



HAUTE AUTORITÉ DE SANTÉ

ÉVALUER

LES TECHNOLOGIES DE SANTÉ

**RAPPORT
D'ELABORATION**


Intégration des dispositifs médicaux numériques à usage professionnel dans la pratique

État des lieux et perspectives
d'aide au choix

Validé par le Collège le 9 novembre 2022

Descriptif de la publication

Titre	Intégration des dispositifs médicaux numériques à usage professionnel dans la pratique État des lieux et perspectives d'aide au choix
Méthode de travail	Analyse de la littérature Entretiens avec parties prenantes
Objectif(s)	Réaliser un état des lieux du développement des dispositifs médicaux numériques à usage professionnel, de leur intégration dans la pratique auprès des professionnels de santé et du positionnement des régulateurs et agences d'évaluation sur ces outils
Cibles concernées	Professionnels de santé utilisateurs de dispositifs médicaux numériques à usage professionnel Décideur public : ministère de la Santé et de la Prévention, Assurance maladie Organes décisionnaires d'achat d'équipements des établissements de santé Industriels et éditeurs de logiciels
Demandeur	Direction de la Sécurité sociale
Promoteur(s)	Haute Autorité de santé (HAS)
Pilotage du projet	Mission Numérique en santé
Recherche documentaire	RENNESSON Marina ; LEFEVRE Maud
Auteurs	ROCHEREAU Aude ; RENNER Simon ; COLLIGNON Corinne
Validation	Version du 9 novembre 2022

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 – novembre 2022 – ISBN : 978-2-11-167563-6

Sommaire

Saisine	9
1. Méthode de travail	11
1.1. Analyse bibliographique et enquête	11
1.2. Périmètre visé	11
2. Analyse de la littérature	15
2.1. Méthodologie	15
2.2. Objectifs	15
2.3. Résultats de la recherche	15
3. Enquête via les membres du réseau INAHTA	30
3.1. Méthodologie	30
3.2. Résultats	30
4. Entretiens parties prenantes	32
4.1. Méthodologie	32
4.2. Résultats	32
4.3. Discussion	37
5. Perspectives	39
5.1. Production à visée informative pour les professionnels : guide d'aide au choix des dispositifs médicaux numériques à usage professionnel	39
5.2. Productions socles permettant la structuration d'une évaluation	40
6. Conclusion	42
Table des annexes	43
Références bibliographiques	46
Participants	49
Abréviations et acronymes	50

Table des figures

Figure 1. Logigramme de sélection des articles	15
Figure 2. Comparaison des phases de développement des thérapies médicamenteuses, de l'IA dans les soins de santé et de l'innovation chirurgicale, Vasey <i>et al.</i> , 2022 (29)	19
Figure 3. Recommandations pour l'évaluation des algorithmes d'intelligence artificielle, Vasey <i>et al.</i> , 2022 (29)	20
Figure 4. Classification des solutions numériques selon le NICE, 2022 (33)	22
Figure 5. Réponses de CNP à la question « Souhaitez-vous que la HAS s'empare du sujet des DMN à usage professionnel ? »	36

Table des tableaux

Tableau 1. Périmètre des DMN à usage professionnel retenus dans le cadre de ce projet, d'après la cartographie fonctionnelle des solutions numériques, HAS, février 2021 (2)	13
Tableau 2. Catégorisation des risques des SaMD, FDA, 2018 (30)	21
Tableau 3. Comparaison des domaines couverts par la grille d'évaluation des technologies avec IA du NHS avec la grille descriptive des fonctionnalités des DM embarquant un système avec IA de la HAS	25
Tableau 4. Recensement des premières évaluations de logiciels à usage professionnel par des agences d'HTA	28

Synthèse générale

Contexte

De très nombreuses solutions numériques sont aujourd'hui utilisées dans le cadre de soins courants. Pourtant, leur utilité ou leur pertinence par rapport à l'arsenal existant ne sont pas toujours réellement établies¹. C'est notamment le cas de certaines solutions numériques à usage professionnel qui n'entrent pas dans les cadres d'évaluation existants pour une prise en charge par la solidarité nationale. De ce fait, **les professionnels de santé peuvent utiliser des dispositifs médicaux numériques (DMN) dans le cadre d'un acte médical sans être pleinement éclairés de leurs performances ou de leurs limites ou, a contrario, être réticents à leur utilisation pour ces mêmes raisons.**

Dans ce contexte, la direction de la Sécurité sociale (DSS) a saisi la HAS « *afin d'établir un état des lieux des solutions numériques existantes et/ou innovantes, de définir leur finalité (diagnostic, prévention, aide à la décision, aide au traitement, etc.), ainsi que les critères de sécurité, de qualité et d'utilisation, leurs spécificités d'évaluation et de réfléchir à la mise en place d'un référentiel qualité* ». La HAS a donc publié le 8 septembre 2022 la [note de cadrage du projet](#) mis en œuvre à partir de cette saisine visant à présenter ses objectifs, le périmètre visé et le déroulé prévu.

Enjeu

Ce projet vise à développer la confiance relative à l'utilisation des DMN à usage professionnel et de favoriser, lorsqu'ils sont utiles et performants, leur intégration dans le système de santé. Pour ce faire, des outils aux professionnels/structures de soins et des recommandations sur le dispositif d'évaluation à instaurer seront proposés.

Il s'agit :

- d'accompagner **les professionnels de santé et les structures de soins dans la sélection de DMN en proposant des approches d'aide au choix les plus adaptées et graduées, sans être un frein à l'innovation ;**
- et de **proposer au ministère en charge de la Santé et de la Prévention un dispositif d'évaluation pour les DMN concernés.**

Objectif de ce document

Parmi les différentes productions à venir pour répondre à cet enjeu, ce premier document de synthèse a pour objectif de réaliser un état des lieux du développement des dispositifs médicaux numériques à usage professionnel, de leur intégration dans la pratique auprès des professionnels de santé et du positionnement des régulateurs et agences d'évaluation sur ces outils, des attentes des professionnels en termes d'aide au choix et de mettre en lumière des perspectives d'aide au choix qui pourront être travaillées dans le cadre de ce projet.

Méthodologie

Face à la diversité des DMN à usage professionnel, le **périmètre visé par la HAS dans le cadre de ce projet se concentre sur les DMN d'aide au dépistage, d'aide au diagnostic, d'aide à la**

¹ Haute Autorité de santé. Numérique : quelle (R)évolution ? Rapport d'analyse prospective 2019. Saint-Denis La Plaine: HAS: 2019.

décision médicale, d'aide à la décision thérapeutique (hors logiciels d'aide à la prescription ou à la dispensation), ainsi que sur les systèmes de gestion autonome de la décision.

Pour cartographier les DMN à usage professionnel et identifier les usages ou besoins du terrain sur l'aide au choix des DMN, la méthodologie suivante a été choisie :

- **une analyse systématique de la littérature et des actions ou productions mises en place par les agences d'évaluation ou systèmes de soins internationaux** pour ce type de dispositif ;
- une enquête auprès du **réseau INAHTA** (*International Network of Agencies for Health Technology Assessment*), réseau international des agences d'évaluation des technologies de santé regroupant 49 agences d'évaluation dans 30 pays ;
- une conduite de réunions **bilatérales informelles avec 28 parties prenantes** : organisations professionnelles, représentants de patients, institutionnels ou experts.

Analyse systématique de la littérature

Les principaux dispositifs médicaux à usage professionnel retrouvés dans la littérature correspondent à des **systèmes d'aide au diagnostic ou au dépistage et les systèmes d'aide à la décision médicale.**

L'analyse de la littérature, de plus en plus prolifique sur les DMN à usage professionnel, met en évidence :

- une **utilisation en pratique des DMN avec intelligence artificielle (IA) encore sporadique** ;
- **l'hétérogénéité des performances** de ces dispositifs, notamment en raison d'un potentiel manque de fiabilité de la donnée d'apprentissage utilisée ;
- un **besoin d'études, en particulier d'études prospectives en vie réelle** venant valider et comparer les DMN ;
- **l'absence de matrice de risques utilisable en l'état pour prioriser les besoins d'évaluation** pour les DMN à usage professionnel ;
- des **premiers cadres d'évaluation** de ces dispositifs.

Peu de travaux sur les DMN à usage professionnel ont été conduits par les régulateurs ou agences d'évaluation. En France, en dehors des situations où un nouvel acte est nécessaire pour leur utilisation, il n'existe pas de cadre d'évaluation national. Trois agences internationales, **l'Institut national d'excellence en santé et en services sociaux (INESSS), la Canadian Agency for Drugs & Technologies in Health (CADTH) et le National Institute for Health and Care Excellence (NICE)**, ont, à ce jour, évalué certains de ces dispositifs en se basant sur les standards habituels d'évaluation technologique (une analyse critique des études cliniques disponibles et différents coûts et modèles économiques proposés). L'INESSS porte notamment une attention particulière à l'analyse de l'efficacité de ces dispositifs et à l'impact de leur intégration sur l'organisation des soins.

Le **National Health Service (NHS)** a quant à lui **mis à disposition un guide spécifique d'aide au choix pour les utilisateurs ou acheteurs de DMN à usage professionnel avec IA.** En plus de la description des données, du modèle et des caractéristiques, ce guide permet aux utilisateurs de **se questionner sur les principaux enjeux** associés à l'utilisation de ces outils, et notamment l'intégration dans le parcours et l'impact organisationnel associé, l'acceptation par les équipes, la maintenance et également la comparaison avec les autres solutions existantes...

Réseau INAHTA

Même si tous les répondants du réseau INAHTA sont susceptibles d'évaluer les DMN à usage professionnel, il y a eu **très peu d'évaluations jusqu'ici pour ce type de technologies**. À ce jour, aucun partenaire du réseau INHATA n'a construit un cadre spécifique pour l'évaluation des DMN à usage professionnel. Certaines agences considèrent que les méthodes classiques d'évaluation s'appliquent. Certaines agences mettent en œuvre une réflexion sur le développement de **cadre spécifique** ou la mise à disposition de **guides méthodologiques** à destination des utilisateurs des DMN à usage professionnel.

Entretiens avec les parties prenantes

Selon les conseils nationaux professionnels (CNP), l'utilisation des DMN à usage professionnel est très **hétérogène**. Les principaux usages relevés sont orientés vers les **systèmes d'aide au diagnostic et au dépistage** utilisés par les spécialités à prédominance diagnostique telles que la radiologie, l'hématologie ou encore la biologie médicale. L'intégration de ces outils est également dépendante de leur acception par les patients. Pour cela, l'information sur l'utilisation des données et le renforcement du dialogue entre patients et professionnels de santé sont primordiaux pour garantir la confiance dans ces outils.

La sélection des outils numériques par les professionnels relève de leur libre choix. Elle n'est pas encadrée et dépend de l'appétence de la profession pour le numérique et de l'intégration dans les pratiques de ces outils.

71 % des 28 parties prenantes interrogées ont indiqué être favorables à une aide de la HAS

avec toutefois des attentes très variables, décrites par niveau d'implication croissant :

- la mise en place de guides d'aide au choix ou de référentiels généraux pour ces outils ;
- une aide à la veille technologique et également une aide à l'expression des besoins des professionnels de santé qui se baserait sur leur pratique du terrain, pour faciliter le repérage en amont des technologies pertinentes ;
- la mise en place d'une aide méthodologique de la HAS dans le cadre de partenariat avec les CNP, notamment avec la production de guides méthodologiques et/ou un soutien personnalisé lors de la mise en place d'études en vie réelle pour évaluer les DMN ;
- la mise en place de « labels » ou équivalents.

Perspectives

Deux principaux types de productions semblent se dégager à ce stade :

- la mise à disposition de productions à destination des professionnels ou structures de soins, comme des **guides généralistes d'aide au choix** des DMN à usage professionnel ;
- la mise en place de productions permettant de construire un socle d'évaluation pour les DMN à usage professionnel, et notamment :
 - **la mise en œuvre de démarches collaboratives avec les CNP, pour accompagner méthodologiquement les CNP qui le désirent dans l'évaluation des DMN à différents degrés ;**
 - **la mise en œuvre d'un dispositif volontaire de labellisation ou équivalent, la production de référentiels d'évaluation** pour les DMN à usage professionnel, directement par la HAS.

D'autres productions, comme la création d'une matrice de risques spécifique aux DMN à usage professionnel, pourraient également être discutées selon les besoins identifiés lors des échanges à

venir avec les parties prenantes. La réalisation de ce type de production reste toutefois complexe à mettre en œuvre, du fait de la notion de risque qui est très hétérogène selon les spécialités médicales et nécessiterait l'implication de l'ensemble des parties prenantes.

Conclusion

À courte échéance et afin d'accompagner les professionnels de santé et les structures de soins dans la sélection de DMN, la **HAS souhaite publier un guide d'aide au choix des DMN à usage professionnel** qui permettra d'orienter les utilisateurs et les acheteurs en leur permettant d'identifier les principales questions clés à se poser ou à poser sur le DMN.

Enfin, selon les besoins qui seront identifiés en termes d'évolution de l'évaluation de ce type de technologies, la HAS **pourra également émettre des propositions aux pouvoirs publics en matière d'évaluation**. Ces réflexions autour de l'évaluation pourront notamment être testées **dans le cadre de plans ou d'expérimentations nationaux**.

Saisine

Demande

La direction de la Sécurité sociale (DSS) a saisi la Haute Autorité de santé (HAS) « *afin d'établir un état des lieux des solutions numériques existantes et/ou innovantes, de définir leur finalité (diagnostic, prévention, aide à la décision, aide au traitement, etc.), ainsi que les critères de sécurité, de qualité et d'utilisation, leurs spécificités d'évaluation et de réfléchir à la mise en place d'un référentiel qualité* ».

Contexte

Le numérique est aujourd'hui présent à toutes les étapes du système de soins. Les technologies numériques ont déjà un rôle majeur dans les organisations de soins et leur place va croître très rapidement. Le numérique est un des leviers susceptibles de contribuer à une amélioration de la prise en charge et du suivi des patients. Le champ de ces solutions est vaste et hétérogène tant dans les fonctionnalités que dans la nature même des technologies utilisées et des champs d'application possibles.

La diversité des solutions numériques est liée notamment aux utilisateurs finaux. Nous pouvons en distinguer deux.

- Les solutions numériques utilisées par les patients eux-mêmes pouvant faire l'objet d'une évaluation par la HAS dès lors qu'elles ont une finalité médicale et que l'entreprise souhaite obtenir son inscription sur la liste des produits remboursables (LPPR) par l'Assurance maladie.

Par exemple, un système qui analyse les données issues d'un moniteur de glucose en continu utilisé par un patient diabétique et qui va automatiquement ajuster le débit basal ou administrer une dose bolus sans que le patient intervienne (pancréas artificiel, également appelé boucle semi-fermée).

- Et les DMN liés à un acte médical ou à usage exclusivement professionnel, utilisés dans le cadre de soins.

Par exemple, les logiciels à finalité médicale embarquant de l'intelligence artificielle, comme un logiciel associé à un scanner ou à un appareil d'IRM avec un algorithme d'apprentissage automatique dont la fonctionnalité est de proposer une aide aux professionnels.

Ces outils peuvent constituer une aide précieuse pour, entre autres, diminuer l'errance diagnostique, aider le professionnel à la décision thérapeutique dans le cadre de la médecine personnalisée ou standardiser une pratique professionnelle. Cependant, leur utilisation ne doit pas engendrer de perte de chance pour les patients et/ou être une source de confusion pour les professionnels. La fiabilité, de leur conception à leur utilisation dans le temps, est donc indispensable pour garantir la qualité et la sécurité des soins impliquant l'utilisation de ces dispositifs médicaux numériques.

Actuellement en France, il n'existe aucun cadre d'évaluation pour guider les professionnels dans leur choix d'utiliser tel ou tel type de DM. Pourtant, à ce jour, de très nombreuses solutions numériques sont utilisées dans le cadre de soins courants sans que leur utilité ou leur pertinence par rapport à l'arsenal existant soient réellement établies dans certains cas (1). En effet, la plupart de ces solutions ne sont pas embarquées dans un DM à usage unique par le patient ; ces technologies sont donc hors champ de la procédure nationale d'évaluation de leur intérêt dans le cadre d'une demande d'inscription sur la LPPR ou sur la liste des activités de télésurveillance médicale à venir. De ce fait, pour ce type de DMN, l'utilisation des divers outils d'évaluation, c'est-à-dire l'organisation de l'évaluation dans son ensemble, souffre d'un défaut de structuration et les professionnels de santé peuvent utiliser des

solutions numériques dans le cadre d'un acte médical sans être pleinement éclairés sur leurs performances ou leurs limites.

Une réflexion sur les spécificités de l'évaluation des DMN à usage professionnel et, au-delà, sur les éventuelles démarches d'évaluation pour éclairer les choix des professionnels de santé et des organes décisionnaires d'achat d'équipements est donc nécessaire.

Ce projet a été divisé en plusieurs étapes. Ce document en est la première production et vise à dresser un état des lieux pour initier cette réflexion *via* la réalisation d'un état des lieux des DMN à usage professionnel, l'analyse de la littérature, les positions des parties prenantes et leurs méthodologies de sélection de ces outils.

1. Méthode de travail

1.1. Analyse bibliographique et enquête

La méthode d'élaboration de cette synthèse repose sur :

- une **revue systématique de la littérature** ;
- une **enquête auprès du réseau *International Network of Agencies for Health Technology Assessment (INAHTA)***, regroupant 49 agences d'évaluation dans 30 pays afin **d'identifier les typologies de DMN à usage professionnel relevant du champ de la saisine** ;
- un état des lieux préliminaire auprès des principales organisations professionnelles susceptibles d'avoir recours à des DMN dans le but de **relever les usages en pratique de ce type d'outils et de recueillir leurs besoins dans l'accompagnement à la sélection des DMN sur des critères de qualité et pertinence**. Des experts et représentants de patients ont également été sollicités.

Les méthodologies mises en œuvre pour chaque axe seront développées dans ce document.

1.2. Périmètre visé

Les solutions numériques liées à un acte médical ou à usage professionnel englobent une diversité d'outils importante. Nous pouvons citer :

- les **logiciels de support métier de gestion administrative** (tels que ceux utilisés pour la prise de rendez-vous ou autres) ;
- les **logiciels d'aide à la prescription (LAP) ou à la dispensation (LAD) ou à l'administration des médicaments** (ayant le statut de dispositif médical) ;
- les **logiciels utilisés dans le cadre de la télémédecine** (téléconsultation, téléexpertise, télésurveillance médicale ou télésoin) ;
- ou encore les **logiciels d'aide au dépistage, d'aide au diagnostic ou d'aide à la décision médicale, avec ou sans intelligence artificielle**.

Dès lors, il convient de délimiter le périmètre de réponse à la saisine.

Ce projet sera focalisé sur les technologies numériques :

- ayant une finalité médicale et relevant du statut de **dispositif médical** ;
- **liées à un acte** médical ;
- et qui ne sont **pas éligibles à un processus d'évaluation national**.

Ne feront pas partie du périmètre du projet :

- les solutions numériques ne rentrant pas dans la catégorie des dispositifs médicaux ;
- les dispositifs médicaux numériques **à usage individuel (c'est-à-dire ceux utilisés par les patients eux-mêmes)** entrant déjà dans un processus d'évaluation *via* la LPPR ou la liste des activités de télésurveillance ;
- les dispositifs médicaux numériques **ayant fait ou faisant déjà l'objet de travaux à la HAS** tels que les logiciels d'aide à la prescription ou à la dispensation de médicaments pour lesquels elle a élaboré une procédure et des référentiels de certification ou la télémédecine (notamment téléconsultation et télésoin), faisant l'objet de référentiels et de guides de bonnes pratiques.

Le périmètre visé sera ainsi concentré sur les DMN d'aide au dépistage, d'aide au diagnostic, d'aide à la décision médicale, d'aide à la décision thérapeutique (hors logiciels LAP/LAD), les DMN visés pouvant cumuler plusieurs fonctionnalités.

La liste finale des DMN concernés par ces travaux est présentée dans le Tableau 1.

En synthèse

Face à la diversité des dispositifs médicaux numériques (DMN) à usage professionnel, le **périmètre visé par la réponse à la saisine se concentre sur les DMN d'aide au dépistage, d'aide au diagnostic, d'aide à la décision médicale, d'aide à la décision thérapeutique (hors logiciels de type LAP/LAD), ainsi que sur les systèmes de gestion autonome de la décision.**

Pour cartographier les DMN à usage professionnel et identifier les usages ou besoins du terrain sur l'aide au choix des DMN, la méthodologie suivante a été choisie :

- **une analyse systématique de la littérature et des actions ou productions mises en place par les agences d'évaluation ou systèmes de soins internationaux** pour ce type de dispositif ;
- une enquête auprès du **réseau INAHTA** ;
- une conduite de réunions **bilatérales informelles avec les parties prenantes.**

Tableau 1. Périmètre des DMN à usage professionnel retenus dans le cadre de ce projet, d'après la cartographie fonctionnelle des solutions numériques, HAS, février 2021 (2)

Catégorie	Fonctionnalités	Finalité d'usage visée	Technologies hors périmètre	Exemples
C5	Aide au dépistage sans autonomie de la solution numérique ²	Permet une aide au dépistage d'une pathologie ou d'une déficience par la production d'options proposées au patient ou au professionnel. Les solutions numériques concernées sont destinées à être utilisées dans une population large et impliquent une confirmation diagnostique, le cas échéant.	Ne sont pas dans cette catégorie les systèmes qui fournissent des conseils de santé généraux non personnalisés.	Logiciel de dépistage de la rétinopathie diabétique Système connecté de détection de maladies génétiques ou de cancers
C6	Aide au diagnostic sans autonomie de la solution numérique	Permet une aide au diagnostic personnalisé d'une pathologie ou d'un état de santé par la production d'options diagnostiques ou d'informations pronostiques au professionnel.	Ne sont pas dans cette catégorie les systèmes qui permettent uniquement de fournir une liste non personnalisée de symptômes et signes d'une maladie.	Système qui utilise des données du patient pour aider un professionnel de santé au diagnostic d'une pathologie (par exemple : un cancer) ou d'un état de santé (par exemple : confirmation et datation d'une grossesse). Les systèmes fournissant des informations pronostiques sont dans cette catégorie. Par exemple : <ul style="list-style-type: none"> – Logiciel permettant la détection de tumeurs par techniques d'imagerie – Logiciel associé à une bande thoracique pour détecter les pauses de respiration afin de diagnostiquer une apnée du sommeil – Logiciel d'aide au diagnostic à partir de résultats de biologie médicale...

² On entend par solution non autonome une technologie qui fonctionne en circuit ouvert, c'est-à-dire une solution qui ne mettra en œuvre aucune action thérapeutique, de dépistage ou de diagnostic sans intervention humaine préalable.

C7	Aide au traitement sans autonomie de la solution numérique	Permet le traitement, la détermination des paramètres de sa mise en œuvre ou l'orientation de la décision médicale. La technologie est utilisée lors des soins ou en amont pour optimiser leur mise en œuvre.	Ne sont pas dans cette catégorie les systèmes qui sont destinés à la prise de décision thérapeutique.	<p>Système qui utilise des données du patient pour calculer directement les paramètres d'un traitement (par exemple en radiothérapie) ou les modalités de mise en œuvre d'un traitement (débit de perfusion, dilutions...). Selon les cas, les utilisateurs peuvent être ceux qui mettent en œuvre le traitement, c'est-à-dire les prescripteurs, ceux qui les dispensent (pharmaciens, préparateurs) ou ceux qui les administrent (infirmier). Par exemple :</p> <ul style="list-style-type: none"> – Application de calcul de dose d'insuline en fonction des données glycémiques – Logiciel d'adaptation de posologie en fonction de la fonction rénale <p>Système utilisé par des professionnels de santé à des fins d'aide au traitement, quel qu'il soit (chirurgical, rééducation...). Par exemple :</p> <ul style="list-style-type: none"> – Solution de gamification utilisée lors de soins en présence d'un professionnel de santé (par exemple : une application dans le cadre d'une rééducation locomotrice ou oculomotrice) – Solution de rééducation du périnée utilisée dans le cadre de séances réalisées en présence d'un professionnel de santé – Solution d'orientation chirurgicale orthopédique...
C8	Aide à la décision thérapeutique	Propose, d'après les données du patient, une ou plusieurs options au professionnel pour aider à la décision thérapeutique d'une pathologie diagnostiquée.	Ne sont pas dans cette catégorie les systèmes qui fournissent des conseils de santé généraux non personnalisés ou les systèmes destinés à améliorer les parcours de soins.	Système qui utilise des données du patient afin de proposer aux professionnels les différentes options de traitement possibles (par exemple en oncologie). Les utilisateurs sont ceux qui prescrivent et ceux qui dispensent le traitement.
D1	Gestion autonome de la décision	Analyse les données, diagnostique puis ajuste automatiquement et de façon autonome le traitement à administrer ou à mettre en œuvre.	Ne sont pas dans cette catégorie les systèmes qui intègrent des fonctionnalités de cette catégorie mais fonctionnant en circuit ouvert.	<p>Système qui utilise des données du patient afin de diagnostiquer une situation pathologique et ensuite de modifier automatiquement, sans intervention humaine préalable, les paramètres d'un traitement.</p> <p>À ce jour, les exemples de ce type de technologies totalement autonomes semblent limités.</p> <p>Par exemple :</p> <ul style="list-style-type: none"> – Défibrillateur cardiaque implanté avec une solution de télésurveillance qui analyse les données issues d'un moniteur cardiaque, délivre un choc en cas d'arrêt cardiaque et peut transmettre les alertes au professionnel qui suit le patient.

2. Analyse de la littérature

2.1. Méthodologie

La recherche documentaire a été limitée aux publications en langue anglaise et française sur la période d'avril 2017 à mai 2022. Ont été recherchés : les sites internet, évaluations technologiques, les recommandations, les conférences de consensus, les guides, les méta-analyses, les revues systématiques et les études cliniques. Une veille a ensuite été réalisée jusqu'à septembre 2022.

La recherche a été réalisée en interrogeant :

- plusieurs bases de données bibliographiques, dont *Medline* et divers sites internet tels que la *Cochrane Library* ;
- les sites internet publiant des recommandations ;
- les sites internet publiant des rapports d'évaluation technologique ;
- les sites internet des sociétés savantes.

2.2. Objectifs

Les objectifs de la revue systématique étaient d'identifier :

- les différentes typologies de DMN en développement pour chacune des grandes spécialités concernées et les principales aires thérapeutiques concernées ;
- les premières évaluations des performances métrologiques de ces DMN ;
- le positionnement des régulateurs et agences d'évaluation HTA.

2.3. Résultats de la recherche

2.3.1. Volumétrie

Au total, 945 références bibliographiques ont été identifiées à partir de la recherche systématique effectuée sur *Medline* et 37 sur les sites web. Parmi ces références, 318 ont été analysées et 49 ont été retenues. Le processus de sélection est illustré dans la figure 1 ci-dessous.

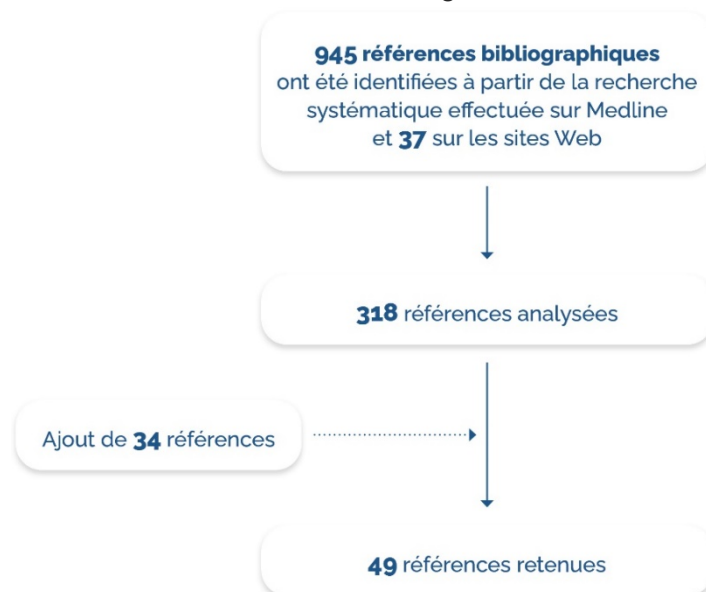


Figure 1. Logigramme de sélection des articles

2.3.2. Principaux éléments issus de la littérature

2.3.2.1. Typologies des dispositifs médicaux numériques à usage professionnel identifiées

En raison de l'hétérogénéité, de la rapidité d'évolution et/ou d'innovation et du caractère parfois éphémère de ces technologies, il n'est pas possible d'en effectuer une cartographie exhaustive. **Un état des lieux des grandes tendances de développement de ce type d'outils** via la revue systématique de la littérature scientifique a été mis en œuvre.

Considérant les différentes catégories de produits, les principaux DMN à usage professionnel mis en évidence par l'analyse de la littérature sont les systèmes d'aide au diagnostic ou au dépistage et les systèmes d'aide à la décision médicale (3). Ces systèmes peuvent correspondre à une matérialisation sous forme numérique d'arbres décisionnels validés et utilisés en pratique clinique, mais ils peuvent également se baser sur des modèles plus complexes tels que l'apprentissage profond ou *deep learning* (4).

S'agissant des usages ou applications cliniques, à ce jour, les systèmes d'aide au diagnostic ou au dépistage utilisant de l'apprentissage profond se concentrent essentiellement sur de l'imagerie (radiographies, scanner, photos...) et concernent des disciplines comme la radiologie, la cancérologie, la pneumologie, la dermatologie, l'anatomopathologie, la gastroentérologie ou encore en ophtalmologie. Parmi ces applications, peuvent être notamment cités les systèmes d'aide au diagnostic ou au dépistage de :

- nodules pulmonaires ou de la tuberculose à partir de radiographies thoraciques (5, 6) ;
- nodules cancéreux à partir de mammographies (7) ;
- mélanomes à partir d'image de lésions cutanées (8, 9) ;
- fractures osseuses à partir de radiographies (10) ;
- rétinopathies diabétiques à partir de fond d'œil (11) ;
- polypes lors de colonoscopie (12).

À noter qu'il existe également des **logiciels intermédiaires de traitement d'images**, avec ou sans IA, utilisés à une étape intermédiaire entre la prise de mesure et le diagnostic : c'est par exemple le cas des logiciels avec IA d'amélioration de la résolution d'images de scanners et de réduction du bruit en radiologie (13).

Les systèmes d'aide à la décision médicale permettent quant à eux de **calculer les paramètres ou les modalités de mise en œuvre d'un traitement directement à partir des données du patient**. D'autres usages pour la prédiction de réponses thérapeutiques se développent.

Plus rarement, des DMN se basant sur les **technologies de réalité virtuelle et augmentée** sont retrouvés en chirurgie dans le but d'accompagner ou de simuler des interventions chirurgicales. C'est le cas des logiciels de navigation chirurgicale utilisés en chirurgie orthopédique notamment (14).

Aucun dispositif médical numérique à usage professionnel intégrant une gestion autonome de la décision médicale n'a été identifié dans cette analyse exploratoire.

2.3.2.2. Des freins à l'utilisation de ces dispositifs médicaux numériques en pratique courante de soins

Malgré un nombre de publications croissant sur les DMN à usage professionnel, leur **utilisation et leur adoption par les professionnels restent encore limitées en pratique courante de soins** (3).

Il existe en effet, en plus des questions liées à la performance des DMN, plusieurs barrières à l'adoption par les professionnels :

- le **positionnement et l'intégration des DMN à usage professionnel dans le parcours de soins et dans le flux de travail des professionnels de santé (15)**. En effet, ces outils peuvent impliquer de modifier la prise en charge des patients ou les pratiques des professionnels de santé ;
- le **temps nécessaire** pour utiliser ces outils, auquel peuvent s'ajouter des **difficultés à maîtriser la prise en main** si leur utilisation n'est pas suffisamment **intuitive et ergonomique** (3) ;
- le **manque de pertinence clinique des informations apportées au regard de la quantité d'informations médicales non utiles en pratique** (4). Ainsi, certaines informations issues des DMN ne sont pas directement transposables en actions médicales utiles (16) et le **risque de multiplication des alertes** inutiles ou non adaptées peut être à l'origine d'une **désensibilisation aux alarmes** (4) ;
- **l'aide limitée apportée par de tels outils en absence de consensus clinique**. En effet, pour les DMN d'aide à la décision médicale, en l'absence de consensus pour la prise en charge d'un cas donné, l'intérêt clinique du DMN sera limité, sans compter les difficultés lors de l'entraînement et l'évaluation de l'outil (3) ;
- à l'inverse, le **manque d'explicabilité** du résultat produit par un DMN, à l'origine d'un effet « **boîte noire** » (3). En effet, un résultat basé sur une prédiction mathématique complexe peut constituer un frein majeur en termes de confiance des professionnels de santé dans les DMN (15) ;
- **la crainte de mauvais fonctionnement du dispositif** (3). La moindre erreur due à l'utilisation d'un DMN à usage professionnel suscitera des perceptions négatives importantes (16).

2.3.2.3. Des technologies aux performances hétérogènes

Une fois le dispositif médical numérique à usage professionnel développé, sa **performance peut être établie de manière :**

- **théorique** avec une performance établie sur un échantillon test ;
- **en condition réelle d'utilisation**, où le DMN doit être intégré dans l'environnement clinique et peut faire l'objet de réglages sur site (sensibilité/spécificité).

Ces performances théoriques et en conditions réelles d'utilisation sont susceptibles de varier en raison de différents facteurs, à l'origine d'un **manque de validité externe des algorithmes :**

- le **volume de données d'apprentissage** nécessaire à l'entraînement d'algorithme d'intelligence artificielle parfois **insuffisant** amenant les éditeurs à réutiliser des données historiques de qualité variable ou labellisées de façon hétérogène (17) ;
- le **biais de représentativité de l'échantillon** utilisé pour l'entraînement d'algorithme d'intelligence artificielle. L'échantillon peut être représentatif d'un type de population mais la performance de l'utilisation du DMN ne sera pas extrapolable à grande échelle auprès de populations différentes. C'est par exemple le cas de certains DMN d'aide au diagnostic de cancer cutané

- en dermatologie : si l'algorithme est entraîné majoritairement sur des photos de « populations blanches », il sera moins performant sur les peaux de couleurs (18) ;
- le risque d'**entraînement du modèle sur des facteurs de confusion, de surapprentissage³ et l'absence de la prise en compte du contexte environnemental du patient lors de l'apprentissage** (16, 19). De plus, les échantillons d'entraînement de ces algorithmes ne contenant pas ou peu de données cliniques, les algorithmes peuvent être peu adaptés ou extrapolables en routine clinique ;
 - pour finir, la **qualité de la donnée d'entrée est essentielle et peut impliquer la nécessité de standardiser la réalisation d'actes et d'interventions** pour garantir les bonnes performances du DMN en conditions réelles. Par exemple, dans la détection de mélanomes en dermatologie à partir de photographies de smartphone, les variations d'angles, de lumière ou de qualité d'image peuvent considérablement modifier les performances du DMN (20) ;
 - par ailleurs, une vigilance doit également être portée sur **les potentielles erreurs de programmation ou une maintenance du système insuffisante pouvant conduire à un DMN rapidement obsolète en pratique** (4).

Un certain nombre d'informations nécessaires pour répondre à des questionnements sur ces dimensions importantes dans la performance d'un DM numérique intégrant un ou plusieurs fonctionnalités apprenantes sont prises en compte dans la grille de la HAS produite en 2020 « [Grille descriptive des fonctionnalités des dispositifs médicaux embarquant un système avec apprentissage automatique \(intelligence artificielle\)](#) » (21).

2.3.2.4. Le besoin d'études robustes

Des méta-analyses et revues systématiques ont souligné le manque d'études robustes pour évaluer la performance des DMN à usage professionnel (22-24). En cause principale, les limites méthodologiques des études mises en œuvre, justifiées par les éditeurs par les spécificités du numérique (25).

Tout d'abord, **il existe peu d'études comparatives machine vs professionnels de santé ou entre différents algorithmes** (19, 25). Lorsqu'elles existent, par exemple pour les systèmes d'aide au diagnostic, il est relevé que l'approche d'évaluation de la précision du diagnostic de modèles d'intelligence artificielle ne reflète pas la pratique clinique (22). En effet, ceux-ci sont évalués de façon isolée sans prendre en compte le cadre clinique d'utilisation de ces outils. D'autre part, certaines publications ont révélé des faiblesses dans la méthodologie de ces études tels que l'absence de calcul de l'échantillon ou encore l'utilisation de jeux de données tests différents entre les groupes comparés (17, 22, 25).

Actuellement, **les études pour évaluer la performance des DMN sont essentiellement rétrospectives, sur un jeu de données déjà préassemblé et issu d'un site unique** (17, 25, 26). L'avantage est qu'elles contiennent des données de routine clinique. Cependant, la qualité de ces données est souvent discutable : les données cliniques sont collectées en routine, issues de dossiers patients et comptes-rendus médicaux (comptes-rendus de radiologie ou d'histologie par exemple). Ces jeux de données, dont l'usage initial ne prévoyait pas ces études, ne sont donc pas harmonisés (opérateur-dépendant) et comportent de nombreuses données manquantes et biais potentiels (25).

Enfin, quel que soit le type d'étude, **la plupart de ces modèles font uniquement l'objet d'une évaluation interne**, c'est-à-dire que l'échantillon de données de test est issu du même échantillon que les

³ Le surapprentissage ou *overfitting* correspond à un problème fréquent dans le domaine de l'apprentissage automatique, où le modèle est très performant lorsqu'il est testé sur des données provenant de la même source que ses données d'apprentissage, mais où ses performances ne se généralisent pas à de nouvelles sources de données (19).

données d'entraînement. « *L'absence de validation de l'algorithme sur un jeu de données indépendant ne présupant pas de l'extrapolation des résultats à d'autres populations, ceci peut engendrer une surestimation des performances des modèles d'intelligence artificielle par rapport aux conditions réelles d'utilisation* » (25).

Finalement, les études d'évaluation des algorithmes sont centrées principalement sur la performance. Or, il existe différents niveaux d'évaluation en plus de l'évaluation théorique de la performance, l'évaluation de la performance en situation réelle d'utilisation peut être nécessaire, et également l'évaluation de l'intérêt clinique et/ou organisationnel apporté par la solution en pratique de soins courante. **Le bénéfice clinique et/ou l'impact organisationnel d'un algorithme peuvent être limités malgré une excellente performance.** Par exemple, un système d'aide au diagnostic doté d'une sensibilité élevée et d'une moins bonne spécificité entraîne plus de faux positifs et donc une potentielle perte de temps pour les professionnels de santé. Sa pertinence dépendra du contexte d'utilisation. Se pose également une question sur le gain réel de ces outils pour le patient. Le choix de l'indicateur de la performance est donc primordial dans la construction d'un algorithme (27).

C'est le cas notamment d'outils d'aide au diagnostic de cancers cutanés paramétrés avec une sensibilité haute dans le but de détecter un plus grand nombre de mélanomes. Un algorithme peut alors classer à tort un plus grand nombre de nævi bénins comme malins (faux positifs). Des diagnostics qui seraient établis à partir de ce type d'outils pourraient donc conduire à des actes et biopsies inutiles, source de préjudice pour les patients (19, 28).

Très récemment, un consensus d'experts a proposé des recommandations (nommées DECIDE-AI) pour l'évaluation précoce des systèmes d'aide à la décision basés sur l'intelligence artificielle (29) et reprend d'autres recommandations publiées pour l'évaluation d'algorithmes d'intelligence artificielle (TRIPOD-AI, STARD-AI, SPIRIT-AI, et CONSORT-AI) préconisées selon les stades de développement de l'algorithme (2). Ainsi, ces recommandations ont vocation à améliorer la standardisation des plans de développement pour l'évaluation des systèmes d'aide à la décision.

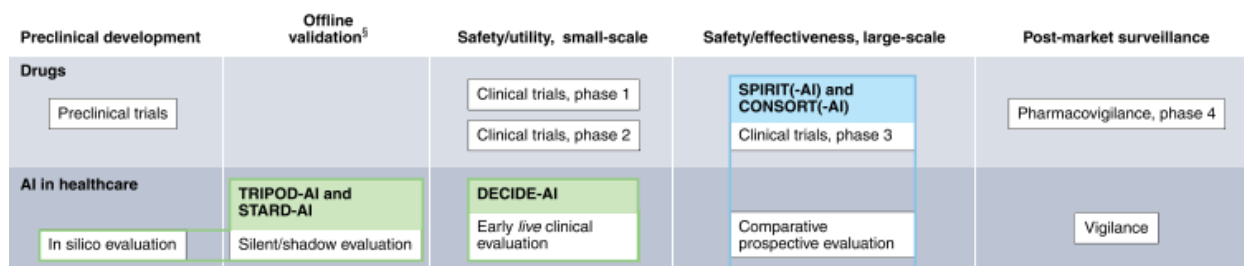


Figure 2. Comparaison des phases de développement des thérapies médicamenteuses, de l'IA dans les soins de santé et de l'innovation chirurgicale, Vasey *et al.*, 2022 (29)

Les recommandations existantes relatives aux schémas d'étude pour l'évaluation des algorithmes d'intelligence artificielle sont résumées dans la figure 3 ci-dessous.

- Les recommandations TRIPOD-AI et STARD-AI concernent **l'évaluation des performances et peuvent être utilisées au stade préclinique.**
- DECIDE-AI est la recommandation développée par Vasey *et al.* (29) pour **évaluer l'intérêt clinique et la sécurité de l'algorithme tout en prenant en compte les conditions réelles d'utilisation (facteurs humains) dès le stade précoce de développement d'un algorithme d'intelligence artificielle.**

- Enfin, les recommandations SPIRIT-AI et CONSORT-IA **concernent la phase 3 du développement clinique et prônent la mise en œuvre de schémas d'étude comparatifs et prospectifs des algorithmes d'intelligence artificielle comme *gold standards***.
- À noter, aucun consensus n'a à ce jour été proposé pour l'évaluation des algorithmes d'intelligence artificielle en phase de surveillance post-mise sur le marché.

AI reporting guidelines			
Name	Stage	Study design	Comment
TRIPOD-AI	Preclinical development	Prediction model evaluation	Extension of TRIPOD. Used to report prediction models (diagnostic or prognostic) development, validation and updates. Focuses on model performance.
STARD-AI	Preclinical development and offline validation	Diagnostic accuracy studies	Extension of STARD. Used to report diagnostic accuracy studies, either at development stage or as an offline validation in clinical settings. Focuses on diagnostic accuracy.
DECIDE-AI	Early live clinical evaluation	Various (prospective cohort studies, non-randomized controlled trials, ...) with additional features, such as modification of intervention, analysis of pre-specified subgroups or learning curve analysis.	Stand-alone guideline. Used to report the early evaluation of AI systems as an intervention in live clinical settings (small-scale, formative evaluation), independently of the study design and AI system modality (diagnostic, prognostic, therapeutic). Focuses on clinical utility, safety and human factors.
SPIRIT-AI	Comparative prospective evaluation	Randomized controlled trials (protocol)	Extension of SPIRIT. Used to report the protocols of randomized controlled trials evaluating AI systems as interventions.
CONSORT-AI	Comparative prospective evaluation	Randomized controlled trials	Extension of CONSORT. Used to report randomized controlled trials evaluating AI systems as interventions (large-scale, summative evaluation), independently of the AI system modality (diagnostic, prognostic, therapeutic). Focuses on effectiveness and safety.

Bold font indicates the primary target of the guidelines, either a specific stage or a specific study design. ^aAlthough existing reporting guidelines exist for some of these study designs (for example, STROBE for cohort studies), none covers all the core aspects of AI system early-stage evaluation, and none would fit all possible study designs; DECIDE-AI was, therefore, developed as a new stand-alone reporting guideline for these early, live, clinical AI studies.

Figure 3. Recommandations pour l'évaluation des algorithmes d'intelligence artificielle, Vasey *et al.*, 2022 (29)

2.3.2.5. Le positionnement des régulateurs, agences d'évaluation des technologies de santé et autres représentants académiques

2.3.2.5.1. Classifications existantes des solutions numériques selon leur niveau de risques

Face à la diversité des technologies et des fonctionnalités des DMN à usage professionnel, leur classification en matrice de risques apparaît utile pour évaluer les risques associés à leur utilisation.

La **littérature n'a pas mis en évidence de classification des DMN à usage professionnel selon leur niveau de risques**. Cependant, trois matrices de risques de solutions numériques existantes, non spécifiques aux DMN à usage professionnel, ont été identifiées :

- **Classification des *Software as a Medical Device* (FDA)**

La première catégorisation des risques concerne les logiciels ayant le statut de dispositif médical ou *Software as a Medical Device* (SaMD) (30). Elle est issue des travaux de l'*International Medical Device Regulators Forum* (IMDRF), groupe de travail composé de régulateurs mondiaux de DM. Un SaMD y est défini comme un logiciel ayant une finalité médicale sans être intégré à un dispositif médical *hardware*. La catégorisation de risques a notamment pour objectif d'attribuer à un SaMD un score de risques afin de déterminer le niveau d'informations à fournir par l'industriel en termes de contrôle qualité ou encore d'évaluation clinique. Les niveaux de risques, présentés dans le Tableau 2, sont classés de I à IV (le quatrième étant le risque le plus critique) et sont fondés sur la situation clinique d'une pathologie ou état de santé (critique, modéré, faible) et le type d'informations fournies par le SaMD pour la prise de décision d'une prise en charge (traiter ou diagnostiquer, orienter les décisions cliniques et transmettre des informations cliniques). Cette classification a été reprise par la *Food and Drug*

Administration (FDA) dans son guide de décembre 2017 (30) « *Software as a Medical Device (SaMD): Clinical Evaluation* » qui a pour objectif d'établir des recommandations sur l'évaluation clinique des SaMD pour démontrer leur sécurité, leur efficacité et leur performance.

State of Healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV	III	II
Serious	III	II	I
Non-serious	II	I	I

Tableau 2. Catégorisation des risques des SaMD, FDA, 2018 (30)

Cette matrice donne quelques indications sur les finalités d'utilisation les plus à risques : par risques décroissants, les DM à finalité diagnostique ou thérapeutique, puis aide à la décision clinique et information à la décision clinique. Elle n'est cependant pas spécifique aux logiciels à destination des professionnels de santé et ne fait pas la distinction entre les utilisateurs du DM, patients ou professionnels de santé. De plus, la catégorisation de la maladie ou de l'état de santé selon les trois niveaux de gravité apparaît comme très subjective. Les critères devant être pris en compte pour apprécier la criticité ne sont en effet pas définis. Enfin, l'exclusion des logiciels embarqués avec un DM de type *hardware* pose également question dans un objectif d'utilisation de cette matrice pour les DMN à usage professionnel.

En parallèle, la FDA a adapté sa surveillance réglementaire aux spécificités des SaMD ou logiciels ayant le statut de dispositif médical, notamment pour ceux intégrant de l'intelligence artificielle. Ainsi, dans le cadre de son programme *Pre-Cert* (31), la FDA propose un nouveau paradigme réglementaire pour l'accès au marché des logiciels avec intelligence artificielle et provenant d'entreprises précertifiées sur des critères de qualité comme d'excellence organisationnelle et s'étant engagées à suivre les performances en vie réelle de leur(s) logiciel(s). Cette surveillance continue vise à assurer le maintien de la sécurité et l'efficacité tout au long du cycle de vie du produit et ouvre la voie à des adaptations algorithmiques après mise sur le marché (32), tant que ces modifications prévues ont été spécifiées lors de la demande de précommercialisation.

– **Evidence Standards Framework for Digital Health Technologies (NICE)**

Dans son document « [Evidence Standards Framework for Digital Health Technologies](#) » (33), mis à jour en août 2022, le NICE propose une classification des solutions numériques stratifiées par niveaux de risques pour les utilisateurs et le système de santé et y fait correspondre un niveau de preuve nécessaire pour chaque niveau fonctionnel de la classification.

Cette classification comprend trois grands niveaux de risques, nommés « Tier » :

- Tier A : solutions numériques utilisées pour des fonctions administratives ou de gestion sans effets directs sur le patient mais qui apportent des bénéfices pour le système de santé ;
- Tier B : solutions numériques visant à aider les patients à la gestion de leur pathologie :

- fournissant des informations aux utilisateurs sur les pathologies ou des conseils de promotion de la santé,
 - permettant le suivi de sa pathologie pour un patient (journal de bord, etc.),
 - de communication entre professionnels de santé ;
- Tier C : solutions numériques à destination des patients et des professionnels utilisées comme systèmes d'aide à la décision clinique ou pour traiter ou diagnostiquer. Ce niveau de risques est divisé en quatre catégories selon les fonctionnalités des solutions :
- transmission d'informations cliniques,
 - aide à la décision clinique,
 - aide au diagnostic,
 - aide au traitement.

Le Tier C reprend la classification proposée par l'IMDRF et se base à la fois sur la nature de l'information apportée par la solution numérique et l'état de la situation ou l'état de santé.

Cette matrice de risques ne concerne donc pas uniquement les DMN à usage professionnel, qui sont en fait tous regroupés dans la catégorie « Tier C », niveau de risques le plus élevé, et ne permet pas l'utilisation de cette matrice de risques en l'état pour discriminer les DMN à usage professionnel selon les risques associés à leur utilisation.

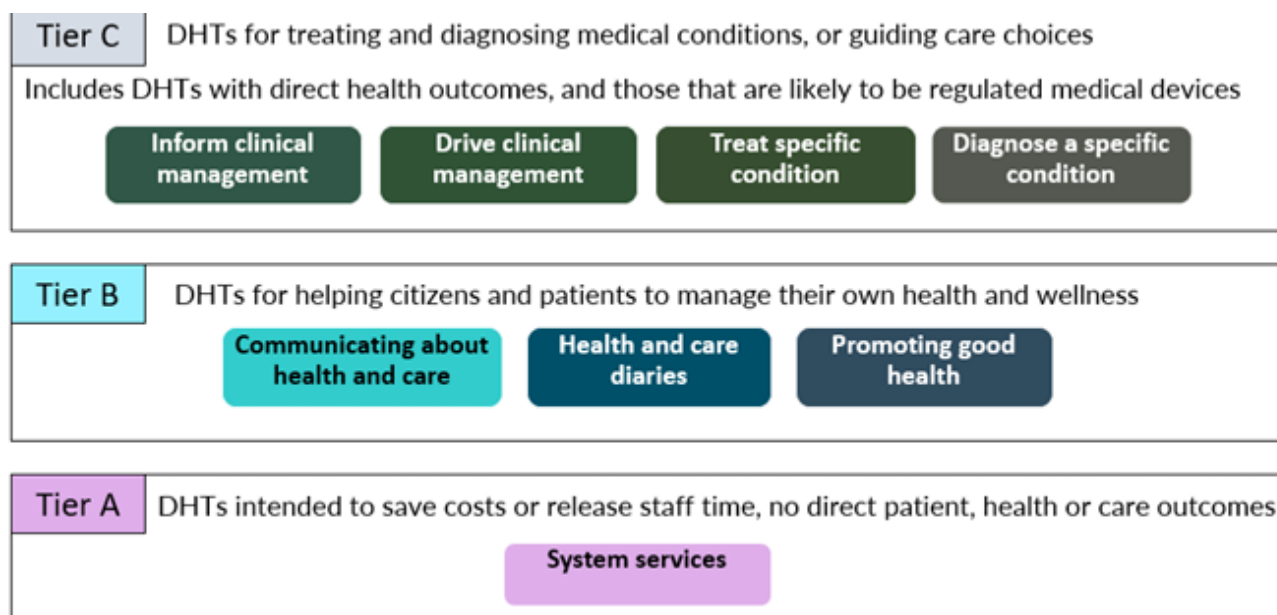


Figure 4. Classification des solutions numériques selon le NICE, 2022 (33)

– **Classification fonctionnelle des solutions numériques selon leur finalité d'usage (HAS)**

Enfin, la HAS a publié, en février 2021 (2), la « [Classification fonctionnelle, selon leur finalité d'usage, des solutions numériques utilisées dans le cadre de soins médicaux ou paramédicaux](#) ». Cette classification compte 11 types de solutions numériques classés en quatre niveaux (A, B, C, D), selon leur finalité d'usage complétée par la capacité de la solution numérique à prendre en compte les paramètres de l'usager/du patient et le caractère autonome ou non de la solution numérique, permettant de distinguer les technologies nécessitant une intervention humaine pour mettre en œuvre une action

thérapeutique, de dépistage ou de diagnostic, et celles générant d'elles-mêmes, c'est-à-dire sans intervention humaine préalable, ce même type d'action. Cette classification a été conçue avant tout en tant qu'outil socle de référence destiné aux différents acteurs potentiels et pour des usages multiples. Elle pourrait être utilisée dans une approche matricielle.

Dans les différentes classifications déjà produites, on note une évolution du niveau de granularité entre les classifications FDA, NICE et HAS. En effet, la classification FDA est fondée sur des critères très « macros », tandis que le NICE affine sa classification en augmentant le nombre de fonctionnalités distinguées mais n'individualise pas les solutions numériques selon qu'elles embarquent ou non de l'intelligence artificielle. Enfin, la classification HAS possède le niveau de granularité le plus fin en termes de fonctionnalités et fait la distinction selon le type d'utilisateurs (professionnels de santé, patients et aidants). Elle sépare également en différentes catégories les solutions embarquant de l'intelligence artificielle ou non. La HAS a indiqué d'emblée que cette classification avait vocation à être développée au fil de ses utilisations.

Les DMN à usage professionnel ne font pas l'objet d'une catégorie propre au sein des différentes classifications précédentes. Cependant, la classification proposée par la HAS, construite dans la perspective d'être réutilisée pour intégrer une composante matricielle, pourrait servir de base à l'élaboration d'une éventuelle matrice de risques, voire d'évaluation, dans le cadre de ce projet en fonction des échanges à venir.

Il est à noter que la HAS a publié en novembre 2016 (34) un « [Référentiel de bonnes pratiques sur les applications et les objets connectés en santé \(mobile Health ou mHealth\)](#) ». Bien que le périmètre de ce référentiel concerne les applications ou objets connectés ayant un effet potentiel sur la santé sans être un dispositif médical, la HAS a considéré que les applications et objets connectés utilisés par les professionnels de santé, même s'ils n'ont pas de finalité médicale annoncée par l'exploitant, avaient déjà le « niveau de criticité » le plus élevé. Au regard de l'utilisation des DMN à usage professionnel comme outils d'aide à la décision clinique, ceci soulève une fois de plus le questionnement sur le potentiel risque d'utiliser ces outils sans que les preuves de leur sécurité et leur performance en condition réelle d'utilisation ne soient réellement établies.

2.3.2.5.2. Travaux sur des méthodes d'aide au choix par les usagers

Récemment, **le NHS AI Lab a publié un guide de l'acheteur de ces technologies à destination des organisations ou des professionnels de santé (35), couplé à une grille d'évaluation (36, 37)**. L'objectif de cette grille d'évaluation est de suggérer les questions importantes à prendre en compte lors de la sélection de dispositifs d'IA. Le guide en lui-même quant à lui **apporte une aide au choix du DMN en permettant notamment à l'acheteur de se questionner sur les aspects réglementaires du produit, de la validation du modèle et de son intégration dans la pratique clinique**.

Le questionnaire comporte près d'une centaine de questions qui sont retrouvées dans la grille d'évaluation associée, et notamment des questions de type :

- Quel problème essayez-vous de résoudre, et l'intelligence artificielle est-elle la bonne solution ?
- Ce produit est-il conforme aux exigences réglementaires ?
- Les performances de ce produit sont-elles conformes aux revendications du fabricant ?
- Ce produit fonctionnera-t-il dans la pratique ?

- Pouvez-vous garantir le soutien dont vous aurez besoin de la part du personnel et des utilisateurs du produit ?
- Pouvez-vous construire et maintenir une culture de responsabilité éthique autour de ce projet ?
- De quels protocoles de protection des données avez-vous besoin pour préserver la vie privée et vous conformer à la loi ?
- Pouvez-vous gérer et maintenir ce produit après son introduction ?
- La procédure de passation de marché est-elle équitable, transparente et concurrentielle ?
- Pouvez-vous garantir un résultat contractuel commercialement et juridiquement solide pour votre organisation et le secteur de la santé et des soins ?

Pour les volets techniques du logiciel et les questions relatives au développement et à la validation de l'outil, ce guide peut être rapproché de la « [Grille descriptive des fonctionnalités des dispositifs médicaux embarquant un système avec apprentissage automatique \(intelligence artificielle\)](#) », produite en 2020 (21). Cette grille a pour objectif de lister les informations nécessaires à fournir par les entreprises dans le cadre d'une demande de prise en charge dans le droit commun pour les dispositifs médicaux embarquant de l'intelligence artificielle et permet de donner un cadre structure à tout industriel pour qu'il décrive notamment le rôle de chaque fonctionnalité concernée, les caractéristiques des données utilisées aux différentes étapes du développement du produit, les résultats obtenus et les modalités de fonctionnement de l'algorithme. Elle aborde les points clés, regroupés par catégories, qui couvrent l'ensemble du processus de développement d'une technologie : l'usage revendiqué, les données d'apprentissage utilisées, le type d'apprentissage de l'algorithme, sa fréquence de réapprentissage, la place accordée à l'intervention humaine dans ce processus ou encore les seuils de performance. Cette grille est avant tout à visée descriptive, construite pour lever l'effet « boîte noire » des algorithmes d'intelligence artificielle. Il ne s'agit pas d'un outil de substitution à l'évaluation de l'intérêt.

La comparaison entre les grilles NHS et HAS, présentée dans le Tableau 4, met en avant qu'au-delà de la description technique de l'algorithme d'intelligence artificielle, la grille de la NHS s'intéresse plutôt à l'ensemble du cycle de vie du DM, comme les aspects réglementaires et contractuels ou encore l'intégration dans la pratique et l'impact organisationnel associé.

Tableau 3. Comparaison des domaines couverts par la grille d'évaluation des technologies avec IA du NHS avec la grille descriptive des fonctionnalités des DM embarquant un système avec IA de la HAS

	Grille NHS	Grille HAS
Utilisateurs de la grille	Professionnels de santé et organes décisionnaires d'achat des centres hospitaliers	Cible 1 ^{re} : industriels et membres de la CNEDiMTS Autres utilisateurs potentiels : professionnels de santé, établissements de santé
Finalité déclarée	Offrir un cadre d'évaluation aux potentiels utilisateurs/acheteurs avec les questions importantes à prendre en compte avant d'acheter des produits avec IA	Description objective des caractéristiques du DM intégrant une fonctionnalité apprenante Les items doivent être renseignés par l'industriel lorsqu'il dépose son dossier de demande de remboursement
Domaines couverts		
• Finalité d'usage/Réponse au besoin	X	X
• Description des données	X	X
• Description du modèle : entraînement, validation, test, mise à jour éventuelle	X	X
• Caractéristiques fonctionnelles : performance, robustesse, résilience, explicabilité	X	X
• Comparaison par rapport aux autres solutions existantes	X	
• Aspects règlementaires	X	
• Intégration dans le parcours et impact organisationnel	X	
• Interopérabilité	X	
• Acceptation de l'outil par les utilisateurs et équipes	X	
• Aspects éthiques et protection des données	X	
• Maintenance et service après-vente	X	
• Base contractuelle : contrat, propriété intellectuelle, responsabilité civile...	X	

Dans la même ligne, un groupe d'experts établi par la Commission européenne (le *High-Level Expert Group on Artificial Intelligence* (AI HLEG)) a publié en 2020 une liste de questions destinées à évaluer la fiabilité des algorithmes d'intelligence artificielle en santé nommée ALTAI, dans un but d'auto-évaluation. Parmi ces questions, on retrouve différentes catégories telles que la fiabilité des données, la confidentialité des données ou encore l'explicabilité de l'algorithme, c'est-à-dire que les décisions prises par l'algorithme peuvent être expliquées aux utilisateurs. Cette liste est issue d'un travail mené par des chercheurs australiens visant à aider les professionnels de santé à évaluer si un algorithme d'intelligence artificielle était utilisable en routine et d'identifier les difficultés techniques à résoudre. Cette liste a permis d'aboutir en 2022 au consensus de recommandations, appelé « *FUTURE-AI* », réalisé par un réseau international de chercheurs et d'experts interdisciplinaires. Ces recommandations sont organisées en six catégories : équité, universalité, traçabilité, utilisabilité, robustesse et explicabilité et sont destinées aux éditeurs, évaluateurs et régulateurs (28).

La FDA a également publié en septembre 2022 le guide « *Clinical Decision Support Software : guidance for industry and Food and Drug administration Staff* » (38). Ce guide est complémentaire de productions précédentes (39) et participe à la définition du périmètre d'action de la FDA sur les logiciels d'aide à la décision clinique, et notamment pour ceux à destination des professionnels de santé et n'ayant pas le statut de dispositif médical, mais n'a pas vocation à établir des prérequis réglementaires pour ce type d'outils. Cependant, ce guide présente des recommandations sur les informations à fournir aux professionnels de santé par les éditeurs de logiciels. L'objectif est de s'assurer que le professionnel de santé établit une décision médicale sur les recommandations établies par le logiciel, mais également sur son propre jugement et ses connaissances. Un algorithme doit donc être explicable et le professionnel de santé doit être en mesure de remettre en cause les recommandations de cet algorithme grâce à certaines informations qui doivent être fournies au professionnel. Les informations à fournir sont notamment celles relatives à la qualité et la pertinence des données patients utilisées par le logiciel ou encore le degré de personnalisation des informations fournies par le logiciel et la manière dont elles sont spécifiques à un patient et prises en compte dans l'aide à la décision proposée.

Enfin, en France, l'Agence nationale d'appui à la performance (ANAP) a développé en mai 2022 une plateforme des solutions d'intelligence artificielle en santé pour référencer des exemples de solutions numériques avec IA à destination des structures de soins ou des professionnels de santé. Ce répertoire descriptif regroupe des exemples de solutions avec IA destinées à faciliter l'organisation et la gestion des soins, le suivi des patients ou des solutions d'aide au diagnostic ou d'aide à la prise de décision. Aucune évaluation des dispositifs référencés n'est réalisée par l'ANAP dans ce cadre, dont l'objectif est de mettre en place une vitrine descriptive de solutions avec IA pouvant être déployées dans les établissements hospitaliers.

En parallèle, l'ouverture en fin 2022 de la plateforme numérique « [Bouquet de services](#) »⁴ pour les professionnels de santé peut constituer une aide au choix dans la mesure où elle vise à simplifier l'accès aux différents services numériques dès lors qu'ils respectent des critères d'interopérabilité, de sécurité et d'éthique. Dans le même objectif, aux États-Unis, un répertoire des systèmes d'aide à la décision médicale ou au diagnostic/dépistage respectant des critères d'interopérabilité et de standardisation a été mis en place par l'*Agency for Healthcare Research and Quality* (AHRQ) (40).

Aucun travail spécifique sur l'utilisation de la réalité virtuelle ou augmentée en chirurgie n'a été identifié.

⁴ G-NIUS. Bouquet de services : le bouquet de services permet aux professionnels de santé d'accéder en un seul point à plusieurs services. [En ligne] 2022. <https://gni.us.esante.gouv.fr/fr/reglementation/fiches-reglementation/bouquet-de-services>

2.3.2.6. Évaluations réalisées

La revue de la littérature met en évidence peu de travaux d'évaluation conduits par les régulateurs ou agences d'évaluation sur les DMN à usage professionnel.

De premières évaluations technologiques de ces dispositifs ont été réalisées au Royaume-Uni et au Canada et sont présentées dans le Tableau 3.

- Le NICE en particulier a produit plusieurs évaluations entre 2020 et 2022 (4 publiées) portant sur des outils d'aide au diagnostic ou d'aide à la décision médicale. Les évaluations issues du NICE, dont la finalité précise n'est pas explicitée (en pratique, information, remboursement ou bon usage), se basent sur un recensement (non exhaustif) des technologies disponibles, une analyse critique des études cliniques disponibles et s'intéresse aux différents coûts et modèles économiques proposés, associés à des avis d'experts.
- Le CADTH a conduit une évaluation sur les systèmes d'aide à la décision clinique en imagerie médicale en 2019 (41). Il indique dans son rapport que les preuves cliniques disponibles sont limitées, en raison notamment des faiblesses méthodologiques des études analysées.
- L'INESSS, dans son rapport de 2021 sur le télédepistage de la rétinopathie diabétique (TDRD) avec intelligence artificielle, a réalisé une analyse critique des différents systèmes disponibles. Ce rapport décrit en détail pour chaque dispositif : la technologie utilisée, les données cliniques associées, et les coûts que son utilisation implique. Une attention particulière a également été portée à l'évaluation de l'efficacité de ces dispositifs et à l'impact de leur intégration sur l'organisation des soins.

Rapport télédepistage de la rétinopathie diabétique à l'aide d'une solution d'intelligence artificielle, INESSS

« L'intégration de l'IA au sein d'un programme de TDRD a des impacts majeurs sur les plans technologique (par exemple, interopérabilité, pourcentage d'images non analysables), clinique (par exemple, sensibilité et spécificité), organisationnel (par exemple, accès aux soins dans les délais impartis), économique (efficacité), éthique (par exemple, équité, confidentialité et sécurité des données) et humain (par exemple, expérience des patients, expérience des cliniciens). Les données tirées de la littérature montrent que la plupart des études ne font qu'évaluer les performances techniques et cliniques de solutions d'IA destinées au dépistage de la rétinopathie diabétique, et ce, le plus souvent, dans un contexte expérimental. Les défis que pose l'utilisation de solutions d'IA destinées au dépistage de la rétinopathie diabétique dans un contexte réel de soins et services restent donc peu explorés. »

Tableau 4. Recensement des premières évaluations de logiciels à usage professionnel par des agences d'HTA

Type de technologie évaluée	Organisme évaluateur	Pays	Année
Logiciels d'aide à la décision clinique au sujet de l'imagerie médicale (41)	CADTH	Canada	2019
Logiciels d'aide au diagnostic à partir de scanners cérébraux (42)	NICE	Royaume-Uni	2020
Logiciels d'aide au diagnostic à partir de scanners thoraciques (43)	NICE	Royaume-Uni	2021
Logiciels d'aide au diagnostic à partir de mammographie (44)	NICE	Royaume-Uni	2021
Logiciels de télédépistage de la rétinopathie diabétique à l'aide d'une solution d'intelligence artificielle (45)	INESSS	Canada	2021
Logiciels d'aide au diagnostic à partir de radiographies thoraciques (46)	NICE	Royaume-Uni	2022
Logiciels d'aide à la décision médicale pour la prise en charge des accidents vasculaires cérébraux (47)	NICE	Royaume-Uni	En cours
Logiciels d'aide au diagnostic de nodules pulmonaires à partir de scanners thoraciques (48)	NICE	Royaume-Uni	En cours

2.3.2.7. Perspectives d'évaluation et propositions d'encadrement

Plusieurs productions proposant des perspectives d'évaluation ou d'encadrement des DMN à usage professionnel commencent à voir le jour.

Ainsi, le Parlement européen a publié en juin 2022 un rapport d'étude intitulé « *Artificial intelligence in healthcare: application, risks and ethical and societal impacts* ». Sept options stratégiques y sont décrites pour mieux développer, évaluer, déployer et exploiter des algorithmes d'IA techniquement, cliniquement et éthiquement solides dans les systèmes de santé. Parmi elles, sont retrouvées :

- la promotion du **co-développement** des algorithmes d'IA entre éditeurs, professionnels de santé et/ou patients ;
- la création d'un « **passport IA** » associé aux algorithmes visant à favoriser la transparence et la confiance en présentant des informations relatives au modèle, aux données utilisées, à l'évaluation (biais, limites, cas extrêmes, etc.), à l'utilisation (par exemple, distributions statistiques, (dés)accords avec les cliniciens, défaillances identifiées, utilisation de la mémoire, etc.) et à la maintenance (dernières mises à jour, version du logiciel, dernière évaluation périodique, dates, etc.) ;
- l'élaboration d'un cadre légal concernant la **responsabilité** médicale ;
- l'amélioration de la **formation** au numérique des professionnels de santé et l'information auprès du grand public dans le but de favoriser l'adoption de ces outils et de minimiser les erreurs d'utilisation ;
- la mise en œuvre d'une stratégie visant à réduire la fracture européenne en matière d'IA médicale.

Ces travaux sont à mettre en regard du rapport « *Artificial intelligence in healthcare : Benefits and Challenges of Machine Learning Technologies for Medical Diagnostics* » publié en septembre 2022 (49) par l'Académie de médecine américaine. Trois grandes perspectives stratégiques à destination des décideurs politiques y sont détaillées pour améliorer le développement des outils et limiter les difficultés adressées par les technologies diagnostiques embarquant de l'IA :

- la **production de guides et des standards de bonnes pratiques d'évaluation** de ce type de technologie par les décideurs politiques ;
- la facilitation de **l'accès à des jeux de données médicales de qualité** dans le but de développer et tester des technologies embarquant de l'IA. L'objectif des bases de données de haute qualité est de faciliter l'entraînement et les tests des algorithmes d'IA dans des conditions diverses et représentatives des conditions en vie réelle ;
- enfin, l'Académie de médecine américaine promeut la **collaboration** entre éditeurs, utilisateurs et régulateurs pour l'intégration de ces technologies dans la pratique médicale et leur évaluation à travers la **co-création** de ces outils.

En synthèse

Les principaux dispositifs médicaux à usage professionnel retrouvés dans la littérature correspondent à des **systèmes d'aide au diagnostic ou au dépistage et les systèmes d'aide à la décision médicale**.

L'analyse de la littérature, de plus en plus prolifique sur les DMN à usage professionnel, met en évidence :

- une **utilisation en pratique des DMN avec IA encore sporadique en pratique courante des soins** ;
- **l'hétérogénéité des performances** des dispositifs médicaux proposés, notamment en raison d'un potentiel manque de fiabilité ou de représentativité de la donnée d'apprentissage utilisée ;
- un **besoin de véritables études prospectives en vie réelle** venant valider et comparer les DMN ;
- **l'absence de matrice de risques utilisable en l'état pour prioriser les besoins d'évaluation** pour les DMN à usage professionnel ;
- des **premiers cadres d'évaluation** de ces dispositifs.

Les travaux sur les DMN à usage professionnel conduits par les régulateurs ou agences d'évaluation sont encore limités mais se développent. Seuls **l'INESSS et le NICE** ont à ce jour évalué certains de ces dispositifs en se basant sur une analyse critique des études cliniques disponibles et différents coûts et modèles économiques proposés. L'INESSS a notamment analysé l'efficacité de ces dispositifs et l'impact de leur intégration sur l'organisation des soins.

Le **NHS** a quant à lui **mis à disposition un guide d'aide au choix pour les utilisateurs ou acheteurs de DMN à usage professionnel avec IA**. En plus de la description des données, du modèle et des caractéristiques, ce guide permet aux utilisateurs de **se questionner sur les principaux enjeux** associés à l'utilisation de ces outils, et notamment l'intégration dans le parcours et l'impact organisationnel associé, l'éthique, l'acceptation par les équipes, la comparaison avec les autres solutions existantes, mais aussi sur des questions essentielles pour un acheteur telles que les questions de maintenance et service après-vente, d'interopérabilité, etc.

3. Enquête via les membres du réseau INAHTA

3.1. Méthodologie

Afin d'identifier le positionnement des agences d'évaluation sur le sujet des DMN à usage professionnel, une enquête au sein du réseau INAHTA – *International Network of Agencies for Health Technology Assessment* – (49 agences d'évaluation – 30 pays) a été effectuée.

Le questionnaire a été partagé avec les agences du réseau et avait pour objectif d'identifier :

- le **positionnement actuel** des agences et notamment si elles évaluent ce type de dispositifs médicaux et, si tel est le cas, les critères et les modalités d'évaluation (optionnelle ou obligatoire) en vigueur ;
- les **travaux en cours ou prévus** sur le sujet.

L'analyse s'est limitée aux données apportées par les agences ayant répondu au questionnaire.

3.2. Résultats

Au total, dix agences de sept pays ont répondu au questionnaire partagé au réseau INAHTA : G-BA et IQWiG en Allemagne, AIHTA et GoeG en Autriche, CADTH au Canada, AQuAS en Espagne, FinCCHTA en Finlande, ZonMW et ZIN aux Pays-Bas et AOTMiT en Pologne.

Parmi les organisations ayant répondu au questionnaire, la diversité des réponses témoigne de l'hétérogénéité des positionnements des agences d'évaluation sur le numérique en santé de manière générale.

L'analyse des contributions met en évidence que tous les répondants sont susceptibles de participer à l'évaluation des DMN à usage professionnel :

- sur demande du ministère de la Santé (aux Pays-Bas ; en Pologne) ;
- de manière optionnelle (DMN à usage professionnel associés à un acte à l'hôpital en Allemagne ; en cours de mise en place en Autriche ; au Canada ; en Finlande) ;
- de manière obligatoire, notamment en cas de volonté d'obtenir un remboursement (en Allemagne pour les DMN à usage professionnel associés à un acte en ville ; en Espagne en théorie).

Pourtant, à ce jour, peu d'évaluations de DMN à usage professionnel ont été déclarées par les répondants. Les répondants indiquent d'ailleurs ne pas avoir mis en place de cadre spécifique d'évaluation dédié aux DMN à usage professionnel, sans approche spécifique à l'utilisateur de la technologie numérique. Lorsque des évaluations sont réalisées, celles-ci se basent sur des cadres méthodologiques adaptés aux dispositifs médicaux quels qu'ils soient, numériques ou non.

Par exemple, en Finlande, la méthode d'évaluation Digi-HTA est utilisable pour tout type de DM numérique. L'objectif de cette évaluation est d'émettre une recommandation d'utilisation (avec un score de 0 à 10) d'une solution digitale, rendue publique. Un score ≤ 4 correspond à une recommandation de considérer attentivement avant d'utiliser le produit (une organisation envisageant le déploiement d'un produit doit considérer qu'il existe des lacunes dans un ou plusieurs des domaines clés) alors qu'un score de 10 correspond à une recommandation d'utilisation.

Lorsqu'elles sont réalisées, ces évaluations se basent essentiellement sur des critères cliniques. Cependant, plusieurs pays ont également indiqué se baser également sur des critères éthiques, légaux, sociaux (Allemagne, Espagne, Pologne), de coûts (Finlande, Espagne, Pologne) ou d'impact organisationnel (Espagne). Deux pays évaluent la sécurité et protection des données : par l'obtention d'un certificat (Pologne) ou par une évaluation d'experts (Finlande).

Des groupes de travail spécifiques sur ces sujets sont en cours dans plusieurs pays, avec des réflexions sur le développement de cadres spécifiques (Autriche, Espagne) ou la mise à disposition de guides méthodologiques à destination des utilisateurs (Pays-Bas).

L'ensemble des positions des agences sont résumées en annexe.

En synthèse

Même si tous les répondants du réseau INAHTA sont susceptibles d'évaluer les DMN à usage professionnel, en pratique, il y a eu **très peu d'évaluations par les membres du réseau à ce jour pour ce type de technologies**. Aucun partenaire du réseau INHATA n'a construit un cadre spécifique pour l'évaluation des DMN à usage professionnel. Certaines agences considèrent que les méthodes classiques d'évaluation s'appliquent. Certaines agences mettent en œuvre une réflexion sur le développement de **cadre spécifique** ou la mise à disposition de **guides méthodologiques** à destination des utilisateurs des DMN à usage professionnel.

4. Entretiens parties prenantes

4.1. Méthodologie

L'ensemble des professions médicales est aujourd'hui concerné par le numérique.

Afin d'évaluer les besoins des professionnels de santé sur le numérique et leur éventuel besoin d'aide à la sélection des dispositifs médicaux numériques à usage professionnel, des réunions bilatérales ont été conduites avec les conseils nationaux professionnels les plus concernés par le sujet, ainsi qu'avec des experts ou institutions. Une attention particulière a été portée lors du choix des spécialités contactées en vue de l'intégration de spécialités cliniques et techniques.

Ces entretiens, menés de manière informelle, avaient pour objectif d'aborder les thématiques suivantes :

- Quels types de dispositifs médicaux numériques sont utilisés par les professionnels ?
- Quelle est la finalité d'usage de ces dispositifs : aide au dépistage, aide au diagnostic, aide au traitement... ?
- Quelle est la technologie utilisée et quel est le rôle du professionnel de santé ?
- Comment se font le choix et la sélection des DMN par les professionnels de santé ?

En raison de la nature informelle des entretiens, les résultats sont présentés de manière agrégée.

4.2. Résultats

4.2.1. Résultats quantitatifs

Les conseils nationaux professionnels (CNP) sollicités ont été particulièrement réactifs et s'investissent sur les sujets du numérique.

Au total, 28 entretiens ont été menés entre avril 2022 et octobre 2022 avec la répartition suivante :

- 18 CNP ;
- 7 experts ou parties prenantes : 3 pharmaciens hospitaliers, 2 radiologues, 1 rythmologue, 1 collectif de jeunes professionnels de santé ;
- 2 structures institutionnelles ;
- 1 association d'usagers et représentants de patients.

4.2.2. Résultats qualitatifs

4.2.2.1. Une hétérogénéité de l'utilisation dans les pratiques

L'intégration des DMN à usage professionnel est très hétérogène :

- six CNP et un expert ont indiqué ne pas utiliser d'outils numériques entrant dans le champ de la saisie, leur scope et leurs besoins étant concentrés sur la télésurveillance médicale ;
- quatre CNP et un expert ont déclaré utiliser des outils de ce type de façon ponctuelle dans le cadre de recherches et d'expérimentations ;
- deux CNP ont connaissance de l'utilisation des outils numériques sans connaître le volume d'utilisation au sein de leur profession ;

- enfin, cinq CNP et un expert ont indiqué utiliser des outils numériques dans le cadre de leur pratique en routine.

Il est à noter qu'au sein d'un même CNP, l'utilisation peut être également hétérogène et cibler certaines sous-spécialités.

De manière générale, les CPN des spécialités à prédominance clinique se considèrent peu concernés par l'usage de dispositifs médicaux numériques à usage strictement professionnel. Pour ces spécialités, les usages du numérique se concentrent essentiellement sur la télésanté, en particulier la télésurveillance médicale et les applications destinées aux patients. Cette tendance s'explique également par l'exercice pratique et en particulier la combinaison nécessaire entre les éléments documentant la situation clinique du patient et la prise en compte de son ressenti des symptômes et de l'efficacité des traitements.

A contrario, les spécialités particulièrement impliquées à l'étape du diagnostic, et donc plus objectives, sont particulièrement au fait des évolutions et développements sur ces technologies. Lorsqu'ils sont utilisés, ces outils servent d'aide à l'interprétation ou à l'organisation de l'activité, notamment en termes de priorisation, et restent sous supervision humaine.

Certains CNP indiquent faire face à des sollicitations croissantes de la part des industriels pour l'utilisation de DMN à usage professionnel.

4.2.2.2. Cartographie de la finalité d'usage

Quatre finalités d'usage des DM numériques à usage strictement professionnel ont été identifiées au cours de ces entretiens : aide au diagnostic, aide au dépistage, aide à la décision médicale et aide à la validation.

Si les DMN mentionnés lors des échanges et utilisés comme aide au diagnostic se basent tous sur des technologies d'intelligence artificielle, les DMN utilisés comme aide à la décision médicale révèlent une plus grande variété dans les technologies employées, comme la réalité augmentée ou encore les applications sur smartphone. La quasi-totalité des spécialités a ainsi indiqué utiliser des applications ou outils d'aide au diagnostic ou à la décision médicale (score, échelles...) basés sur des arbres décisionnels cliniquement validés et issus de recommandations par les sociétés savantes ou autorités compétentes. Les performances de ces applications, qui n'utilisent pas d'IA et consistent en une transposition numérique d'outils existants et validés, vont dépendre de la qualité de la source d'informations, du développement technique de l'outil et de l'identification rapide des sources à l'origine des recommandations.

Concernant les modalités d'accès à la technologie, deux principaux types d'utilisation se dégagent :

- un accès direct à l'outil avec une intégration au sein du système informatique propre à l'établissement ;
- plus rarement, un accès à une plateforme centralisée, notamment en radiologie, avec une banque d'algorithmes à disposition dans certains cas. Dans ces cas, les professionnels ne semblent pas systématiquement avoir d'information quant aux performances métrologiques des algorithmes qui leur sont proposés ou qui sont utilisés par le prestataire. Ce type d'accès à la technologie n'est pas le plus répandu à ce jour : il a été mentionné par deux des CNP interrogés.

4.2.2.3. Les freins à l'intégration des dispositifs médicaux numériques à usage professionnel

De nombreux freins ont été relevés expliquant l'hétérogénéité des pratiques.

- Six CNP et experts ont mentionné une **inquiétude grandissante sur les actes associés à l'utilisation de ces outils et leur rémunération**. En effet, l'utilisation de ces DMN peut modifier ou créer de nouvelles pratiques qui ne sont pas encore associées à un acte médical et à une tarification.
- Cinq CNP ont déploré que les outils ne soient pas toujours adaptés à leur pratique et donc non adaptés aux besoins des professionnels pour plusieurs raisons. D'une part, les logiciels sont souvent très hospitalo-centrés et ne sont pas adaptés au suivi du patient dans le temps, ce qui peut expliquer le fait que les usages de DMN à usage professionnel ne soient pas développés dans certaines spécialités en libéral. D'autre part, les professionnels de santé ne sont pas toujours impliqués dans la conception des logiciels.
- L'apport réel pour le professionnel de santé est une préoccupation importante de ces CNP et également pour les patients. En effet, la volonté de maximiser certaines performances théoriques, au détriment de l'usage pratique, semble desservir l'usage des DMN. Par exemple, une sensibilité accrue d'un DMN peut engendrer un taux de faux positifs élevé et donc amener à une quantité démesurée d'alertes inutiles pour les professionnels de santé.

De même, l'association de patients et représentants des usagers a pointé du doigt « l'hyperproduction » de données qui ne sont pas toujours utiles et pertinentes dans la prise en charge et dont la quantité peut prendre le pas sur la qualité.

- Enfin, le caractère chronophage du numérique a également été rapporté par certains CNP qui voient dans le numérique une perte de temps au détriment du temps passé auprès du malade. Cinq CNP ont relevé l'**interopérabilité** comme enjeu majeur face à l'intégration des DMN à usage professionnel, en particulier à l'hôpital du fait de la multiplicité des logiciels, la plupart ciblant une spécialité ou étant spécifiques d'une pathologie et à ce jour pas systématiquement interopérables. Les systèmes d'information hospitaliers se retrouvent dans l'incapacité de gérer en parallèle la multitude de logiciels et la sécurité nécessaire pour les données de santé.
- Trois CNP ont cité le **coût élevé** et l'**absence de modèle économique** adapté pour ces solutions.
- Quatre CNP ont relevé la **fiabilité médicale** de ces outils pouvant soulever des questionnements et parfois une méfiance vis-à-vis de ces technologies. En cause, le peu ou l'absence de preuves cliniques, mais également une réflexion sur le gain réel pour le patient de ces nouvelles pratiques.
- Le collectif de jeunes professionnels a souligné l'effet « **boîte noire** » des outils embarquant de l'intelligence artificielle. En effet, les professionnels n'ont pas d'informations sur les populations de l'échantillon d'entraînement (pouvant ne pas être représentatives de leur patientèle) et déplorent le manque de transparence des éditeurs sur la précision des algorithmes ou le stockage des données. Il a été soulevé le **manque de transparence vis-à-vis du risque d'utilisation et de la façon dont l'outil a été validé**. Ceci pose des questions sur la possibilité d'extrapoler l'utilisation de ces outils d'un hôpital à l'autre qui est contraire à la volonté de standardisation des pratiques.

Cette notion a également été abordée par l'association de patients et représentants des usagers. Le défaut d'explicabilité des résultats et de transparence sur l'utilisation des données recueillies par un DMN peut entraîner une rupture dans l'utilisation de ces outils par refus du

patient. Le défaut d'information entraîne une défiance qui peut nuire au dialogue entre patients et professionnels de santé.

D'autres questionnements ont également été abordés par les CNP :

- des questionnements à clarifier sur la responsabilité médicale associée à l'utilisation de ces DMN ;
- une question éthique sur les nouvelles informations « non attendues » que peut générer l'utilisation de DMN à usage professionnel et de la manière dont les professionnels de santé peuvent les traiter. Par exemple, une anomalie biologique détectée par un algorithme d'IA qui n'était pas identifiable auparavant ;
- l'émergence de sociétés pouvant démarcher directement des patients pour l'interprétation de leurs résultats ;
- six CNP ont insisté sur la place des professionnels de santé face aux changements profonds apportés par ces outils numériques et sur l'avenir de leur profession, même si le numérique peut être vu comme un levier pour une meilleure organisation des parcours ville-hôpital avec une potentielle délégation de tâches vers des paramédicaux ;
- dans la même ligne, l'association de patients et représentants des usagers a insisté sur l'importance de renforcer le dialogue entre patients et professionnels de santé malgré les changements profonds apportés par ce type d'outils. Il est important pour les professionnels de santé d'être transparents sur leur interprétation des résultats apportés par ceux-ci. Par exemple, si un professionnel de santé émet un diagnostic différent de celui du logiciel d'aide au diagnostic utilisé, celui-ci doit pouvoir expliquer son interprétation auprès du patient. L'utilisation de ces outils doit donc rester sous supervision humaine mais la relation entre patients et professionnels de santé se doit d'être protégée ;
- enfin, l'association de patients et représentants des usagers a relevé la nécessité de recueillir le consentement des patients dans un souci d'améliorer l'information et la transparence dans l'utilisation des données des patients.

Finalement, ces freins révèlent l'enjeu majeur de **l'organisation de l'intégration des DMN à usage professionnel** et les nombreux questionnements auxquels font face les professionnels de santé : comment intégrer les DMN dans la pratique professionnelle, où les place-t-on et comment réorganise-t-on les soins ?

4.2.2.4. Le choix et la sélection des dispositifs médicaux numériques à usage professionnel

À l'image de l'hétérogénéité des pratiques, le choix et la sélection des DMN par les professionnels sont très variables.

La plupart des DMN étant développés pour une seule spécialité, les acteurs à l'origine du choix sont les services hospitaliers concernés eux-mêmes. Ainsi, les deux experts pharmaciens interrogés ont indiqué ne pas être sollicités ou ne pas avoir d'actions particulières dans l'aide au choix de ces dispositifs, qui sont développés essentiellement par les services sans intervention de leur part.

Cette sélection des DMN se base sur des critères de choix qui apparaissent comme peu formalisés dans les différentes organisations rapportées et peu véritablement fondés sur la démonstration clinique.

Les critères les plus cités sont les suivants :

- **ergonomie** de la solution ;
- **performances de l'outil annoncées par l'industriel**, parfois complétées par des tests pratiques sur site ;
- **recommandations entre utilisateurs** ;
- **prix** et **nombre de licences**.

Les CNP interrogés ont souligné la difficulté de guider leur choix sur la médecine fondée sur les preuves en raison de la rareté des données disponibles, et notamment dans un contexte d'utilisation en vie réelle. Ainsi, le faible nombre d'études cliniques robustes, l'évolution rapide des technologies face à leur évaluation trop lente ou encore la difficulté d'extrapolation des données d'un hôpital à l'autre étaient souvent mentionnés par les CNP.

Par ailleurs, de nombreux CNP ont mis en avant les mises à jour régulières des logiciels ou DMN, qui posent la question de la validité dans le temps d'une évaluation à un instant T.

Pour pallier le manque de données cliniques disponibles, des CNP se sont organisés ou sont en cours d'organisation pour évaluer les DMN à usage professionnel. Certains CNP sont déjà autonomes dans l'évaluation des DMN à usage professionnel. D'autres mettent en place des études observationnelles ou suivent les performances de leurs outils en vie réelle afin de s'assurer du bon paramétrage des logiciels. Enfin, certains mènent des réflexions sur l'évaluation de ces solutions en se basant sur des critères cliniques, techniques, d'intégration, de prix et de maintenance.

4.2.2.5. Le positionnement attendu de la Haute Autorité de santé

Les CNP ont été interrogés spécifiquement sur leurs attentes vis-à-vis de la HAS sur le sujet des DMN à usage professionnel :

- treize CNP, cinq experts, le collectif de jeunes professionnels et l'association de patients et représentants des usagers seraient favorables à un positionnement de la HAS sur ces sujets ;
- trois CNP ne souhaitent pas que la HAS intervienne dans le choix des outils numériques ;
- deux CNP, un expert et deux structures institutionnelles ne se sont pas prononcés ou n'ont pas été interrogés sur ce sujet.

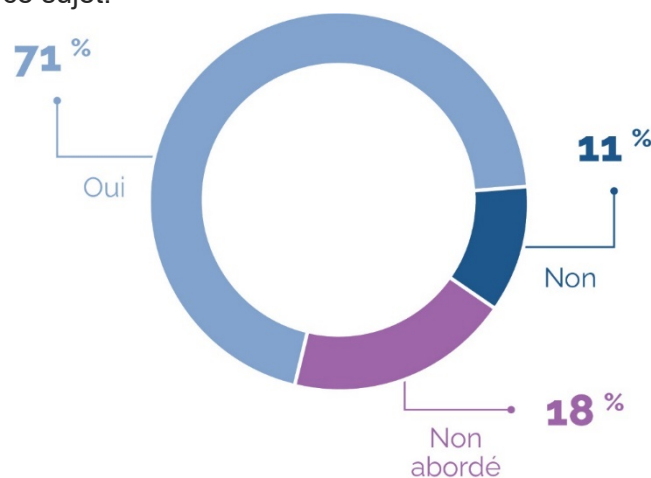


Figure 5. Réponses de CNP à la question « Souhaitez-vous que la HAS s'empare du sujet des DMN à usage professionnel ? »

Parmi les CNP favorables à un positionnement d'aide de la HAS, quatre grands types d'actions différents ont été soulignés par les CNP :

- la mise en place d'un guide d'aide au choix ou de référentiels généraux pour ces outils ;
- une aide à la veille technologique et également l'expression des besoins des professionnels de santé en se basant sur leur pratique du terrain, pour faciliter le repérage en amont des technologies pertinentes ;
- la mise en place d'une aide méthodologique de la HAS dans le cadre d'un partenariat avec les CNP, notamment avec la production de guides méthodologiques et/ou un soutien personnalisé lors de la mise en place d'études en vie réelle pour évaluer les DMN ;
- la mise en place de « labels » ou tout autre dispositif de reconnaissance.

4.3. Discussion

L'intégration des DMN à usage professionnel est encore au début de son développement et reste très variable selon les spécialités et au sein de mêmes spécialités médicales selon les professionnels. Par ailleurs, certains CNP ont particulièrement eu du mal à se projeter dans des exemples de dispositifs médicaux numériques évoqués dans la saisine. Ceci témoigne notamment de l'hétérogénéité de l'apparition de ces technologies et des pratiques. Le taux de recours à l'utilisation de ces outils à l'échelle nationale reste peu connu. L'utilisation est très liée aux moyens financiers et matériels des hôpitaux ou structures médicales qui sont inégaux au sein du territoire français.

Les CNP font face à des évolutions de leur profession et sont très concernés par les sujets du numérique. Les entretiens ont permis de faire émerger les sujets prioritaires pour les CNP. **Peu de CNP se questionnent sur la performance des DMN à usage professionnel. Quelques-uns ont toutefois déjà organisé leur évaluation.** Pour de nombreux CNP, la problématique est tournée avant tout sur le déploiement de la télésanté (hors périmètre de ce projet), l'absence de modèle économique pour intégrer les logiciels dans les pratiques, la responsabilité médicale et la nécessité d'une révision des tarifs des actes faisant appel à l'utilisation de ce type de technologie.

L'utilisation des DMN à usage professionnel étant à ses débuts, de nombreux CNP ont souligné l'importance de ne pas freiner l'accès à l'innovation et de laisser se développer les usages et les pratiques associées. La variation dans les réponses sur le positionnement attendu de la HAS ouvre la voie à plusieurs scénarios qui peuvent être discutés et travaillés.

En synthèse

L'utilisation des DMN à usage professionnel est très **hétérogène** entre les CNP. Dans le périmètre des DMN visés dans ce projet (excluant la télésanté et les LAP/LAD en particulier), les principaux usages sont orientés vers les **systèmes d'aide au diagnostic et au dépistage** utilisés par les spécialités à prédominance diagnostique. L'intégration de ces outils est également dépendante de leur acceptation par les patients. Pour cela, l'information sur l'utilisation des données et le renforcement du dialogue entre patients et professionnels de santé sont primordiaux pour garantir la confiance en ces outils.

La sélection de ces outils par les professionnels n'est pas encadrée et dépend de l'appétence de la profession pour le numérique et de l'intégration dans les pratiques de ces outils.

71 % des CNP interrogés ont indiqué être favorables à une aide de la HAS avec des attentes variables, décrites par niveau d'implication décroissant :

- la mise en place de « labels » ou équivalents ;
- la mise en place d'une aide méthodologique de la HAS dans le cadre d'un partenariat avec les CNP, notamment avec la production de guides méthodologiques ou un soutien personnalisé lors de la mise en place d'études en vie réelle pour évaluer les DMN ;
- la mise en place de référentiels pour ces outils ;
- une aide à la veille technologique ou une aide au préchoix pour repérer en amont ces outils et les proposer ensuite aux professionnels de santé en se basant sur leur pratique du terrain.

5. Perspectives

L'analyse de la littérature et des réunions bilatérales d'échange avec les professionnels met en évidence la forte disparité des attentes vis-à-vis de la HAS selon les spécialités, dont les besoins d'aide au choix des DMN apparaissent différents. Face à ce constat, la HAS propose divers scénarios concernant les positionnements possibles de la HAS dans la promotion du bon usage de ces DMN. La construction de ce projet avec les parties prenantes, pouvant passer par un appel à contributions, permettra de faire émerger d'autres productions possibles et/ou de structurer ces premières pistes de réflexion. Les différents scénarios identifiés ci-dessous pourront être combinés. Ils ne sont pas forcément exclusifs et seront explorés dans le cadre des travaux qui démarrent.

5.1. Production à visée informative pour les professionnels : guide d'aide au choix des dispositifs médicaux numériques à usage professionnel

Ce guide généraliste aura pour objectif d'orienter les utilisateurs et des acheteurs de DMN à usage professionnel dans leur choix en leur permettant d'identifier les **principales questions clés à se poser sur le DMN** :

- son développement : provenance et fiabilité de l'information disponible, populations sources dont sont issues les données utilisées ;
- la validation de l'outil et les performances associées ;
- l'intégration dans les parcours ;
- le modèle économique ;
- les questions de maintenance, de devenir des données...

Ces questions clés pourraient être accompagnées d'une **liste de prérequis nécessaires** sur l'interopérabilité avec les systèmes d'information, le marquage CE, l'ergonomie du DMN, le stockage des données...

La production d'un guide d'aide au choix de ces dispositifs correspond notamment au positionnement actuel du NHS sur le sujet des DMN à usage professionnel.

Avantages

Ce guide d'aide au choix serait généraliste et commun à l'ensemble des spécialités médicales. Il pourrait donc permettre de couvrir le champ large des DMN à usage professionnel avec une mise en œuvre rapide.

Il constituerait une aide directe d'aide au choix pour les professionnels de santé ou les établissements par une structuration des informations clés requises pour éclairer leurs choix.

Inconvénient

Ce guide aurait un caractère forcément généraliste et ne constituerait pas une évaluation.

5.2. Productions socles permettant la structuration d'une évaluation

Deux scénarios ont été envisagés, à ce stade, pour la structuration d'une démarche d'évaluation, mais de nouveaux scénarios pourront émerger au fil des travaux.

5.2.1. Scénario 1 : la mise en place d'une méthode collaborative d'évaluation, en appui des CNP

La mise en place d'une méthode collaborative pourrait prendre différents cadres possibles :

1. la mise à disposition par la HAS d'un **guide généraliste sur les méthodes d'évaluation adaptées** pour les DMN à usage professionnel ;
2. la **mise en place de veilles technologiques collaboratives** par la HAS et le CNP ;
3. un **appui méthodologique** spécifique pour les CNP souhaitant réaliser des **évaluations des performances** annoncées des technologies ;
4. un **appui méthodologique** spécifique pour les CNP souhaitant réaliser des **évaluations de l'apport de la technologie par rapport à la pratique sans ces technologies** ;
5. un **appui méthodologique** spécifique pour les CNP souhaitant **valider le déploiement national** de ces technologies.

Les évaluations mises en œuvre par les CNP pourraient consister en :

- une **(pré-)sélection des algorithmes** avec intelligence artificielle sur un **jeu de données indépendantes, labellisées par les CNP** ;
- une **évaluation basée sur la mise en place de sites pilotes** pour le **test en vie réelle des outils pendant plusieurs mois pour chaque nouvelle version des logiciels** ;
- une évaluation basée sur la **mobilisation de l'ensemble de la profession à l'échelle nationale pour l'évaluation sur la base du volontariat**.

La possibilité de créer des partenariats entre plusieurs CNP a été évoquée pour mutualiser les expériences et réaliser une évaluation collaborative de ces solutions avec l'appui de la HAS.

Avantage

Cette approche personnalisée permettrait d'accompagner les besoins des CNP de manière individualisée en impliquant l'ensemble de l'écosystème.

Inconvénient

La question de la faisabilité et des moyens à mobiliser pour la HAS reste à explorer.

5.2.2. Scénario 2 : la mise en place d'une labellisation (ou équivalent) ou de référentiels d'évaluation par la HAS pour les DMN à usage professionnel

Même si aucune demande d'évaluation classique des DMN à usage professionnel n'a été formulée par les CNP, la réalisation de référentiels d'évaluation ou la mise en place d'un dispositif volontaire de labellisation ou équivalent sur des critères cliniques et organisationnels ont pu être demandées par certains CNP ou experts. Les modalités restent cependant à définir. Si cette démarche était mise en place, elle devrait être testée sur les catégories de DMN soulevant le plus de questionnements en termes de validation et pourrait s'appuyer sur une matrice de risques qui serait à construire avec les professionnels.

Avantage

Ce scénario pourrait répondre au besoin spécifique de certains CNP et donne un éclairage clés en main pour les professionnels demandeurs.

Inconvénients

Au vu des divergences de points de vue entre CNP (et au sein de plusieurs CNP), un accueil mitigé par l'écosystème est possible.

Cette évaluation ne devra pas être un frein au développement des innovations.

L'obtention d'un label aura un coût pour les entreprises avec un impact sur les prix des DMN qui seraient labellisés, dans un contexte où le modèle économique des DMN est encore à trouver.

La question de la faisabilité et des moyens à mobiliser par la HAS reste à explorer compte tenu du nombre de technologies et de leur *turnover* et des compétences nécessaires.

En synthèse

Deux principaux types de productions semblent se dégager :

- la mise à disposition de productions à destination des professionnels ou structures de soins, comme les **guides d'aide au choix** des DMN à usage professionnel généralistes ;
- la mise en place de productions permettant de construire un socle d'évaluation pour les DMN à usage professionnel, et notamment :
 - la mise en place de **démarches collaboratives** avec les CNP, pour accompagner méthodologiquement les CNP dans l'évaluation des DMN à différents degrés,
 - la mise en place d'une **labellisation (ou équivalent) ou la production de référentiels d'évaluation** pour les DMN à usage professionnel, directement par la HAS pour les CNP demandeurs.

6. Conclusion

L'état des lieux réalisé dans le cadre de cette production met en avant la nécessité du positionnement de la HAS et des pouvoirs publics pour promouvoir le bon usage des DMN à usage professionnel. Il est nécessaire de construire un cadre de confiance dédié aux DMN à usage professionnel, afin de favoriser l'intégration de ceux qui sont utiles et innovants dans le système de santé, sans être un frein à l'innovation.

Comme explicité dans [la note de cadrage](#) du projet, plusieurs pistes de réflexion ont émergé au cours de ces travaux préliminaires, dont les modalités pratiques seront à définir.

À courte échéance et afin d'accompagner les professionnels de santé et les structures de soins dans la sélection de DMN, la HAS souhaite publier le guide d'aide au choix des DMN à usage professionnel précédemment évoqué. Il permettra d'orienter les utilisateurs et les acheteurs en leur permettant d'identifier les principales questions clés à se poser ou à poser sur le DMN (développement, validation de l'outil et performances associées, intégration dans les parcours, modèle économique, questions de maintenance, de devenir des données, etc.). D'autres productions, comme la création d'une matrice de risques spécifique aux DMN à usage professionnel, pourraient également être discutées selon les besoins identifiés lors des échanges à venir avec les parties prenantes. La réalisation de ce type de production reste toutefois complexe à mettre en œuvre, du fait de la notion de risque qui est très hétérogène selon les spécialités médicales, et nécessiterait l'implication de l'ensemble des parties prenantes.

Selon les besoins qui seront identifiés en termes d'évolution de l'évaluation de ce type de technologies, la HAS pourra émettre de nouvelles propositions aux pouvoirs publics. Ces réflexions autour de l'évaluation pourront notamment être impulsées dans le cadre de plans ou expérimentations nationaux.

Table des annexes

Annexe 1. Résumé des réponses des agences d'évaluation du réseau INAHTA au questionnaire partagé

44

Annexe 1. Résumé des réponses des agences d'évaluation du réseau INAHTA au questionnaire partagé

Pays	Position sur les DMN à usage professionnel
Allemagne (G-BA et IQWiG)	L'Allemagne évalue les DMN à usage professionnel s'ils sont associés à un acte diagnostique ou thérapeutique. Ce sont essentiellement des DMN IIb ou III. Cette évaluation est obligatoire en ville pour le remboursement mais optionnelle dans le cadre de soins à l'hôpital. L'IQWiG sera chargé de l'évaluation en se basant sur des critères d'évaluation clinique comparative (<i>comparative effectiveness (i.e. additional patient-relevant benefit)</i>). Le rapport produit par l'IQWiG sera ensuite utilisé par le G-BA, en complément de critères éthiques, sociaux et juridiques pour statuer sur la prise en charge de la solution par la collectivité et négocier son prix avec l'industriel.
Autriche (AIHTA et GoeG)	L'AIHTA est en train de mettre en place une évaluation optionnelle des DMN à usage professionnel dans un objectif de visibilité sur les outils utilisés et de catégorisation (pas dans un cadre de remboursement). Les critères d'évaluation seront cliniques, issus de RCT. L'AIHTA envisage une méthode d'évaluation standard basée sur l'analyse PICO. À la date de réponse au questionnaire, aucune évaluation n'a été faite sur le sujet et les DMN à usage professionnel sont choisis à l'hôpital directement par les services utilisateurs.
Canada (CADTH)	<p>Le périmètre d'évaluation actuelle du CADTH concerne les dispositifs numériques qui sont utilisés à des fins médicales susceptibles d'avoir un impact direct sur la prévention, le diagnostic ou le traitement d'une maladie ou d'une affection, y compris les logiciels (par exemple, les algorithmes d'intelligence artificielle, les applications mobiles).</p> <p>Les méthodes d'évaluation utilisées sont les méthodes classiques d'HTA. Des travaux sont toutefois en cours sur une éventuelle mise à jour des méthodes et processus d'évaluation pour adapter les méthodologies aux DMN. L'évaluation est facultative (non nécessaire pour l'intégration ou le remboursement des dispositifs) et a pour but d'éclairer le décideur public.</p>
Espagne (AQuAS)	L'AQuAS déclare en théorie vouloir évaluer l'ensemble des DMN entrant dans le périmètre suivant : <i>mHealth</i> (applications de santé), robotique, systèmes d'aide à la décision clinique, services de télémédecine, algorithmes. Cependant, en pratique, l'AQuAS indique ne pas réaliser d'évaluation complète pour l'intégralité de ces DMN. Cette évaluation se base sur des critères de sécurité, efficacité/efficience, impact organisationnel, éthique, légal et social. Un cadre spécifique dédié aux DMN à usage professionnel est en réflexion, notamment quant aux domaines à prendre en compte pour l'évaluation et le niveau d'exigence attendu pour les études. Sur son site internet, l'AQuAS indique procéder à une adaptation du cadre dédié aux DMN du NICE et avoir mis en place un groupe de travail avec le NICE.
Finlande (FinCCHTA)	À ce jour, le FinCCHTA évalue des DM numériques (de type application mobile ou web application) destinés au recueil des symptômes par les patients eux-mêmes ou à la mesure de <i>Patient Reported Outcomes</i> (PRO) dans le cadre du suivi à des fins d'échange avec les professionnels de santé. L'agence annonce

qu'elle pourrait évaluer tout type de service numérique, qu'il s'agisse d'une application, d'un robot ou d'intelligence artificielle. La méthode d'évaluation dédiée aux DMN est la Digi-HTA. L'objectif de cette évaluation est d'émettre une recommandation d'utilisation (avec un score de 0 à 10) d'une solution digitale, rendue publique. Un score ≤ 4 correspond à une recommandation de considérer attentivement avant d'utiliser le produit (une organisation envisageant le déploiement d'un produit doit considérer qu'il existe des lacunes dans un ou plusieurs des domaines clés) alors qu'un score de 10 correspond à une recommandation d'utilisation.

Les recommandations sont basées sur les informations fournies par les entreprises sur leur produit, qui sont complétées par une revue de la littérature, des évaluations d'experts et d'éventuelles questions supplémentaires pour les entreprises. Cette évaluation est optionnelle et se base sur des critères d'efficacité, de coût, de sécurité, de confidentialité et protection des données, de facilité d'utilisation et de convivialité/accessibilité. Même si la Digi-HTA peut s'appliquer aux outils avec IA, aucun DMN à usage professionnel n'a encore fait l'objet d'une évaluation dans ce cadre.

Pays-Bas (ZonMW et ZIN)	<p>Les DMN ne sont pas évalués aux Pays-Bas, sauf exceptions (notamment si demande particulière du ministère ou d'une société savante).</p> <p>Le ZIN, notamment en charge de l'évaluation des produits de santé aux Pays-Bas, élabore actuellement un guide méthodologique à destination des professionnels de santé et des acheteurs pour les accompagner dans la génération de preuve et dans l'évaluation clinique de la valeur ajoutée par les DMN (notamment ceux avec IA).</p> <p>Le ZonMW, responsable de la recherche en santé aux Pays-Bas, développe de son côté un guide méthodologique pour les études cliniques et médico-économiques à destination des industriels des technologies de santé, et donc non spécifique aux DMN à usage professionnel.</p>
Pologne (AOTMiT)	<p>L'<i>Agency of Health Technology Assessment and Tariff System</i> (AOTMiT) conduit des évaluations en réponse aux demandes du ministère de la Santé. Elles peuvent concerner tout type de DMN. Son évaluation est obligatoire pour obtenir un remboursement et se base sur des critères d'efficacité, de sécurité, de coûts et d'impact budgétaire. Un certificat concernant la sécurité des données est également exigé (sans évaluation par l'AOTMiT). Pour les solutions déjà évaluées, l'AOTMiT précise qu'en raison de l'absence de données suffisantes ou de leur manque de robustesse, la plupart des DMN évalués ont reçu un avis négatif ou des études pivots supplémentaires ont été demandées afin de collecter les données nécessaires. Parmi ces DMN, un seul DMN à usage professionnel (aide au diagnostic précoce d'infarctus à partir d'EKG) a été évalué et a reçu une recommandation négative de l'agence.</p>

Références bibliographiques

1. Haute Autorité de santé. Numérique : quelle (R) évolution ? Rapport d'analyse prospective 2019. Saint-Denis La Plaine: HAS; 2019.
https://www.has-sante.fr/jcms/p_3223636/fr/numerique-quelle-r-evolution-rapport-d-analyse-prospective-2019
2. Haute Autorité de santé. Classification fonctionnelle, selon leur finalité d'usage, des solutions numériques utilisées dans le cadre de soins médicaux ou paramédicaux. Saint-Denis La Plaine: HAS; 2021.
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
3. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. *JAMA* 2018;320(21):2199-200.
<http://dx.doi.org/10.1001/jama.2018.17163>
4. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med* 2020;3:17.
<http://dx.doi.org/10.1038/s41746-020-0221-y>
5. Setio AAA, Traverso A, de Bel T, Berens MSN, Bogaard CVD, Cerello P, *et al.* Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Med Image Anal* 2017;42:1-13.
<http://dx.doi.org/10.1016/j.media.2017.06.015>
6. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional Neural Networks. *Radiology* 2017;284(2):574-82.
<http://dx.doi.org/10.1148/radiol.2017162326>
7. Becker AS, Marcon M, Ghafoor S, Wurnig MC, Frauenfelder T, Boss A. Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. *Invest Radiol* 2017;52(7):434-40.
<http://dx.doi.org/10.1097/rli.0000000000000358>
8. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542(7639):115-8.
<http://dx.doi.org/10.1038/nature21056>
9. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, *et al.* Corrigendum: Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;546(7660):686.
<http://dx.doi.org/10.1038/nature22985>
10. Olczak J, Fahlberg N, Maki A, Razavian AS, Jilert A, Stark A, *et al.* Artificial intelligence for analyzing orthopedic trauma radiographs. *Acta Orthop* 2017;88(6):581-6.
<http://dx.doi.org/10.1080/17453674.2017.1344459>
11. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, *et al.* Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016;316(22):2402-10.
<http://dx.doi.org/10.1001/jama.2016.17216>
12. Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, *et al.* Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019;68(10):1813-9.
<http://dx.doi.org/10.1136/gutjnl-2018-317500>
13. Kulkarni S, Seneviratne N, Baig MS, Khan AHA. Artificial intelligence in medicine: where are we now? *Acad Radiol* 2020;27(1):62-70.
<http://dx.doi.org/10.1016/j.acra.2019.10.001>
14. Ewurum CH, Guo Y, Pagnha S, Feng Z, Luo X. Surgical navigation in orthopedics: workflow and system review. *Adv Exp Med Biol* 2018;1093:47-63.
http://dx.doi.org/10.1007/978-981-13-1396-7_4
15. Stead WW. Clinical Implications and challenges of artificial intelligence and deep learning. *JAMA* 2018;320(11):1107-8.
<http://dx.doi.org/10.1001/jama.2018.11029>
16. Maddox TM, Rumsfeld JS, Payne PRO. Questions for artificial intelligence in health care. *JAMA* 2019;321(1):31-2.
<http://dx.doi.org/10.1001/jama.2018.18932>
17. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019;17(1):195.
<http://dx.doi.org/10.1186/s12916-019-1426-2>
18. Adamson AS, Smith A. Machine learning and health care disparities in Dermatology. *JAMA dermatology* 2018;154(11):1247-8.
<http://dx.doi.org/10.1001/jamadermatol.2018.2348>
19. Du-Harpur X, Watt FM, Luscombe NM, Lynch MD. What is AI? Applications of artificial intelligence to dermatology. *Br J Dermatol* 2020;183(3):423-30.
<http://dx.doi.org/10.1111/bjd.18880>
20. Zaar O, Larson A, Polesie S, Saleh K, Tarstedt M, Olives A, *et al.* Evaluation of the diagnostic accuracy of an online artificial intelligence application for skin disease diagnosis. *Acta Derm Venereol* 2020;100(16):adv00260.
<http://dx.doi.org/10.2340/00015555-3624>
21. Haute Autorité de santé. Grille descriptive des fonctionnalités des dispositifs médicaux embarquant un système avec apprentissage automatique (intelligence artificielle). Saint-Denis La Plaine: HAS; 2020.
https://www.has-sante.fr/jcms/p_3318028/fr/grille-descriptive-des-fonctionnalites-des-dispositifs-medicaux-embarquant-un-systeme-avec-apprentissage-automatique-intelligence-artificielle

22. Freeman K, Geppert J, Stinton C, Todkill D, Johnson S, Clarke A, *et al.* Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy. *BMJ* 2021;374:n1872. <http://dx.doi.org/10.1136/bmj.n1872>
23. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, *et al.* Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ* 2020;368:m689. <http://dx.doi.org/10.1136/bmj.m689>
24. Dhiman P, Ma J, Andaur Navarro CL, Speich B, Bullock G, Damen JAA, *et al.* Methodological conduct of prognostic prediction models developed using machine learning in oncology: a systematic review. *BMC Med Res Methodol* 2022;22(1):101. <http://dx.doi.org/10.1186/s12874-022-01577-x>
25. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, *et al.* A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health* 2019;1(6):e271-e97. [http://dx.doi.org/10.1016/s2589-7500\(19\)30123-2](http://dx.doi.org/10.1016/s2589-7500(19)30123-2)
26. Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nat Med* 2021;27(4):582-4. <http://dx.doi.org/10.1038/s41591-021-01312-x>
27. Conseil d'Etat. Intelligence artificielle et action publique : construire la confiance, servir la performance. Etude adoptée en assemblée générale plénière du 31/03/2022. Paris: CE; 2022. <https://www.conseil-etat.fr/Media/actualites/documents/2022/08-aout/etude-a-la-demande-du-pm-ia-et-action-publique.pdf>
28. European Parliamentary Research Service. Artificial intelligence in healthcare. Applications, risks, and ethical and societal impacts. Brussels: EU; 2022. [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729512/EPRS_STU\(2022\)729512_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729512/EPRS_STU(2022)729512_EN.pdf)
29. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, *et al.* Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med* 2022;28(5):924-33. <http://dx.doi.org/10.1038/s41591-022-01772-9>
30. Food Drug and Administration. Software as a Medical Device (SaMD). Silver Spring: FDA; 2018. <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd>
31. Food Drug and Administration. Digital Health Software Precertification (Pre-Cert) Program. Silver Spring: FDA; 2021. <https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-software-precertification-pre-cert-program>
32. Baumfeld Andre E, Carrington N, Siami FS, Hiatt JC, McWilliams C, Hiller C, *et al.* The current landscape and emerging applications for real-world data in diagnostics and clinical decision support and its impact on regulatory decision making. *Clin Pharmacol Ther* 2022. <http://dx.doi.org/10.1002/cpt.2565>
33. National Institute for Health and Care Excellence. Evidence standards framework (ESF) for digital health technologies [En ligne]. London: NICE; 2021. <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies>
34. 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/jcms/c_2681915/fr/referentiel-de-bonnes-pratiques-sur-les-applications-et-les-objets-connectes-en-sante-mobile-health-ou-mhealth
35. NHSx. A buyer's guide to AI in health and care. London: NHS; 2020. <https://www.nhsx.nhs.uk/ai-lab/explore-all-resources/adopt-ai/a-buyers-guide-to-ai-in-health-and-care/a-buyers-guide-to-ai-in-health-and-care/>
36. NHSx. AI buyer's guide assessment template. London: NHS; 2022. <https://www.nhsx.nhs.uk/ai-lab/explore-all-resources/adopt-ai/a-buyers-guide-to-ai-in-health-and-care/ai-buyers-guide-assessment-template/>
37. National Institute for Health and Care Excellence. NHS AI Lab roadmap [En ligne]. London: NICE; 2021. <https://www.nhsx.nhs.uk/ai-lab/nhs-ai-lab-roadmap/>
38. Food and Drug Administration. Clinical decision support software. Guidance for Industry and Food and Drug Administration Staff. Silver Spring: FDA; 2022. <https://www.fda.gov/media/109618/download>
39. Food and Drug Administration. Changes to existing medical software policies resulting from section 3060 of the 21st century cures act. Guidance for Industry and Food and Drug Administration Staff. Silver Spring: FDA; 2019. <https://www.fda.gov/media/109622/download>
40. Agency for Healthcare Research and Quality. Clinical decision support [En ligne] 2022. <https://cds.ahrq.gov/cdsconnect/faq>
41. Canadian Agency for Drugs and Technologies in Health, Edge R, Ford C. Clinical decision support systems for appropriate medical imaging: Clinical evidence and cost effectiveness. Ottawa: CADTH; 2019. <https://www.cadth.ca/fr/les-systemes-daide-la-decision-clinique-au-sujet-de-limagerie-medicale-appropriee-donnees-probantes>
42. National Institute for Health and Care Excellence. Artificial intelligence for analysing CT brain scans. London: NICE; 2020. <https://www.nice.org.uk/advice/mib292>
43. National Institute for Health and Care Excellence. Artificial intelligence for analysing chest CT images. London: NICE; 2021. <https://www.nice.org.uk/advice/mib243>
44. National Institute for Health and Care Excellence. Artificial intelligence in mammography. London: NICE; 2021. <https://www.nice.org.uk/advice/mib242>

45. Institut national d'excellence en santé et en services sociaux (INESSS), Mombo NN, Arbour S, Brabant J, Alami H. Télédépistage de la rétinopathie diabétique à l'aide d'une solution d'intelligence artificielle. Québec: INESSS; 2021.

https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/Technologies/INESSS_IA_TDRD_EC.pdf

46. National Institute for Health and Care Excellence. Artificial intelligence for analysing chest X-ray images. London: NICE; 2022.

<https://www.nice.org.uk/advice/mib292>

47. National Institute for Health and Care Excellence. Artificial intelligence (AI) software to help clinical decision making in stroke [statut In progress]. London: NICE; 2022.

<https://www.nice.org.uk/guidance/indevelopment/gid-dg10044>

48. National Institute for Health and Care Excellence. Software with artificial intelligence derived algorithms for automated detection and analysis of lung nodules from CT scan images. In development [GID-DG10046]Expected publication date: 28 June 2023 [En ligne] 2023.

<https://www.nice.org.uk/guidance/indevelopment/gid-dg10046>

49. Government Accountability Office (GAO), National Academy of Medicine (NAM). Artificial intelligence in health care benefits and challenges of machine learning technologies for medical diagnostics. Washington: GAO; NAM; 2022.

<https://www.gao.gov/assets/gao-22-104629.pdf>

Participants

Les organismes professionnels, institutions et experts suivants ont participé aux réunions bilatérales informelles entre le 26 avril 2022 et le 27 juin 2022 :

Agence nationale d'appui à la performance	Conseil national professionnel en dermato-vénéréologie
Ateliers Mercure	Conseil national professionnel d'hématologie
Collège de médecine générale	Conseil national professionnel de radiologie et d'imagerie médicale (G4)
Conseil national professionnel d'anatomie et cytologie pathologiques (CNPath)	Conseil national professionnel de pneumologie (Fédération française de pneumologie)
Conseil national professionnel de biologie médicale	Conseil national professionnel de médecine d'urgence (Collège français de médecine d'urgence)
Conseil national professionnel de médecine nucléaire	Institut national du cancer
Conseil national professionnel d'ophtalmologie (Académie française de l'ophtalmologie)	Pr Jean-Paul Bérégi, CHU de Nîmes
Conseil national professionnel de rhumatologie	Pr Pierre-Yves Chambrin, AP-HP
Conseil national professionnel de chirurgie orthopédique et traumatologique (CNP-COT)	Pr Pierre Champsaur, AP-HM, président du directoire DRIM France IA
Conseil national professionnel d'ORL et CCF (CNP ORL)	Pr Pascal Defaye, CHU Grenoble Alpes
Conseil national professionnel cardiovasculaire	Dr Charlotte Gourio, CHU de Caen, membre de la Commission nationale d'évaluation des dispositifs médicaux et des technologies de santé (CNEDiMTS)
Conseil national professionnel endocrinologie diabétologie nutrition	Dr Xavier Dode, Hospices civils de Lyon, président du CHNIM
Conseil national professionnel de psychiatrie	
Conseil national professionnel d'hépatogastro-entérologie	
Conseil national professionnel d'oncologie	

Les représentants de patients suivants ont également été sollicités pour participer à ces travaux :

France Assos Santé
Conseil pour l'engagement des usagers de la Haute Autorité de santé

Remerciements

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

Abréviations et acronymes

AHRQ	<i>Agency for Healthcare Research and Quality</i>
AIHTA	<i>Austrian Institute for Health Technology Assessment, Autriche</i>
ANAP	Agence nationale d'appui à la performance
AOTMiT	<i>Agency for Health Technology Assessment and Tariff System, Pologne</i>
AQuAS	<i>Agència de Qualitat i Avaluació Sanitàries de Catalunya, Espagne</i>
CADTH	<i>Canadian Agency for Drugs & Technologies in Health</i>
CNP	Conseil national professionnel
DM	Dispositif médical
DMN	Dispositif médical numérique
DSS	Direction de la Sécurité sociale
FDA	<i>Food and Drug Administration</i>
FinCCHTA	<i>Finnish Coordinating Center for Health Technology Assessment</i>
G-BA	<i>Federal Joint Committee, Allemagne</i>
GoeG	<i>Gesundheit Österreich GmbH, Autriche</i>
HAS	Haute Autorité de santé
IA	Intelligence artificielle
IQWiG	<i>Institute for Quality and Efficiency in Health Care, Allemagne</i>
IMDRF	<i>International Medical Device Regulators Forum</i>
INAHTA	<i>International Network of Agencies for Health Technology Assessment</i>
INESSS	Institut national d'excellence en santé et en services sociaux
IQWiG	<i>Institute for Quality and Efficiency in Health Care, Allemagne</i>
LAD	Logiciel d'aide à la dispensation
LAP	Logiciel d'aide à la prescription
LPPR	Liste des prestations et produits remboursables
NHS	<i>National Health Service</i>
NICE	<i>National Institute for Health and Care Excellence</i>
PRO	<i>Patient Reported Outcomes</i>
SaMD	<i>Software as a Medical Device</i>
TDRD	Télédepistage de la rétinopathie diabétique
ZIN	<i>National Health Care Institute (Zorginstituut Nederland)</i>
ZonMW	<i>Dutch Organization for Health Research and Development</i>

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

