - support utilisateur
<p>Contenu interprété/typologie d'interprétation :</p> <ul style="list-style-type: none"> - types d'algorithmes - interprétation humaine d'un contenu de santé - interprétation automatisée d'un contenu de santé 	<p>Contenu interprété/typologie d'interprétation :</p> <ul style="list-style-type: none"> - interprétation humaine d'un contenu de santé - interprétation automatisée d'un contenu de santé
<p>Contenu affiché/performance de l'affichage du contenu :</p> <ul style="list-style-type: none"> - ergonomie - aide à l'utilisation/instructions - convivialité et intuitivité - lisibilité texte et image <p>Contenu affiché/adaptation du contenu à l'utilisateur :</p> <ul style="list-style-type: none"> - niveau d'utilisation - prévention des erreurs - cas d'usage, scénarios métier 	<p>Contenu initial de santé/compréhension et implications des utilisateurs :</p> <ul style="list-style-type: none"> - implications des utilisateurs (patients, professionnels, personne spécifique) - description de la destination d'usage - contre-indications, risques potentiels, limitations d'usage - compréhension du contenu de santé <p>Contenu affiché/performance de l'affichage du contenu :</p> <ul style="list-style-type: none"> - aide à l'utilisation/instructions - lisibilité texte et image et navigation - prévention des erreurs

Au total, 17 critères ont été retenus ou sont issus de fusion des 28 critères adressés au groupe de lecture (tableau 13).

Sur la forme, de petites modifications de rédaction ont été réalisées pour améliorer la description de ce qui est attendu comme critère.

Concernant les éléments de preuve à fournir pour pouvoir évaluer le critère, des propositions ont été effectuées pour rendre l'évaluation plus réaliste et plus compréhensible par le concepteur. Une certaine latitude étant proposée au concepteur en fonction du type d'Applications à évaluer.

En fonction des Applications et du niveau de risque qu'elles peuvent présenter, le niveau de précision de ces éléments de preuve nécessite une évaluation par des professionnels de santé ou des personnes compétentes pour juger de leurs pertinences.

Annexe 8. Recherche documentaire

La recherche documentaire, limitée aux publications en langue anglaise et française, a été réalisée sur la base Medline. La stratégie d'interrogation des bases de données précises pour chaque question et/ou type d'étude, les termes de recherche utilisés, les opérateurs booléens et la période de recherche.

Les termes de recherche utilisés sont soit des termes issus de thésaurus (descripteurs), soit des termes libres (du titre ou du résumé). Ils sont combinés avec les termes décrivant les types d'études.

Le tableau 1 présente de façon synthétique les étapes de cette interrogation dans la base de données Medline. Le nombre total de références obtenues par interrogation de cette base de données bibliographiques est 2 073.

Tableau 1 : Stratégie de recherche dans la base de données Medline

Type d'étude/sujet	Termes utilisés	Période
Recommandations & conférences de consensus		01/2015 – 12/2020
Étape 1	Mobile Applications/de OR (mobile application OR mobile applications OR mobile app OR mobile apps OR smartphone application OR smartphone applications OR Smartphone app OR Smartphone apps OR app stores OR Mobile Medical Application OR Mobile Medical Applications OR medical apps OR medical app OR standalone software OR health apps OR health app OR mhealth OR mobile health OR ehealth OR apps OR app)/ti OR ((Medical Informatics Applications/de OR Software/majr OR application/ti OR medical/ti) AND (mobile OR smartphone OR phone)/ti)	
ET		
Étape 2	(recommendation* OR guideline* OR statement* OR consensus OR position paper)/ti OR health planning guidelines/de OR (practice guideline OR guideline OR Consensus Development Conference OR Consensus Development Conference, NIH)/pt	
Méta-analyses & revues systématiques		01/2015 – 12/2020
Étape 1		
ET		
Étape 3	(metaanalys* OR meta-analys* OR meta-analysis OR systematic review* OR systematic overview* OR systematic literature review* OR systematical review* OR systematical overview* OR systematical literature review* OR systematic literature search OR pooled analysis)/ti OR (meta-analysis OR "Systematic Review")/pt OR cochrane database syst rev/sd	

Évaluation des applications		01/2015 12/2020	–
Étape 1			
ET			
Étape 4	(framework OR frameworks OR certificat* OR label* OR standard OR criteria OR scoring OR quality assessment)/ti OR (Mobile App Rating scale OR MARS)/ti,ab OR Software Validation/de		
OU			
Étape 5	(mhealth OR mobile health OR apps OR app OR standalone software)/ti AND (evaluat* OR Assessment)/ti		
Échelles et applications		01/2015 01/2021	–
Étape 1			
ET			
Étape 6	(scale OR reliabilit* OR validit* OR assessment)/ti OR (scale development) OR (app ability)/ab		

* : troncature ; de : descriptor ; ti : title ; ab : abstract ; pt : publication type ; so : journal title

Résultats

Nombre de références identifiées : 2073

Nombre de références analysées : 483

Nombre de références retenues : 121

Références bibliographiques

1. Haute Autorité de Santé. Référentiel de bonnes pratiques sur les applications et les objets connectés en santé (mobile Health ou mHealth). Saint-Denis La Plaine: HAS; 2016. https://www.has-sante.fr/upload/docs/application/pdf/2016-11/has_ref_apps_oc.pdf
2. Pagliari C, Sloan D, Gregor P, Sullivan F, Detmer D, Kahan JP, et al. What is eHealth (4): a scoping exercise to map the field. *J Med Internet Res* 2005;7(1):e9. <http://dx.doi.org/10.2196/jmir.7.1.e9>
3. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010;12(1):e4. <http://dx.doi.org/10.2196/jmir.1376>
4. Salmani H, Ahmadi M, Shahrokhi N. The impact of mobile health on cancer screening: A systematic review. *Cancer Inform* 2020;19:1-11. <http://dx.doi.org/10.1177/1176935120954191>
5. Leppin AL, Montori VM, Gionfriddo MR. Minimally disruptive medicine: a pragmatically comprehensive model for delivering care to patients with multiple chronic conditions. *Healthcare* 2015;3(1):50-63. <http://dx.doi.org/10.3390/healthcare3010050>
6. Wallace D, Perry J, Yu J, Mehta J, Hunter P, Cross KM. Assessing the Need for Mobile Health (mHealth) in Monitoring the Diabetic Lower Extremity. *JMIR Mhealth Uhealth* 2019;7(4):e11879. <http://dx.doi.org/10.2196/11879>
7. Scott K, Richards D, Adhikar R. A review and comparative analysis of security risks and safety measures of mobile health apps. *Australas J Inf Syst* 2015;19:1-17. <http://dx.doi.org/10.3127/ajis.v19i0.1210>
8. 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. <http://dx.doi.org/10.1016/j.amepre.2013.03.017>
9. IQVIA Institute for Human Data Science. The Growing value of digital health. Parsippany: IQVIA Institute; 2017. <https://regresearchnetwork.org/wp-content/uploads/the-growing-value-of-digital-health.pdf>
10. Fagherazzi G, Ravaud P. Digital diabetes: Perspectives for diabetes prevention, management and research. *Diabetes Metab* 2019;45(4):322-9. <http://dx.doi.org/10.1016/j.diabet.2018.08.012>
11. PricewaterhouseCoopers. Socio-economic impact of mHealth. An assessment report for the European Union. Ahmedabad: PwC; 2013. http://www.gsma.com/connectedliving/wp-content/uploads/2013/06/Socio-economic_impact-of-mHealth_EU_14062013V2.pdf
12. Powell A, Torous J. A patient-centered framework for measuring the economic value of the clinical benefits of digital health apps: Theoretical modeling. *JMIR Ment Health* 2020;7(10):e18812. <http://dx.doi.org/10.2196/18812>
13. Wicks P, Chiauzzi E. 'Trust but verify'--five approaches to ensure safe medical apps. *BMC Med* 2015;13:205. <http://dx.doi.org/10.1186/s12916-015-0451-z>
14. Dick S, O'Connor Y, Thompson MJ, O'Donoghue J, Hardy V, Wu TJ, et al. Considerations for Improved Mobile Health Evaluation: Retrospective Qualitative Investigation. *JMIR Mhealth Uhealth* 2020;8(1):e12424. <http://dx.doi.org/10.2196/12424>
15. Albrecht UV. Créer la transparence et fournir une orientation. Méthodes et outils d'aide décisionnelle pour l'utilisation des applications mobiles de santé. Bern: eHealth Suisse; 2019. https://mhh-publikationsserver.gbv.de/receive/mhh_mods_00000003
16. Albrecht UV, Hillebrand U, von Jan U. Relevance of trust marks and CE labels in German-language store descriptions of health apps: Analysis. *JMIR Mhealth Uhealth* 2018;6(4):e10394. <http://dx.doi.org/10.2196/10394>
17. World Health Organization. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. Geneva: WHO; 2016. <https://apps.who.int/iris/bitstream/handle/10665/252183/9789241511766-eng.pdf>
18. World Health Organization. WHO guideline: recommendations on digital interventions for health system strengthening. Geneva: WHO; 2019.
19. Pham Q, Wiljer D, Cafazzo JA. Beyond the randomized controlled trial: a review of alternatives in mhealth clinical trial methods. *JMIR Mhealth Uhealth* 2016;4(3):e107. <http://dx.doi.org/10.2196/mhealth.5720>
20. Byambasuren O, Sanders S, Beller E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. *NPJ Digit Med* 2018;1:12. <http://dx.doi.org/10.1038/s41746-018-0021-9>
21. 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. <http://dx.doi.org/10.1016/j.amepre.2013.06.006>
22. Grundy QH, Wang Z, Bero LA. Challenges in assessing mobile health app quality: A systematic review of prevalent and innovative methods. *Am J Prev Med* 2016;51(6):1051-9. <http://dx.doi.org/10.1016/j.amepre.2016.07.009>
23. Ologeanu-Taddei R. Assessment of mhealth interventions: Need for new studies, methods, and guidelines for study designs. *JMIR Med Inform* 2020;8(11):e21874. <http://dx.doi.org/10.2196/21874>
24. Huckvale K, Morrison C, Ouyang J, Ghaghda A, Car J. The evolution of mobile apps for asthma: an updated systematic assessment of content and tools. *BMC Med* 2015;13:58. <http://dx.doi.org/10.1186/s12916-015-0303-x>

25. Asthana S, Jones R, Sheaff R. Why does the NHS struggle to adopt eHealth innovations? A review of macro, meso and micro factors. *BMC Health Serv Res* 2019;19(1):984.
<http://dx.doi.org/10.1186/s12913-019-4790-x>
26. Wierda E, Blok S, Somsen GA, van der Velde ET, Tulevski II, Stavrov B, et al. Protecting patient privacy in digital health technology: the Dutch m-Health infrastructure of Hartwacht as a learning case. *BMJ Innov* 2020;6:170-6.
27. Zelmer J, van Hoof K, Notarianni M, van Mierlo T, Schellenberg M, Tannenbaum C. An assessment framework for e-mental health apps in canada: Results of a modified delphi process. *JMIR Mhealth Uhealth* 2018;6(7):e10016.
<http://dx.doi.org/10.2196/10016>
28. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. *BMC Public Health* 2016;16(1):1158.
<http://dx.doi.org/10.1186/s12889-016-3808-0>
29. Pustozarov E, von Jan U, Albrecht UV. Evaluation of mHealth applications quality based on user ratings. *Stud Health Technol Inform* 2016;226:237-40.
30. Singh K, Drouin K, Newmark LP, Lee J, Faxvaag A, Rozenblum R, et al. Many mobile health apps target high-need, high-cost populations, but gaps remain. *Health Aff* 2016;35(12):2310-8.
<http://dx.doi.org/10.1377/hlthaff.2016.0578>
31. Ali EE, Chew L, Yap KY-L. Evolution and current status of mhealth research: a systematic review. *BMJ Innov* 2016;2:33-40.
32. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. *J Med Internet Res* 2013;15(5):e95.
<http://dx.doi.org/10.2196/jmir.2430>
33. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. *JMIR Mhealth Uhealth* 2018;6(1):e23.
<http://dx.doi.org/10.2196/mhealth.8873>
34. Grainger R, Devan H, Sangelaji B, Hay-Smith J. Issues in reporting of systematic review methods in health app-focused reviews: A scoping review. *Health Informatics J* 2020;1460458220952917.
<http://dx.doi.org/10.1177/1460458220952917>
35. World Health Organization. Rapid risk assessment of acute public health events. Geneva: WHO; 2012.
https://apps.who.int/iris/bitstream/handle/10665/70810/WHO_HSE_GAR_ARO_2012.1_eng.pdf?sequence=1
36. Lewis TL, Wyatt JC. mHealth and mobile medical Apps: a framework to assess risk and promote safer use. *J Med Internet Res* 2014;16(9):e210.
<http://dx.doi.org/10.2196/jmir.3133>
37. Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, et al. mHealth Assessment: Conceptualization of a Global Framework. *JMIR Mhealth Uhealth* 2017;5(5):e60.
<http://dx.doi.org/10.2196/mhealth.7291>
38. Rasmussen J. Risk management in a dynamic society: a modelling problem. *Saf Sci* 1997;27:183-213.
39. Neumann K, Larisch K, Dietzel J, Kurepkat M, Wenzlau V. Digital healthcare products. Leveraging opportunities – developing safe routes to market. Berlin: IGES Institute; 2016.
https://www.iges.com/e6/e1621/e10211/e13470/e15278/e15279/e15281/attr_objjs15282/IGES_Publication_Digital_health_care_products_WEB_eng.pdf
40. Therapeutic Goods Administration. Actual and potential harm caused by medical software. A rapid literature review of safety and performance issues. Woden: TGA; 2020.
<https://www.tga.gov.au/sites/default/files/actual-and-potential-harm-caused-medical-software.pdf>
41. Huckvale K, Prieto JT, Tilney M, Benghozi PJ, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Med* 2015;13:214.
<http://dx.doi.org/10.1186/s12916-015-0444-y>
42. Kearney AT. Mobile Health: Fata Morgana oder Wachstumstreiber? Düsseldorf: AT Kearney GmbH; 2013.
43. Benjumea J, Ropero J, Rivera-Romero O, Dorrnoro-Zubiete E, Carrasco A. Privacy Assessment in Mobile Health Apps: Scoping Review. *JMIR Mhealth Uhealth* 2020;8(7):e18868.
<http://dx.doi.org/10.2196/18868>
44. Timmers T, Janssen L, Kool RB, Kremer JA. Educating patients by providing timely information using smartphone and tablet apps: Systematic review. *J Med Internet Res* 2020;22(4):e17342.
<http://dx.doi.org/10.2196/17342>
45. Schmidt-Kraepelin M, Toussaint PA, Thiebes S, Hamari J, Sunyaev A. Archetypes of gamification: Analysis of mHealth Apps. *JMIR Mhealth Uhealth* 2020;8(10):e19280.
46. Thimbleby H, Oladimeji P, Cairns P. Unreliable numbers: error and harm induced by bad design can be reduced by better design. *Journal of the Royal Society, Interface* 2015;12(110):0685.
<http://dx.doi.org/10.1098/rsif.2015.0685>
47. Council of Europe. Guidelines on artificial intelligence and data protection. Strasbourg: CE; 2019.
<https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8>
48. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. *BMC Med* 2012;10:144.
<http://dx.doi.org/10.1186/1741-7015-10-144>
49. Huckvale K, Adomaviciute S, Prieto JT, Leow MK, Car J. Smartphone apps for calculating insulin dose: a systematic assessment. *BMC Med* 2015;13:106.
<http://dx.doi.org/10.1186/s12916-015-0314-7>
50. Bierbrier R, Lo V, Wu RC. Evaluation of the accuracy of smartphone medical calculation apps. *J Med Internet Res* 2014;16(2):e32.
<http://dx.doi.org/10.2196/jmir.3062>
51. Suso-Ribera C, Castilla D, Zaragoza I, Mesas Á, Server A, Medel J, et al. Telemonitoring in chronic pain management using smartphone apps: A randomized controlled trial comparing usual assessment against app-based monitoring with and without clinical alarms. *Int J Environ Res Public Health* 2020;17(18).
<http://dx.doi.org/10.3390/ijerph17186568>

52. Lewis TL, Boissaud-Cooke MA, Aungst TD, Eysenbach G. Consensus on use of the term "App" versus "Application" for reporting of mHealth research. *J Med Internet Res* 2014;16(7):e174; discussion e. <http://dx.doi.org/10.2196/jmir.3460>
53. Riezebos RJ. Peer-reviewing of mHealth applications. Requirements for peer-reviewing mobile health applications and development of an online peer review tool Amsterdam: University of Amsterdam; 2014. <http://dare.uva.nl/cgi/arno/show.cgi?fid=573074>
54. Eysenbach G. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011;13(4):e126. <http://dx.doi.org/10.2196/jmir.1923>
55. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. *BMJ* 2016;352:i1174. <http://dx.doi.org/10.1136/bmj.i1174>
56. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Transl Behav Med* 2014;4(4):363-71. <http://dx.doi.org/10.1007/s13142-014-0293-9>
57. Guidance on evaluating or developing a health app. Wellington: Ministry of Health Ministry of Health; 2017. <https://www.health.govt.nz/system/files/documents/pages/guidance-evaluating-developing-health-app-oct17-v2.pdf>
58. Nielsen SL, Rimpiläinen S. T1.2.1 Report on international practice on digital apps. T1 Technology : Digital Health & Care Institute; 2018.
59. Institut für Gesundheitsförderung und Prävention GmbH, Kraus-Füederer H. Gesundheits Apps. Grundlagenpapier unter besonderer Berücksichtigung des Aspekts Gesundheitskompetenz. Graz: IFGP ; 2018.
60. Nouri R, S RNK, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. *Journal of the American Medical Informatics Association : JAMIA* 2018;25(8):1089-98. <http://dx.doi.org/10.1093/jamia/ocv050>
61. Urmet H. Mapping the nice evidence standards framework and international experiences into a guideline for the assessment of mobile health solutions in Estonia [Thesis] Tallinn: University of technology School of Information Technologies.
62. European mHealth hub. D2.1 knowledge tool 1. health apps assessment frameworks (Pending EC approval). Seville: mHealth-hub; 2020. <https://mhealth-hub.org/download/d2-1-knowledge-tool-1-health-apps-assessment-frameworks-pending-ec-approval>
63. Llorens-Vernet P, Miró J. The Mobile App Development and Assessment Guide (MAG): Delphi-Based Validity Study. *JMIR Mhealth Uhealth* 2020;8(7):e17760. <http://dx.doi.org/10.2196/17760>
64. Levine DM, Co Z, Newmark LP, Groisser AR, Holmgren AJ, Haas JS, et al. Design and testing of a mobile health application rating tool. *NPJ Digit Med* 2020;3:74. <http://dx.doi.org/10.1038/s41746-020-0268-9>
65. Albrecht UV. Chances and risks of mobile health apps (CHARISMHA). Hannover: Hannover Medical School; 2016. https://publikationsserver.tu-braunschweig.de/receive/dbbs_mods_00060024
66. Azad-Khaneghah P, Neubauer N, Miguel Cruz A, Liu L. Mobile health app usability and quality rating scales: a systematic review. *Disabil Rehabil Assist Technol* 2020;1-10. <http://dx.doi.org/10.1080/17483107.2019.1701103>
67. Haute Autorité de Santé. Évaluation de la qualité des sites e-santé et de la qualité de l'information de santé diffusée sur Internet (Revue de la littérature des outils d'évaluation). Saint-Denis La Plaine: HAS; 2007. https://www.has-sante.fr/upload/docs/application/pdf/evaluation_qualite_site_sante_internet.pdf
68. Van Singer M, Chatton A, Khazaal Y. Quality of smartphone apps related to panic disorder. *Front Psychiatry* 2015;6:96. <http://dx.doi.org/10.3389/fpsy.2015.00096>
69. Butcher R, MacKinnon M, Gadd K, LeBlanc-Duchin D. Development and examination of a rubric for evaluating point-of-care medical applications for mobile devices. *Med Ref Serv Q* 2015;34(1):75-87. <http://dx.doi.org/10.1080/02763869.2015.986794>
70. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewer--Let the reader and viewer beware. *JAMA* 1997;277(15):1244-5.
71. Zhang MW, Ho RC, Loh A, Wing T, Wynne O, Chan SWC, et al. Current status of postnatal depression smartphone applications available on application stores: an information quality analysis. *BMJ Open* 2017;7(11):e015655. <http://dx.doi.org/10.1136/bmjopen-2016-015655>
72. Jeon E, Park HA, Min YH, Kim HY. Analysis of the information quality of korean obesity-management smartphone applications. *Healthc Inform Res* 2014;20(1):23-9. <http://dx.doi.org/10.4258/hir.2014.20.1.23>
73. Khazaal Y, Chatton A, Cochand S, Coquard O, Fernandez S, Khan R, et al. Brief DISCERN, six questions for the evaluation of evidence-based content of health-related websites. *Patient Educ Couns* 2009;77(1):33-7. <http://dx.doi.org/10.1016/j.pec.2009.02.016>
74. Guan Z, Sun L, Xiao Q, Wang Y. Constructing an assessment framework for the quality of asthma smartphone applications. *BMC Med Inform Decis Mak* 2019;19(1):192. <http://dx.doi.org/10.1186/s12911-019-0923-8>
75. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015;3(1):e27. <http://dx.doi.org/10.2196/mhealth.3422>
76. Domnich A, Arata L, Amicizia D, Signori A, Patrick B, Stoyanov S, et al. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. *BMC Med Inform Decis Mak* 2016;16:83. <http://dx.doi.org/10.1186/s12911-016-0323-2>
77. Martín Payo R, Fernández Álvarez MM, Blanco Díaz M, Cuesta Izquierdo M, Stoyanov SR, Llana Suárez E. Spanish adaptation and validation of the Mobile Application Rating Scale questionnaire. *Int J Med Inform* 2019;129:95-9.

<http://dx.doi.org/10.1016/j.ijmedinf.2019.06.005>

78. Messner EM, Terhorst Y, Barke A, Baumeister H, Stoyanov S, Hides L, et al. The German version of the mobile app rating scale (MARS-G): Development and validation study. *JMIR Mhealth Uhealth* 2020;8(3):e14479. <http://dx.doi.org/10.2196/14479>

79. Bardus M, Awada N, Ghandour LA, Fares EJ, Gherbal T, Al-Zanati T, et al. The arabic version of the mobile app rating scale: development and validation study. *JMIR Mhealth Uhealth* 2020;8(3):e16956. <http://dx.doi.org/10.2196/16956>

80. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). *PloS one* 2020;15(11):e0241480. <http://dx.doi.org/10.1371/journal.pone.0241480>

81. Terhorst Y, Messner EM, Schultchen D, Paganini S, Portenhausner A, Eder AS, et al. Systematic evaluation of content and quality of English and German pain apps in European app stores. *Internet Interv* 2021;24:100376. <http://dx.doi.org/10.1016/j.invent.2021.100376>

82. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: Systematic search and evaluation using the mobile app rating scale. *JMIR Mhealth Uhealth* 2018;6(10):e10718. <http://dx.doi.org/10.2196/10718>

83. Dawson RM, Felder TM, Donevant SB, McDonnell KK, Card EB, 3rd, King CC, et al. What makes a good health 'app'? Identifying the strengths and limitations of existing mobile application evaluation tools. *Nurs Inq* 2020;27(2):e12333. <http://dx.doi.org/10.1111/nin.12333>

84. Stach M, Kraft R, Probst T, Messner EM, Terhorst Y, Baumeister H, et al. Mobile health app database - A repository for quality ratings of mhealth apps. *Dans: 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS) 2020.* 2020. 427-32.

85. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the User Version of the Mobile Application Rating Scale (uMARS). *JMIR Mhealth Uhealth* 2016;4(2):e72. <http://dx.doi.org/10.2196/mhealth.5849>

86. Wyatt JC, Thimbleby H, Rastall P, Hoogewerf J, Wooldridge D, Williams J. What makes a good clinical app? Introducing the RCP Health Informatics Unit checklist. *Clin Med* 2015;15(6):519-21. <http://dx.doi.org/10.7861/clinmedicine.15-6-519>

87. Leigh S, Ouyang J, Mimmagh C. Effective? Engaging? Secure? Applying the ORCHA-24 framework to evaluate apps for chronic insomnia disorder. *Evid Based Ment Health* 2017;20(4):e20. <http://dx.doi.org/10.1136/eb-2017-102751>

88. Jin M, Kim J. Development and evaluation of an evaluation tool for healthcare smartphone applications. *Telemed J E-Health* 2015;21(10):831-7. <http://dx.doi.org/10.1089/tmj.2014.0151>

89. Jake-Schoffman DE, Silfee VJ, Waring ME, Boudreaux ED, Sadasivam RS, Mullen SP, et al. Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. *JMIR Mhealth Uhealth* 2017;5(12):e190. <http://dx.doi.org/10.2196/mhealth.8758>

90. Modave F, Bian J, Leavitt T, Bromwell J, Harris Iii C, Vincent H. Low quality of free coaching apps with respect to the American College of Sports Medicine guidelines: A review of current mobile apps. *JMIR Mhealth Uhealth* 2015;3(3):e77. <http://dx.doi.org/10.2196/mhealth.4669>

91. Guo Y, Bian J, Leavitt T, Vincent HK, Vander Zalm L, Teurlings TL, et al. Assessing the Quality of Mobile Exercise Apps Based on the American College of Sports Medicine Guidelines: A Reliable and Valid Scoring Instrument. *J Med Internet Res* 2017;19(3):e67. <http://dx.doi.org/10.2196/jmir.6976>

92. Choi J, Noh GY, Park DJ. Smoking cessation apps for smartphones: content analysis with the self-determination theory. *J Med Internet Res* 2014;16(2):e44. <http://dx.doi.org/10.2196/jmir.3061>

93. Ouhbi S, Fernández-Alemán JL, Pozo JR, Bajta ME, Toval A, Idrí A. Compliance of blood donation apps with mobile OS usability guidelines. *J Med Syst* 2015;39(6):63. <http://dx.doi.org/10.1007/s10916-015-0243-1>

94. van Galen LS, Xu X, Koh MJA, Thng S, Car J. Eczema apps conformance with clinical guidelines: a systematic assessment of functions, tools and content. *Br J Dermatol* 2020;182(2):444-53. <http://dx.doi.org/10.1111/bjd.18152>

95. Reynoldson C, Stones C, Allsop M, Gardner P, Bennett MI, Closs SJ, et al. Assessing the quality and usability of smartphone apps for pain self-management. *Pain Med* 2014;15(6):898-909. <http://dx.doi.org/10.1111/pme.12327>

96. Myint M, Adam A, Herath S, Smith G. Mobile phone applications in management of enuresis: The good, the bad, and the unreliable! *J Pediatr Urol* 2016;12(2):112 e1-6. <http://dx.doi.org/10.1016/j.jpuro.2015.09.011>

97. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. *Int J Behav Nutr Phys Act* 2014;11:97. <http://dx.doi.org/10.1186/s12966-014-0097-9>

98. McMillan B, Hickey E, Patel MG, Mitchell C. Quality assessment of a sample of mobile app-based health behavior change interventions using a tool based on the National Institute of Health and Care Excellence behavior change guidance. *Patient Educ Couns* 2015. <http://dx.doi.org/10.1016/j.pec.2015.10.023>

99. Siebert JN, Lacroix L, Cantais A, Manzano S, Ehrler F. The impact of a tablet app on adherence to american heart association guidelines during simulated pediatric cardiopulmonary resuscitation: Randomized controlled trial. *J Med Internet Res* 2020;22(5):e17792. <http://dx.doi.org/10.2196/17792>

100. Watson P, Watson D, Dhese A, New HV, Davidson A, Armstrong R, et al. Improving blood-prescribing decisions: Evaluating the efficacy of a blood guidelines app. *Transfus Med* 2020. <http://dx.doi.org/10.1111/tme.12736>

101. DiFilippo KN, Huang W, Chapman-Novakofski KM. A new tool for nutrition App Quality Evaluation (AQEL): Development, validation, and reliability testing. *JMIR Mhealth Uhealth* 2017;5(10):e163. <http://dx.doi.org/10.2196/mhealth.7441>

102. Anderson K, Burford O, Emmerton L. App chronic disease checklist: Protocol to evaluate mobile apps for chronic disease self-management. *JMIR Res Protoc* 2016;5(4):e204.
<http://dx.doi.org/10.2196/resprot.6194>
103. Hanraha C, Dy Aungst T, Cole S. Evaluating mobile medical applications [En ligne]: ASHP eReports; 2014.
<https://www.ashp.org/-/media/store%20files/mobile-medical-apps.pdf>
104. Loy JS, Ali EE, Yap KY. Quality assessment of medical apps that target medication-related problems. *J Manag Care Spec Pharm* 2016;22(10):1124-40.
<http://dx.doi.org/10.18553/jmcp.2016.22.10.1124>
105. Camacho E, Hoffman L, Lagan S, Rodriguez-Villa E, Rauseo-Ricupero N, Wisniewski H, et al. Technology Evaluation and Assessment Criteria for Health Apps (TEACH-Apps): Pilot Study. *J Med Internet Res* 2020;22(8):e18346.
<http://dx.doi.org/10.2196/18346>
106. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. *Lancet Digit Health* 2019;1(2):e52-e4.
[http://dx.doi.org/10.1016/s2589-7500\(19\)30013-5](http://dx.doi.org/10.1016/s2589-7500(19)30013-5)
107. van Haasteren A, Vayena E, Powell J. The mobile health app trustworthiness checklist: Usability assessment. *JMIR Mhealth Uhealth* 2020;8(7):e16844.
<http://dx.doi.org/10.2196/16844>
108. Kaltheuner M, Drossel D, Heinemann L. DiaDigital Apps: Evaluation of smartphone apps using a quality rating methodology for use by patients and diabetologists in Germany. *J Diabetes Sci Technol* 2019;13(4):756-62.
<http://dx.doi.org/10.1177/1932296818803098>
109. Albrecht UV. Catalogue de critères uniformes pour l'autodéclaration de la qualité des applications de santé. Bern: eHealth Suisse; 2019.
https://mhh-publicationsserver.gbv.de/receive/mhh_mods_00000279
110. Nwe K, Larsen ME, Nelissen N, Wong DC. Medical mobile app classification using the national institute for health and care excellence evidence standards framework for digital health technologies: Interrater reliability study. *J Med Internet Res* 2020;22(6):e17457.
<http://dx.doi.org/10.2196/17457>
111. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? *NPJ Digit Med* 2020;3:4.
<http://dx.doi.org/10.1038/s41746-019-0206-x>
112. Gessa A, Jiménez A, Sancha P. Open innovation in digital healthcare: Users' discrimination between certified and non-certified mhealth applications. *J Open Innov Technol Mark Complex* 2020;6(4):130.
113. Haverinen J, Keränen N, Falkenbach P, Maijala A, Kolehmainen T, Reponen J. Digi-HTA: Health technology assessment framework for digital healthcare services. *FinJeHeW* 2019;11:326-41.
114. Vis C, Bührmann L, Riper H, Ossebaard HC. Health technology assessment frameworks for eHealth: A systematic review. *Int J Technol Assess Health Care* 2020;36(3):204-16.
<http://dx.doi.org/10.1017/s026646232000015x>
115. Kolasa K, Kozinski G. How to Value Digital Health Interventions? A Systematic Literature Review. *Int J Environ Res Public Health* 2020;17(6).
<http://dx.doi.org/10.3390/ijerph17062119>
116. Oortwijn W, Determann D, Schiffrs K, Tan SS, van der Tuin J. Towards integrated health technology assessment for improving decision making in selected countries. *Value Health* 2017;20(8):1121-30.
<http://dx.doi.org/10.1016/j.jval.2017.03.011>
117. Jacob C, Sanchez-Vazquez A, Ivory C. Understanding Clinicians' Adoption of Mobile Health Tools: A Qualitative Review of the Most Used Frameworks. *JMIR Mhealth Uhealth* 2020;8(7):e18072.
<http://dx.doi.org/10.2196/18072>
118. Moshi MR, Tooher R, Merlin T. Development of a health technology assessment module for evaluating mobile medical applications. *Int J Technol Assess Health Care* 2020;36(3):252-61.
<http://dx.doi.org/10.1017/s0266462320000288>
119. Larsen ME, Nicholas J, Christensen H. Quantifying app store dynamics: Longitudinal tracking of mental health apps. *JMIR Mhealth Uhealth* 2016;4(3):e96.
<http://dx.doi.org/10.2196/mhealth.6020>
120. Alon N, Stern AD, Torous J. Assessing the food and drug administration's risk-based framework for software precertification with top health apps in the united states: Quality improvement study. *JMIR Mhealth Uhealth* 2020;8(10):e20482.
<http://dx.doi.org/10.2196/20482>
121. Queensland University of Technology. Mobile Application Rating Scale (MARS). App classification. Brisbane: QUT; 2015.

Participants

Les participants à l'élaboration du référentiel HAS de 2016 (1) ont été sollicités pour la relecture de ce travail avec l'apport de nouveaux participants.

Groupe de lecture

- Professeur Marie-Christine Bene, hématologue, Nantes
- Monsieur Jean-Luc Bernard, représentante d'usagers, Pontault-Combault
- Monsieur Olivier Blot, ingénieur, président de Santexium Consulting, Clermont-Ferrand
- Madame Élisabeth Bouchez, cadre supérieure médico-technique, Créteil
- Madame Anne Buisson, représentante d'usagers, Paris
- Docteur François Carbonnel, médecin généraliste, Cabestany, maître de conférence, Montpellier
- Docteur Pascal Charbonnel, médecin généraliste, Paris
- Docteur Didier Cugy, médecin généraliste, Bordeaux
- Professeur Fabrice Denis, médecin oncologue, Le Mans
- Monsieur Guillaume de Durat, chargé de cours en droit de la santé, membre de la cellule éthique du conseil du numérique en santé, Paris
- Professeur Patrick Disdier, médecine interne, Marseille
- Monsieur Alexandre Falzon, ingénieur, président de Nouveal e-santé, Irigny
- Docteur Matthieu Faure, docteur en informatique, ingénieur, Nîmes
- Professeur Sébastien Faure, pharmacien, Angers
- Docteur Sylvia Franc, endocrinologue, Corbeil-Essones
- Docteur Jean-Louis Fraysse, pharmacien, membre de la cellule éthique du conseil du numérique en santé, Toulouse
- Madame Claire-Charlotte Gaulier, infirmière puéricultrice, Strasbourg
- Docteur Leïla Gofti-Laroche, pharmacienne praticien hospitalier, docteur en épidémiologie, Grenoble
- Professeur Sinem-Ezgi Gülmez, médecin et pharmacienne, Istanbul
- Monsieur Marin Guy, masseur-kinésithérapeute, Pessac
- Monsieur Philippe Haïk, ingénieur chercheur, Paris
- Docteur Olivier Heloir, pharmacien, Lille
- Professeur Bruno Housset, pneumologue, Créteil
- Docteur Cécile Hubsch, neurologue, Paris
- Docteur Benjamin Kretz, chirurgien vasculaire, Colmar
- Madame Lyra Ladurée, infirmière, cadre de santé, Paris
- Docteur Loïck Menvielle, professeur, chercheur en sciences de gestion, EDHEC, Nice
- Monsieur Jean-David Miquel, président de Monitorling France, Paris
- Docteur Alexandre Morel, santé publique, Brest
- Monsieur Jean-Philippe Nanteuil, chef de projets solutions digitales innovantes, Paris
- Professeur Gregory Ninot, universitaire, Montpellier
- Monsieur Serge Payeur, ingénieur informatique, Rochefort
- Docteur Gregory Perrard, cardiologue, Bailleul
- Madame Tatiana Rak, responsable juridique, Hyères
- Monsieur Philippe Ruyer, masseur-kinésithérapeute, Les Angles
- Monsieur Bernard Sauvestre, représentant d'usagers, Évry-Courcourones
- Docteur Éric Sermet, psychiatre, Châtillon-en-Michaille
- Docteur Catherine Sztal-Kutas, représentante d'usagers, La Rochelle
- Madame Soumia Taoui, cheffe de projet en recherche clinique, Toulouse
- Monsieur Romain Tavignot, infirmier-anesthésiste, Cannes

Remerciements

La HAS tient à remercier l'ensemble des participants cités ci-dessus pour leur implication.

Abréviations et acronymes

5G	Cinquième génération
ACDC	App Chronic Disease Checklist
ANS	Agence du numérique en santé
ANSM	Agence nationale de sécurité du médicament et des produits de santé
App	Application
AQEL	App Quality Evaluation
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
BPCO	Bronchopneumopathie chronique obstructive
CE	Conformité aux exigences fixées par la réglementation communautaire
CEEBIT	Continuous Evaluation of Evolving Behavioral Intervention Technologies
CNAM	Caisse nationale de l'assurance maladie
CTTI	Clinical Trials Transformation Initiative
CV	App curriculum vitæ
DAP	Digital Assessment Portal
DAQs	Digital Assessment Questions
DGCCRF	Direction générale de la concurrence, de la consommation et de la répression des fraudes
DiGA	Digital Health Applications
DM	Dispositif médical
DNS	Délégation ministérielle au numérique en santé
DTAC	Digital Technical Assessment Criteria
ECR	Essai contrôlé randomisé
ENS	Espace numérique de santé
EQUATOR	Enhancing the QUALity and Transparency Of health Research
EUnetHTA	European network for Health Technology Assessment
FDA	Food and Drug Administration
GIE	Groupement d'intérêt économique
GSMA	Global System for Mobile Communications
HAS	Haute Autorité de santé
HON	Health On the Net
HTA	Health Technology Assessment
IA	Intelligence artificielle
JAMA	Journal of the American Medical Association
JMIR	Journal of Medical Internet Research
MARS	Mobile Application Rating Score

MDM	Minimal disruptive, care
mERA	mHealth Evidence Reporting and Assessment
MHAD	Mobile Health App Database
mHealth	Mobile Health
MMA	Mobile Medical Application
NCSC-FI	National Cyber Security Centre Finland
NHS	National Health Service
NICE	The National Institute for Health and Care Excellence
OBR	ORCHA Baseline Review
OMH	Our Mobile Health
OMS	Organisation mondiale de la santé
ORCHA	Organisation for the Review of Care and Health Applications
PREMs	Patient-reported outcomes measures
PREMs	Patient-reported experience measures
PwC	PricewaterhouseCoopers
RCP	Royal College of Physicians
RGPD	Règlement général sur la protection des données
RPM	Remote patient monitoring
Sesam-Vitale	Système électronique de saisie de l'assurance maladie
UK	United Kingdom
Wi-Fi	Wireless fidelity

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

