



Votre IA est-elle prête pour l'évaluation clinique en vue du marquage CE ?

Medi'Nov 1er Juillet 2026

© Rath Consultant [2026]. Tous droits réservés. Reproduction interdite.

Qu'est-ce que l'évaluation clinique ?

Un **processus systématique et planifié** visant à produire, collecter, analyser et évaluer en continu les données cliniques relatives à un dispositif, afin de vérifier sa **sécurité, ses performances et ses bénéfices cliniques** conformément à la destination prévue par le fabricant.

« L'évaluation clinique est approfondie et objective, et prend en compte les données favorables et défavorables. Son ampleur est proportionnée à la nature, à la classification, à la destination, aux risques et aux allégations du dispositif. »

Obligation réglementaire

Exigence incontournable dans le cadre du Règlement européen sur les dispositifs médicaux (MDR 2017/745).

Documentation technique

Partie intégrante du dossier technique du fabricant, soumis à l'organisme notifié.

Preuve de conformité

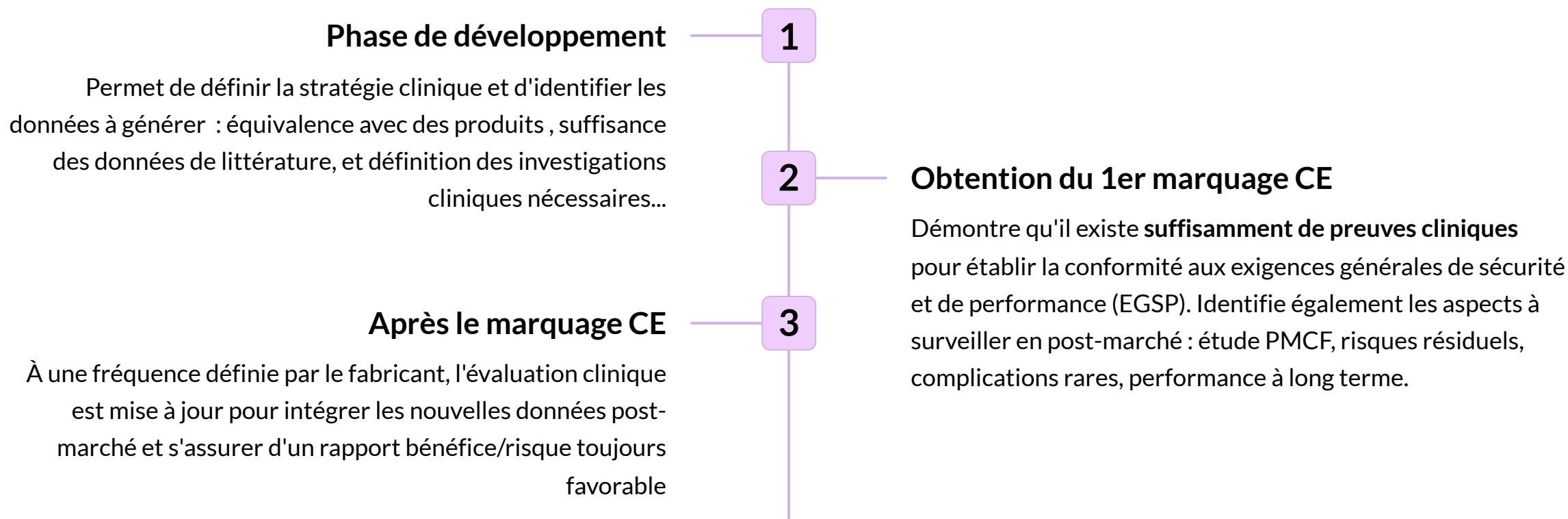
Démontre la sécurité et les performances cliniques tout au long du cycle de vie du dispositif.

Processus continu

Ne s'arrête pas au marquage CE , l'évaluation se poursuit en phase de surveillance post-marché.

Quand réaliser une évaluation clinique ?

L'évaluation clinique intervient à **trois moments clés** du cycle de vie du dispositif médical. Elle n'est pas un exercice ponctuel, mais une démarche structurée et itérative.



DM IA : Un produit, deux réglementations

MDR

Le côté médical

- Destination prévue
- Sécurité et performances
- Preuves cliniques
- Gestion des risques
- Surveillance après commercialisation

AI Act

Le côté algorithmique

- Qualité et représentativité des données
- Biais et équité
- Transparence et explicabilité
- Robustesse et cybersécurité
- Surveillance humaine et monitoring

📄 Le guide MDCG 2025-6 le confirme : si votre logiciel est un Dispositif Médical à Intelligence Artificielle (MDAI), vous êtes soumis aux deux cadres réglementaires.

Une seule évaluation suffit à couvrir les exigences du MDR et de l'IA act

la correspondance entre les exigences MDR et AIA doit être visible

CEP (Clinical Evaluation Plan)

Définit la stratégie et la méthode pour évaluer les données cliniques d'un dispositif médical.

CER (Clinical evaluation report)

Présente les résultats de cette évaluation notamment les données cliniques et démontre sa sécurité et ses performances.

Données cliniques selon le MDR

Investigations/études cliniques du dispositif concerné

Investigation(s) clinique(s) du dispositif concerné.

Études sur des dispositifs équivalents

Investigation(s) clinique(s) ou autres études rapportées dans la littérature scientifique, portant sur un dispositif dont l'équivalence avec le dispositif en question peut être démontrée.

Littérature scientifique à comité de lecture

Rapports publiés dans des revues scientifiques à comité de lecture sur d'autres expériences cliniques du dispositif en question ou d'un dispositif dont l'équivalence avec le dispositif en question peut être démontrée.

Surveillance après commercialisation

Informations cliniquement pertinentes issues de la surveillance après commercialisation, notamment le suivi clinique après commercialisation.

Données définies par l'AI Act

L'AIA introduit quatre définitions relatives aux données, à savoir :

Données d'entraînement

Article 3 (29) AIA : Données utilisées pour entraîner un système d'IA en ajustant ses paramètres apprenables.

Données de validation

Article 3 (30) AIA : Données utilisées pour évaluer le système d'IA entraîné et pour ajuster ses paramètres non apprenables ainsi que son processus d'apprentissage, afin notamment d'éviter le sous-apprentissage ou le surapprentissage.

Jeu de données de validation

Article 3 (31) AIA : Un jeu de données distinct ou une partie du jeu de données d'entraînement, sous forme de division fixe ou variable.

Données de test

Article 3 (32) AIA : Données utilisées pour fournir une évaluation indépendante du système d'IA afin de confirmer les performances attendues de ce système avant sa mise sur le marché ou sa mise en service.

Articulation entre le RDM et le RIA :

Entraînement IA

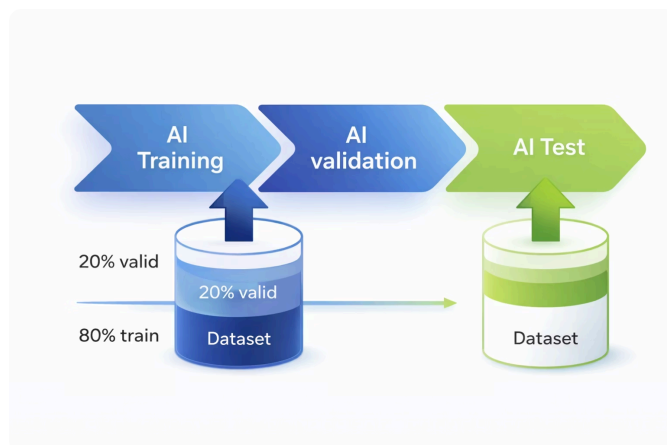
C'est une étape de la conception de l'IA et donc de la conception du dispositif. Le jeu de données d'entraînement est utilisé à cette étape.

Validation IA

C'est une étape de la conception de l'IA. Le jeu de données de validation IA est utilisé à cette étape. Ce jeu peut être extrait de la même source de données que le jeu d'entraînement (en général, ratio 80 - 20).

Tests

Les tests de l'IA sont effectués afin de démontrer que la performance du modèle d'IA est celle attendue. C'est donc une étape de tests de performance de l'IA, qui peut concourir à la validation du DM. Le jeu de données de test IA (ou test de performance, en langage du dispositif médical) est utilisé à cette étape. Ce jeu de données doit être strictement indépendant du jeu de données entraînement et de validation IA. Il ne doit pas provenir de la même source de données.



Mettre en place une stratégie dès le départ

La stratégie d'évaluation clinique ne doit pas être une réflexion tardive. Elle doit être **intégrée dès les premières phases de conception** du dispositif médical pour garantir une trajectoire réglementaire cohérente et maîtrisée.



Planification précoce

Définir dès la conception les allégations cliniques, les populations cibles et les endpoints pertinents .



Analyse des données disponibles

Évaluer l'état de l'art scientifique, les équivalences possibles et les lacunes à combler par des études/ investigations cliniques spécifiques.



Alignement réglementaire

Anticiper les exigences de l'organisme notifié et préparer la documentation technique en cohérence avec les EGSP applicables.

❏ Une stratégie clinique bien construite dès le départ évite des retards coûteux lors de la soumission au marquage CE.

Etat de l'Art (SOTA)

État de l'art : Stade de développement des capacités techniques actuelles et/ou de la pratique clinique acceptée concernant les produits, les procédés et la prise en charge des patients, fondé sur les résultats consolidés pertinents de la science, de la technologie et de l'expérience.

Ce qui est connu sur la pathologie ciblée par votre dispositif

Quelles solutions et méthodes sont actuellement utilisées

Quels niveaux de performance, de sécurité sont déjà atteints et quelles sont les limites et besoins non couverts

❏ **Note :** L'ÉTAT DE L'ART représente ce qui est actuellement et généralement accepté comme bonne pratique en technologie et en médecine. L'état de l'art n'implique pas nécessairement la solution la plus avancée sur le plan technologique.

Les preuves cliniques : quantité et qualité

Les données cliniques doivent être **suffisantes en volume et en qualité** pour permettre d'évaluer en connaissance de cause si le dispositif est sûr et offre les bénéfices cliniques attendus. **Chaque allégation doit être soutenue par des données.**

Quantité des données

- Données suffisantes pour l'usage prévu, les indications et les groupes cibles
- Risques cliniques et performance analytique évalués
- Caractéristiques du DM IA prises en compte dans la génération des données
- Degré d'innovation et historique sur le marché considérés
- Volume proportionné à la complexité et aux risques du dispositif

Qualité des données

- Type et conception d'étude adaptés aux objectifs de recherche
- Jeu de données approprié, à jour et représentatif de l'état de l'art
- Approche statistique valide pour conclure de manière robuste
- Considérations éthiques, juridiques et réglementaires respectées
- Absence de conflits d'intérêts vérifiée et documentée

❏ Pour les logiciels de dispositifs médicaux intégrant de l'IA, une attention particulière doit être portée à la représentativité des données d'entraînement et à la validité clinique des algorithmes dans les populations cibles définies.

Stratégies d'évaluation clinique

Equivalence

Sans données du dispositif en évaluation



Investigation clinique ou autres études

Basée sur les données du dispositif en évaluation (rétrospective, base de données...)



Exemption de données clinique ou Stratégie du 61(10)

Basée sur les données du dispositif en évaluation



Littérature scientifique

Basée sur les données du dispositif en évaluation





Stratégie d'évaluation clinique sans données du dispositif en évaluation : Equivalence

« L'évaluation clinique peut reposer sur les données cliniques relatives à un dispositif pour lequel l'équivalence avec le dispositif en question peut être démontrée »

- Pas de différence cliniquement significative en ce qui concerne la sécurité et les performances cliniques du dispositif.
- La démonstration repose sur une justification scientifique appropriée.
- Il est démontré clairement que les fabricants ont un accès suffisant aux données relatives aux dispositifs qu'ils considèrent comme équivalents

1

Sur le plan technique

Le dispositif est de conception similaire; est utilisé dans des conditions similaires; possède des spécifications et des propriétés similaires, notamment des propriétés physicochimiques telles que l'intensité énergétique, la résistance à la traction, la viscosité, les propriétés de surface, la longueur d'onde et les algorithmes; utilise des méthodes d'installation similaires, le cas échéant; est sujet aux mêmes principes opératoires et exigences en matière de performances critiques,

2

Sur le plan biologique

Le dispositif emploie les mêmes matériaux ou substances en contact avec les mêmes tissus humains ou fluides corporels pour un type et une durée de contact similaires et des caractéristiques de libération des substances similaires, y compris les produits de dégradation et substances relargables,

3

Sur le plan clinique

Le dispositif est utilisé pour un même état clinique, ou une même destination, notamment sévérité et stade similaires de la maladie, en un même endroit du corps, auprès d'une population similaire, notamment pour ce qui est de l'âge, l'anatomie et la physiologie; a le même type d'utilisateur; a des performances critiques pertinentes similaires au regard de l'effet clinique escompté pour une destination donnée.

Investigation clinique

ARTICLE 2 (45) MDR

Investigation clinique : Toute investigation systématique impliquant un ou plusieurs sujets humains, entreprise pour évaluer la sécurité ou les performances d'un dispositif.

Stratégie la plus fiable actuellement → Permet de couvrir l'ensemble des revendications à court/moyen termes.

Incontournable pour les dispositifs très **innovants**

© Rath Consultant [2026]. Tous droits réservés. Reproduction interdite.



Stratégie d'évaluation clinique basée sur les données du dispositif en évaluation : autres études

ÉTUDES CLINIQUES HORS MDR : ATTENTION RÉGLEMENTATION NATIONALE !

Etude
rétrospective

Étude sur base de
données



Enquêtes

Registres

Étude en vie réelle
(observationnelle
/
interventionnelle)



Stratégie d'évaluation clinique basée sur les données du dispositif en évaluation : littérature scientifique

Littérature scientifique sur le dispositif en évaluation

- Preuve que le dispositif est utilisé dans la publication (nom cité)
- Publication complète, pas d'abstract, pas de poster.

Nécessité de documents publiés, revus par les pairs

Moins de liberté en termes de revendications mais stratégie fiable.

© Rath Consultant [2026]. Tous droits réservés. Reproduction interdite.

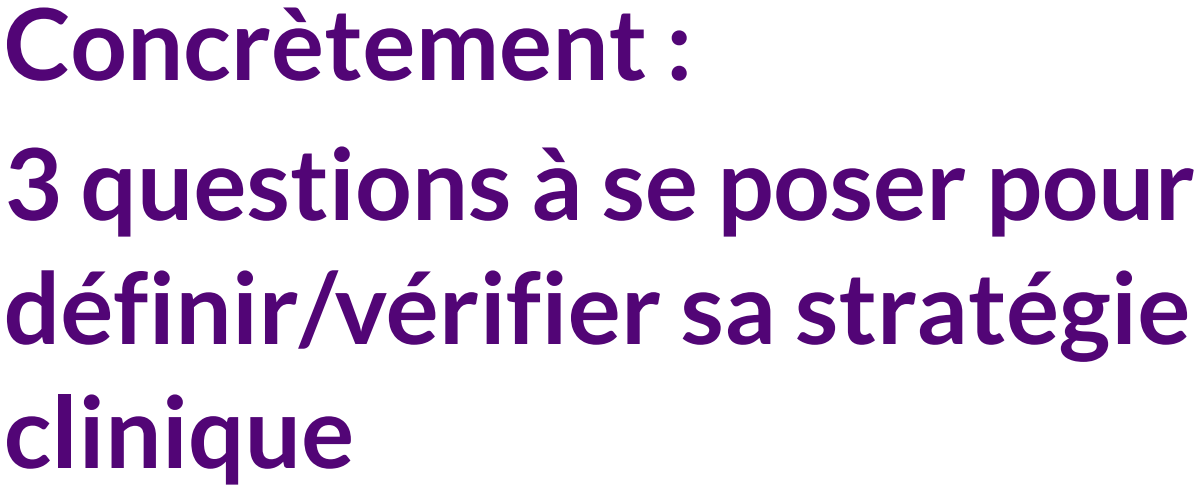




Exemption de données cliniques ou Stratégie du 61(10)

- **Justification :**
- La gestion des risques (risques résiduels faibles; démontrer que les méthodes de test non cliniques sont appropriées au vu des risques identifiés)
- Peu/pas d'interaction entre le DM et le corps humain
- Revendications du fabricant (performances techniques, bénéfice clinique indirect)
- Pertinence/Faisabilité d'une investigation clinique

= SOTA + résultats de la gestion des risques + usability + essais non cliniques (bancs d'essai, évaluation préclinique)



© Rath Consultant [2026]. Tous droits réservés. Reproduction interdite.

1) Y-a t-il suffisamment de données dans la littérature (SOTA) pour documenter l'association clinique valide/validité scientifique ?

Signifie l'association d'une sortie logiciel à une condition clinique ou à un état physiologique.

ASSOCIATION CLINIQUE VALIDE / VALIDITÉ SCIENTIFIQUE : la mesure dans laquelle la sortie du DM IA (p. ex. concept, conclusion, calculs), fondée sur les entrées et les algorithmes sélectionnés, est associée à l'état physiologique cible ou à la condition clinique ciblée. Cette association doit être solidement fondée ou cliniquement acceptée (p. ex. existence d'un cadre scientifique ou niveau de preuve suffisant).

L'ASSOCIATION CLINIQUE VALIDE / VALIDITÉ SCIENTIFIQUE d'un DM IA doit démontrer qu'elle correspond à la situation clinique, à la condition, à l'indication ou au paramètre défini dans la destination prévue du DM IA..

Les preuves à l'appui de l'ASSOCIATION CLINIQUE VALIDE / VALIDITÉ SCIENTIFIQUE peuvent être générées, par exemple, au moyen d'une recherche bibliographique, de lignes directrices professionnelles, d'une preuve de concept, ou des propres investigations cliniques/études de performance clinique du fabricant.

Données de démonstration d'Association Clinique valide

- ❏ Le fabricant doit vérifier l'association entre la sortie du DM IA et la condition physiologique/clinique cible, la situation clinique ou le paramètre clinique, tels que définis dans l'usage prévu du DM IA.

Le CEP doit prévoir comment sera démontrée l'association clinique

Exemples de données existantes (sans ordre particulier) :

- Normes techniques
- Recommandations des sociétés médicales professionnelles
- INVESTIGATIONS CLINIQUES / ÉTUDES DE PERFORMANCE CLINIQUE
- Données cliniques publiées (p. ex. Summary of Safety and Clinical Performance (SSCP) / Summary of Safety and Performance (SSP), registres et bases de données des autorités)

—> Faire une revue de la littérature (SOTA)

Exemples de génération de nouvelles preuves (sans ordre particulier) :

- Analyse de données secondaires (analyse de données du monde réel)
- Réaliser une INVESTIGATION CLINIQUE

Exemples DM IA



Exemple 1 — Sommeil & PPC

Analyse des données d'**accéléromètre** et de **microphone** (téléphone / dispositif portable) pour :

- Déterminer la qualité du sommeil et son impact sur le bien-être général
- Surveiller le sommeil chez les patients atteints d'apnée du sommeil
- Estimer le taux de réussite attendu de la **thérapie PPC (pression positive continue)**



Exemple 2 — Imagerie

Détection automatique d'organes et structures anatomiques sur scanner pour:

- Détecter les **anévrismes de l'aorte abdominale** sur scanners abdominaux
- Détecter les **fractures de compression vertébrales**



Exemple 3 — TCC Digitale

Accompagnement des adultes souffrant de **dépression légère à modérée**, en complément d'une prise en charge médicale :

- Évaluation de la sévérité de la dépression via **questionnaires cliniques validés**
- Adaptation des modules de TCC au profil évolutif du patient
- **Réduction des symptômes dépressifs**

Exemples

Exemple 1 – Sommeil et PCC

La qualité du sommeil (durée, efficacité, fragmentation) et son impact sur le bien-être sont des notions **bien établies**. En revanche, **le lien entre surveillance de la qualité du sommeil et prédiction du succès de la thérapie PPC n'est pas solidement établi**. L'association entre données d'accéléromètre/microphone et paramètres polysomnographiques de référence présente des lacunes. La **génération de données cliniques complémentaires est requise** pour combler ces manques.

Exemple 2 – Imagerie

La forme et la taille normales des structures anatomiques (aorte, vertèbres) sont **bien documentées dans la littérature**. Les techniques de segmentation sur images en coupe transversale CT sont corrélées aux dimensions réelles avec un haut niveau de preuve. **Aucune lacune majeure n'est identifiée** : l'association clinique valide est considérée établie, permettant de s'appuyer sur les données existantes.

Exemple 3 – TCC digitale

La corrélation entre scores de questionnaires cliniques validés (ex. PHQ-9) et sévérité clinique de la dépression est **bien établie**. Le point à démontrer spécifiquement est **l'adéquation entre le score calculé et le module de TCC sélectionné par l'algorithme**. Des données manquantes devront être générées pour valider cette correspondance.

2) Le fabricant a-t-il suffisamment de données pour démontrer la Performance Technique

Capacité d'un MDSW à générer avec précision et fiabilité la sortie technique prévue à partir des données d'entrée.

La validation de la **Performance technique** est la démonstration de la capacité du MDSW à générer avec précision, fiabilité et exactitude la sortie prévue, à partir des données d'entrée.

Les éléments probants soutenant la **Performance technique** peuvent être générés par des activités de vérification et de validation, par exemple des tests au niveau unitaire, d'intégration et système, ou par la génération de nouvelles preuves à l'aide de bases de données, de registres, de bases de données de référence ou de l'utilisation de données patients collectées précédemment.

Exemples

La **performance technique** vise à confirmer que le logiciel calcule de manière **fiable, reproductible et précise** les résultats qu'il est censé produire, Elle repose sur des tests de vérification (conformité aux spécifications) et/ou de validation (adéquation à l'usage prévu).

Exemple 1 — Score de qualité du sommeil

- Confirmer que l'application calcule de façon fiable et reproductible un score de qualité du sommeil à partir des capteurs
- Valider la compatibilité entre le DM IA et les appareils équipés des capteurs (accéléromètre, microphone)
- S'assurer que les données capteurs peuvent être utilisées conformément à l'usage prévu

Exemple 2 — Segmentation d'images CT

Confirmer, par des tests de vérification et de validation, les performances techniques de base de l'outil d'imagerie :

- **Affichage et modification** des images CT (réglage de fenêtre, contraste)
- **Précision des mesures** anatomiques (taille, forme des structures segmentées)
- **Sensibilité et spécificité** de la détection des anévrismes et fractures de compression
- **Fiabilité et reproductibilité** des résultats de segmentation automatique, conformément aux performances attendues déclarées par le fabricant

Exemple 3 — Score TCC & sélection de module

- Tests de reproductibilité et exactitude du calcul du score depuis les réponses au questionnaire
- Tests de cohérence de l'algorithme de sélection : un même profil de réponses doit systématiquement orienter vers le même module de TCC

3) Le fabricant a-t-il suffisamment de données pour démontrer la Performance clinique

Performance clinique : La **capacité** d'un dispositif, du fait de tout effet médical direct ou indirect résultant de ses caractéristiques techniques ou de fonctionnement, y compris en matière de diagnostic, **à atteindre sa destination comme indiqué par le fabricant**, et **à produire de ce fait un bénéfice clinique** pour les patients, lorsqu'il est utilisé comme prévu par le fabricant.

Effets médicaux directs / indirects

Découlant des caractéristiques techniques, fonctionnelles et diagnostiques.

Destination prévue

Atteindre la destination prévue revendiquée par le fabricant.

Bénéfice clinique

Conduire à un bénéfice clinique pour les patients lors d'une utilisation conforme.

Bénéfice clinique

BÉNÉFICE CLINIQUE : L'incidence positive d'un dispositif sur la santé d'une personne physique, se traduisant par un (des) résultat(s) clinique(s) significatif(s), mesurable(s) et pertinent(s) pour le patient, y compris le(s) résultat(s) en matière de diagnostic, ou une incidence positive sur la prise en charge de la santé du patient ou sur la santé publique.

Bénéfice indirect

Le dispositif lui-même n'a pas d'impact positif direct sur les patients, mais il permet à une procédure d'être entreprise, à une thérapie d'être délivrée ou à un autre dispositif d'atteindre son objectif.

- Logiciel d'aide à la prescription
- Un logiciel d'analyse d'imagerie qui revendique la précision diagnostique (attention un logiciel d'analyse d'imagerie peut aussi avoir un bénéfice clinique direct selon ses revendications)

Validation des performances cliniques

démontrer que le DM IA a été testé :

- pour les utilisations prévues, les populations cibles,
- les conditions d'utilisation,
- les environnements de fonctionnement et d'utilisation, ainsi qu'avec
- tous les groupes d'utilisateurs prévus.

La validation des performances cliniques comprend

- l'évaluation de la sécurité clinique,
- de l'efficacité et des performances,
- peut appuyer la démonstration du bénéfice clinique.

La validation des performances cliniques devrait être envisagée à chaque changement du logiciel vers une nouvelle version. Si aucune validation n'est effectuée, une justification devrait être indiquée dans la documentation technique.

Exemples

1

Exemple 1 – Sommeil et PCC

En complément de l'évaluation de l'**aptitude à l'utilisation (usability)**, le fabricant réalise une **étude rétrospective** sur données préexistantes pour confirmer que le succès de la thérapie PPC peut être prédit à partir de la surveillance de la qualité du sommeil.

Si des données rétrospectives ne sont pas disponibles, une investigation clinique serait requise.

2

Exemple 2 – Imagerie

L'évaluation de l'**aptitude à l'utilisation** combinée à l'association clinique valide établie et aux résultats de performance technique pourrait être jugée **suffisante pour répondre aux GSPR**.

Une analyse rétrospective peut compléter le dossier si des données sont disponibles.

Une **investigation clinique prospective** peut être requise si la variabilité des paramètres d'entrée n'est pas couverte par les données existantes.

3

Exemple 3 – TCC digitale

Une **investigation clinique** est conduite pour démontrer l'évolution du score de dépression chez les patients utilisant les **modules TCC adaptés par l'algorithme**, comparés à des modules standards ou à une prise en charge habituelle seule.

La sélection de la stratégie de performance clinique doit être **justifiée et proportionnée** au niveau de risque du dispositif, aux lacunes identifiées lors des étapes précédentes, et à la disponibilité des données existantes. Toute lacune non couverte par la littérature ou les données rétrospectives/ base de données impose la génération de preuves cliniques prospectives.

Conclusion

L'évaluation clinique d'un DM IA n'est pas une case à cocher en fin de développement. C'est une démarche stratégique qui se construit **dès la conception**, se nourrit de l'état de l'art, et s'adapte au niveau de risque réel du dispositif et aux données déjà disponibles/exploitable.

MDR & AI Act : une superposition, pas une opposition

Une seule évaluation peut couvrir les deux cadres réglementaires, à condition que la correspondance entre les exigences soit **explicite et documentée**. L'anticipation de cette articulation est un levier décisif pour éviter retards et refus au marquage CE.

Trois questions structurantes

- L'**association clinique** est-elle déjà établie dans la littérature ?
- La **performance technique** peut-elle être démontrée par vérification et validation ?
- La **performance clinique** repose-t-elle sur des données suffisantes (équivalence, littérature...), ou impose-t-elle une génération de preuves nouvelles (étude rétrospective, base de données, investigatin clinique...)?

📌 **Principe fondamental** : Chaque revendication doit pouvoir s'appuyer sur une source de données identifiée. Pas d'exception, pas d'approximation.



Merci de votre attention

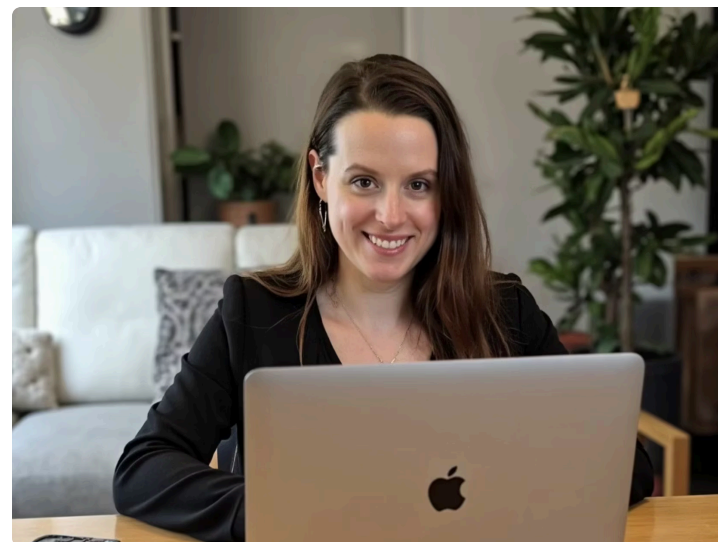
Des questions ?

© Rath Consultant [2026]. Tous droits réservés. Reproduction interdite.

Audeline Rath-Lavialle



© Rath Consultant [2026]. Tous droits réservés. Reproduction interdite.



QARA experte en Clinique et Remboursement

audeline@rathconsultant.com