

Intelligence artificielle en addictologie : intérêts et limites

Alain Dervaux

DANS **L'INFORMATION PSYCHIATRIQUE** 2025/9 Vol. 101, PAGES 707 À 713
ÉDITIONS JLE ÉDITIONS

ISSN 0020-0204

DOI 10.1684/ipe.2025.2951

Article disponible en ligne à l'adresse

<https://stm.cairn.info/revue-linformation-psychiatrique-2025-9-page-707?lang=fr>



Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...
Scannez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



Distribution électronique Cairn.info pour JLE Éditions.

Vous avez l'autorisation de reproduire cet article dans les limites des conditions d'utilisation de Cairn.info ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Détails et conditions sur cairn.info/copyright.

Sauf dispositions légales contraires, les usages numériques à des fins pédagogiques des présentes ressources sont soumises à l'autorisation de l'Éditeur ou, le cas échéant, de l'organisme de gestion collective habilité à cet effet. Il en est ainsi notamment en France avec le CFC qui est l'organisme agréé en la matière.

Intelligence artificielle en addictologie : intérêts et limites

Alain Dervaux

Professeur
Université Paris-Saclay
(<http://www.universite-paris-saclay.fr/>)
Filière universitaire d'addictologie,
EPS Barthélémy Durand, Étampes
PSYCHOMAAD Research Unit,
Hôpital Paul Brousse, Villejuif, France
Institut de psychiatrie (CNRS GDR 3557),
Paris, France

En addictologie, l'intelligence artificielle (IA) n'est pas aussi développée que dans d'autres disciplines médicales. Cependant, plusieurs études ont montré que l'IA pourrait avoir un intérêt dans le repérage, l'évaluation, le diagnostic et la prise en charge des patients souffrant d'addictions, notamment en raison des très nombreux paramètres biologiques, psychologiques et sociaux qui les déterminent. Les agents conversationnels (chatbots) ont un intérêt potentiel dans la prise en charge des addictions.

Toutefois, leur utilisation comporte des limites et des risques bien identifiés. Les erreurs ou « hallucinations » de l'IA, générant des informations fausses ou obsolètes, sont relativement fréquentes. La question de la confidentialité des données est également majeure, nécessitant anonymisation et consentement. En outre, l'IA présente divers biais qui peuvent influencer les résultats, entre autres, biais de représentativité, d'acquiescence, de formulation et de confirmation. Pour bien utiliser l'IA, il faut en comprendre ses forces autant que ses limites.

intelligence artificielle, addiction, addictologie, application numérique, robot

Artificial intelligence in addictology: Potential uses and limitations.

Artificial intelligence (AI) remains less developed in addiction medicine compared to other medical fields. However, several studies have suggested its potential usefulness in screening, assessment, diagnosis, and management of patients with substance use disorders, given the numerous biological, psychological, and social factors involved. Conversational agents, also known as chatbots, also show promise as tools for supporting patients. However, the use of AI entails significant risks, including algorithmic errors, "hallucinations," and the generation of outdated information. Data confidentiality also represents a major concern, requiring both anonymization and informed consent. Furthermore, AI systems are susceptible to multiple biases, such as representativeness, acquiescence, framing, or confirmation biases, that may influence results. AI should therefore be considered a complementary tool to clinical expertise, one that requires a clear understanding of both its strengths and its limitations.

artificial intelligence, addiction, substance use disorders, digital application, robot

Inteligencia artificial en el tratamiento de adicciones: ventajas y limitaciones.

En el tratamiento de adicciones, la inteligencia artificial (IA) no está tan desarrollada como en otras disciplinas médicas. Sin embargo, varios estudios han demostrado que la IA podría ser de algún interés para la detección, evaluación, diagnóstico y tratamiento de pacientes con adicciones, especialmente debido a los numerosos parámetros biológicos, psicológicos y sociales que las determinan. Los agentes conversacionales (chatbots) tienen un interés potencial en el tratamiento de las adicciones.

Sin embargo, su uso conlleva límites y riesgos bien identificados. Los errores o «alucinaciones» de la IA, que generan información falsa u obsoleta, son relativamente frecuentes. La cuestión de la confidencialidad de los datos también es esencial, ya que requiere el anonimato y el consentimiento. Además, la IA presenta diversos sesgos que pueden influir en los resultados, entre otros, el sesgo de representatividad, de aquiescencia, de formulación y de confirmación. Para utilizar correctamente la IA, es necesario comprender tanto sus puntos fuertes como sus limitaciones.

inteligencia artificial, adicción, tratamiento de las adicciones, aplicación digital, robot

Introduction

L'intelligence artificielle (IA) est aujourd'hui bien installée dans notre quotidien. Depuis leur diffusion fin 2022, les agents conversationnels par IA générative, tels que ChatGPT (*Generative Pre-trained Transformer*), développé par OpenAI, sont massivement utilisés par les patients. Ceux-ci y cherchent des informations sur

Correspondance : A. Dervaux
[alain.dervaux@universite-paris-saclay.fr](mailto:<alain.dervaux@universite-paris-saclay.fr>)

les pathologies et les traitements, dont une majorité qui les utilisent avant d'aller consulter un médecin [1]. Une enquête a retrouvé que 44 % des psychiatres ont déjà utilisé ChatGPT-3.5 et 33 % ChatGPT-4.0 pour chercher des informations cliniques et 70 % pour la recherche documentaire [1].

Intérêts de l'intelligence artificielle en addictologie

Dépistage, diagnostics et évolution

Déjà utilisée dans l'aide à la décision dans le domaine médical, notamment en oncologie, dermatologie ou en imagerie médicale, l'intelligence artificielle (IA) peut réaliser des diagnostics et prédire l'évolution avec plus de précision que les cliniciens, mais pas toujours [2]. Le plus grand nombre d'études a été réalisé en radiologie et en biologie, deux domaines où les données sont nombreuses et en pleine expansion. En psychiatrie, certaines études ont retrouvé que l'IA pouvait prédire la survenue de dépression et de comportements suicidaires avec plus de précision que les cliniciens. En addictologie, l'IA n'est pas aussi développée mais pourrait avoir un intérêt pour le repérage, l'évaluation, le diagnostic et la prise en charge.

Un domaine particulièrement pertinent en addictologie est le repérage. L'IA pourrait faciliter l'identification précoce des consommations à risque et des addictions en population générale [3] ou dans des populations de patients psychiatriques, notamment les troubles de l'usage d'alcool [4] ou d'opioïdes [5]. En addictologie, à ce jour, ce sont surtout des études pilotes ou de faisabilité qui ont été réalisées [6], par exemple le repérage des surdoses d'opioïdes en fonction des localisations géographiques, les tendances de consommations de

substances sur les réseaux sociaux, le phénotypage digital (*digital phenotyping*) [7]. Une étude dans une population de sujets recevant des opioïdes à visée antalgique, utilisant l'IA pour analyser des paramètres physiologiques et des tâches cognitives, a retrouvé de bonnes capacités pour prédire les sujets à risque de mésusage de ces médicaments [8]. Une autre a identifié parmi plus de 35 000 dossiers médicaux électroniques de patients hospitalisés, ceux qui présentaient des troubles de l'usage d'opioïdes susceptibles d'être adressés à des équipes d'addictologie [9]. Une étude a identifié parmi les patients présentant des troubles de l'usage d'opioïdes, ceux qui présentaient le plus de risques de surdoses [10].

L'IA pourrait aider les médecins à évaluer les patients présentant une (des) addiction(s) et établir des diagnostics addictologiques, psychiatriques et somatiques qui peuvent être nombreux et très souvent associés. L'évaluation de ces patients nécessite le recueil et l'analyse d'un très grand nombre de données, résumées dans le *tableau 1*, difficilement appréhendées par un seul clinicien addictologue dans un temps limité [2, 3, 11]. L'IA peut améliorer l'évaluation de ces patients, notamment en termes de sévérité des troubles [12], de sévérité des sevrages notamment à l'alcool [13], de repérage de certains troubles comorbides très fréquents tels que la dépression, le trouble déficit de l'attention avec ou sans hyperactivité (TDAH) ou identification des patients à risque de suicide [14, 15] ou de violence [16]. Par exemple, la société USC Institute of Creative Technologies a mis au point le prototype *Multisense* d'analyse des expressions faciales et corporelles pour dépister les patients déprimés : l'application SimSensei simule un psychologue virtuel qui se présente visuellement sous forme d'avatar. Le langage et les mouvements des yeux, la posture, la prosodie, etc. du patient sont analysés automatiquement en temps réel et l'avatar

Tableau 1. Caractéristiques bio-psycho-sociales des patients présentant des troubles de l'usage de substances.

Évaluation addictologique	Troubles cognitifs	Comorbidités psychiatriques	Complications somatiques	Contexte social
<ul style="list-style-type: none"> - Modes de consommation - Niveaux de sévérité - Insight - Addictions associées : tabac, cannabis, addictions comportementales ... 	<ul style="list-style-type: none"> - Attention - Mémoire de travail - Mémoire épisodique - Fonctions exécutives - Biais cognitifs - Niveaux d'impulsivité 	<ul style="list-style-type: none"> - Troubles anxieux - Troubles dépressifs - Troubles bipolaires - Troubles psychotiques - Troubles de personnalité, notamment antisociale et borderline - Événements de vie - Maltraitances physiques, émotionnelles, abus sexuels, négligences - Antécédents familiaux - Traitements psychotropes 	<ul style="list-style-type: none"> - Pratiquement toutes les spécialités médicales (par ex : 60 à 200 maladies induites par l'alcool) - Anomalies biologiques 	<ul style="list-style-type: none"> - Situations sociales complexes - Statut addictologique des conjoints - Activités professionnelles - Médecine au travail - Précarité - Problèmes judiciaires

s'adapte en modulant son discours en tenant compte des réponses aux questions posées. Le logiciel couplé à *Multisense* permettrait de détecter les symptômes de dépression avec une précision de 90 %. Ces études sont néanmoins préliminaires, sur des nombres limités de patients, et ne sont pas mises en œuvre pour l'instant en pratique clinique courante [13].

Concernant l'évolution, une étude a retrouvé que l'IA prédisait l'évolution en termes d'abstinence avec une meilleure spécificité (70 %) que les psychologues cliniciens après psychothérapie cognitivo-comportementale (TCC) [2]. Cependant la sensibilité restait faible [17].

Intérêt de l'IA dans la prise en charge des patients

L'analyse de grandes quantités de données par *machine learning* (apprentissage autonome par l'IA) concernant des rapports d'effets indésirables et de publications scientifiques permet d'améliorer la sécurité des prescriptions médicamenteuses. Concernant la recherche, l'IA peut aider à la gestion de projets, à la planification, à découvrir de nouvelles molécules, à accélérer les phases de conception des essais cliniques, à optimiser l'inclusion de patients et donc la mise sur le marché de nouveaux traitements.

Certaines études ont retrouvé que l'IA pouvait aider aux décisions thérapeutiques par traitements agonistes opioïdes chez des patients présentant des troubles de l'usage d'opioïdes [7]. Un intérêt émergent en addictologie est le suivi du craving, des consommations et comportements associés par des dispositifs digitaux et des systèmes de monitoring utilisés par les patients pour les aider à gérer leurs consommations (*ecological momentary assessment*) [3], notamment dans les addictions aux opioïdes [10, 18]. D'autres études pilotes ont suggéré que certaines techniques de TCC par IA pourraient être efficaces, bien que de façon moins marquée qu'avec les thérapies en face-à-face.

Intérêt des agents conversationnels par IA (chatbots) dans la prise en charge des addictions

Les *large language model* (LLM) peuvent reconnaître, résumer, traduire, prédire et générer des textes ou des images provenant de très grandes bases de données [19, 20]. Ils sont la base des agents conversationnels ou IA générative, dont ChatGPT est le pionnier et le plus connu. Ces agents permettent de répondre rapidement aux questions des utilisateurs. Ils ne se substituent pas aux approches cliniques mais peuvent les faciliter, notamment pour le dépistage, y compris de maladies rares, la surveillance et le repérage des décompensations [6, 21].

Les agents conversationnels ont beaucoup d'intérêt pour la psychoéducation [22]. Plusieurs études ont retrouvé que les réponses de ChatGPT aux requêtes

d'informations par les usagers dans plusieurs spécialités, étaient comparables à celles des médecins, et souvent plus empathiques. L'évaluation de l'IA générative dans les domaines des addictions a été évaluée par des cliniciens spécialistes et a été jugée comme de bonne qualité [23]. ChatGPT peut être un outil complémentaire dans la communication patient-médecin, en particulier pour fournir des explications claires et accessibles sur des sujets médicaux complexes.

Une méta-analyse de 18 études randomisées contrôlées, ayant inclus 3477 sujets, a retrouvé que les agents conversationnels étaient efficaces pour améliorer les troubles anxieux et dépressifs, fréquemment associés aux addictions [24]. Par exemple, l'utilisation de l'agent conversationnel Therabot, élaboré par le Dartmouth's Center for Technology and Behavioral Health, a permis de diminuer les symptômes dans une population de 106 patients souffrant de dépression caractérisée ou de troubles anxieux généralisés [25]. Une étude de faisabilité a retrouvé qu'un agent conversationnel intégrant les entretiens motivationnels ciblant la consommation d'alcool pouvait être utilisé sans risque chez de jeunes adultes [26]. Des agents conversationnels en réalité virtuelle ont été développés dans un but psychothérapeutique (*CBT integrated with AI*) : l'IA favorise en effet la verbalisation, la métapensée et des relations sans jugement. Les agents conversationnels peuvent être utilisés en plusieurs langues [25]. Ils peuvent apporter une écoute bienveillante et proposer des conseils, notamment pour les patients qui ont des difficultés d'accès aux soins ou qui ont des difficultés à communiquer avec les soignants [22]. Il ressort de plusieurs enquêtes que les patients ne se sentent pas jugés par l'IA et rapportent apprécier de pouvoir parler librement [25]. Par exemple, lorsque l'IA commence par engager la conversation avec un patient sur un mode ouvert : « *Qu'est qui t'intéresse le plus en ce moment ?* », les patients abordent leurs problématiques très facilement, y compris sur des sujets que les thérapeutes n'auraient jamais imaginés. Beaucoup de patients gardent conscience qu'ils s'adressent à un robot, ce qui atténue leur peur d'être jugés.

L'utilisation des agents conversationnels pourrait être intéressante pour enseigner la pratique des entretiens motivationnels. Ceux-ci sont aujourd'hui largement utilisés en addictologie [27]. Néanmoins, ils nécessitent un apprentissage et un état d'esprit qui passe de plus en plus par des techniques de simulation qui pourraient être aidées par les agents conversationnels. Le thérapeute simule des scénarios de patients et l'IA l'aide à réfléchir sur les réponses à apporter (*encadré 1*). Il est possible de customiser les IA, notamment en incluant le livre de Miller et Rollnick [27] et les aptitudes nécessaires pour les entretiens motivationnels et autres thérapies : l'empathie, l'acceptation des patients tels qu'ils sont, des buts partagés, l'espérance et un regard positif, l'évocation de la perspective des patients, de l'information et des conseils personnalisés et de l'authenticité.

J'ai 42 ans, j'ai un gros problème d'alcool. Je suis au fond du gouffre, je n'ai plus envie de rien, je ne mange plus, je dors mal, ma femme est partie... Je suis perdu, qu'est-ce que je peux faire ?

Réponse ChatGPT : Je suis vraiment désolé d'entendre ce que tu traverses. Reconnaître que tu as un problème et chercher de l'aide est une première étape importante, et cela demande du courage. Tu n'es pas seul dans cette situation, et il y a des moyens de t'en sortir, même si tout te semble sombre en ce moment.

Voici quelques étapes pour essayer de retrouver un chemin vers le mieux-être :

- Demande de l'aide professionnelle immédiatement.
- Rejoins un groupe de soutien.
- Parle à un proche ou à un ami de confiance.
- Établis des routines, même petites.
- Considère une désintoxication encadrée.
- Sois bienveillant envers toi-même.
- Prends soin de ta santé physique.
- Pense à un traitement médicamenteux si nécessaire.
- Concentre-toi sur une journée à la fois.
- N'abandonne pas.

Intérêt de l'IA dans les tâches médico-administratives

Sur le plan administratif, les agents conversationnels pourraient aider à gagner du temps pour [19,20] :

- Résumer les dossiers (synthèse de l'information).
- Assister l'écriture et corriger des textes (grammaire, syntaxe et style).
- Aider à l'écriture des dossiers et des comptes-rendus.
- Faciliter le codage CIM-10 et CIM-11.
- Réaliser des traductions.
- Automatiser des tâches répétitives, par exemple l'établissement de plannings ou la rédaction de certificats.

Intérêt de l'IA dans la recherche en addictologie

De nombreuses études telles que les études de cohorte longitudinales européennes IMAGEN et nord-américaine ABCD utilisent des techniques de *machine learning* depuis longtemps pour analyser de très grands nombres de données socio-démographiques, génétiques, cliniques, cognitives, de traits de personnalité et d'imagerie cérébrale afin de mettre en évidence les facteurs de risque prédictifs de troubles psychiatriques

ultérieurs, dont les consommations excessives de substances et les addictions.

Par exemple, une des premières études IMAGEN sur 692 adolescents suivis de façon longitudinale a montré qu'à l'âge de 14 ans, les éléments prédictifs de *binge drinking* ultérieur étaient une densité de substance grise plus faible et une activité plus importante dans le gyrus frontal supérieur que chez les sujets témoins [28]. Une autre étude IMAGEN a retrouvé que les facteurs prédictifs les plus pertinents de trouble de l'usage d'alcool chez les adolescents étaient les troubles externalisés et des traits de personnalité tels que la recherche de sensations (*sensation seeking*) [29]. Dans la même cohorte, la consommation de cannabis s'est révélée altérer le neurodéveloppement normal des adolescents [30]. Récemment, des techniques de *machine learning* ont retrouvé dans une sous-population IMAGEN que l'impulsivité, les symptômes de TDAH, des sentiments désespoir et le neuroticisme étaient des facteurs de risque prédictifs partagés entre troubles de l'usage d'alcool, dépression et troubles du comportement alimentaires [31].

Limites de l'IA

Les limites et les risques techniques, éthiques et de confidentialité de l'IA, notamment des agents conversationnels sont maintenant bien connus. Sur les prises en charge, certaines études ont retrouvé que l'IA n'était pas toujours pertinente, notamment aux urgences [32]. En conséquence, l'IA ne peut pas remplacer le raisonnement humain, surtout dans l'interprétation clinique, ni les relations soignants-soignés. En outre, certains auteurs ont souligné que l'utilisation trop fréquente de l'IA générative pouvait conduire à une dégradation de la qualité des dossiers médicaux, à une perte de l'esprit critique, à une perte de compétences et à un appauvrissement de la pensée, notamment chez les « cyborgs », sujets qui utilisent l'IA en permanence [33, 34].

Erreurs et hallucinations

Le risque d'erreurs ou d'« hallucinations » (réponses inventées par l'IA, sans rapport avec la réalité), d'informations fausses ou dépassées sont assez répandues avec les agents conversationnels. Le risque peut être amplifié par l'apprentissage machine [35]. Les hallucinations peuvent aller jusqu'à donner de fausses interactions médicamenteuses, de faux numéros d'urgences ou inventer des études qui vont généralement dans le sens de ceux qui font les requêtes [22, 36]. Ces erreurs peuvent être sous-estimées du fait que les humains ont tendance à croire ce que les machines produisent (illusion d'exactitude). Ces erreurs sont en fait liées aux calculs mathématiques des LLM. Il existe enfin des cas d'utilisation de l'IA à but de désinformation [36, 37].

Limites éthiques

La position dominante des géants du numérique, les Gafam (Google, Apple, Facebook, Amazon, Microsoft) et les BATX chinois (Baidu, Alibaba, Tencent, Xiaomi) peut être problématique, notamment concernant la confidentialité des données, tout particulièrement concernant des patients souffrant d'addiction(s), nécessitant des procédures d'anonymisation et de consentement. Des cas de revente de données à des tiers ont été observés, notamment à des compagnies d'assurances, y compris sans le consentement des patients. Le profilage publicitaire est en principe interdit sur chatGPT, mais rien ne garantit que cela le reste à l'avenir.

À cet égard, l'American Psychiatric Association (APA) et l'American Medical Association (AMA) ont recommandé en 2023 de ne jamais partager les données des patients dans un chatbot type ChatGPT, notamment en raison de l'absence totale de confidentialité et de leur stockage par des tiers. Pour surmonter cet obstacle, Google a mis au point un modèle d'IA (*open MedGemma AI*) pouvant être intégrée par des institutions telles que les hôpitaux sans que leurs données ne puissent en sortir.

Principaux biais de l'IA

La grande majorité des IA conversationnelles ont des algorithmes prédefinis, source d'un certain nombre de biais [22].

Biais liés aux données d'entraînement

– Biais de représentativité : certaines catégories (genre, âge, origine, situation géographique) sont sur-ou sous-représentées. Par exemple, une étude a retrouvé que Gemma IA développée par Google reposait en partie sur des biais sexistes et minimisait les problèmes de santé des femmes [38].

– Biais historique : reproduction ou amplification de discriminations existantes dans les données obsolètes.

– Biais de sélection : les données proviennent de sous-groupes non représentatifs de la réalité.

Biais liés à la formulation des questions ou interactions

– Biais d'acquiescement (*acquiescence bias*) : tendance d'agents conversationnels (et d'humains) à répondre de manière à aller dans le sens de l'interlocuteur, en particulier en acquiesçant à ses propositions, même si elles sont discutables ou incorrectes.

– Biais de formulation (*framing effect*) : la manière dont la question est posée influence la réponse.

– Biais d'ordre : l'ordre des questions ou des propositions influence la réponse.

– Biais conversationnel (ou social) : l'IA cherche à maintenir une conversation fluide, polie et engageante, ce qui peut l'amener à sur-adapter ses réponses aux attentes perçues de l'utilisateur, parfois au détriment de la rigueur ou de la neutralité.

Biais d'interprétation

– Biais de confirmation (*confirmation bias*) : tendance à privilégier les données qui confirment une hypothèse, les croyances ou les attentes de l'utilisateur, plutôt que de les remettre en question ou de proposer une perspective critique.

– Biais d'ancrage : la première information donnée influence toutes les suivantes.

– Biais d'optimisme ou pessimisme : surestimation ou sous-estimation de la probabilité d'un résultat favorable.

Biais algorithmiques spécifiques

– Biais de sur-apprentissage (*overfitting*) : le modèle colle trop aux données d'entraînement, perdant en généralisation.

– Biais de simplification excessive : le modèle privilie une solution facile mais inexacte.

– Biais d'échantillonnage adaptatif : dans les systèmes interactifs, les réponses influencent les futures données collectées.

Ces biais expliquent pourquoi il ne faut pas se fier à tout ce que produit l'IA et certainement pas lui faire confiance aveuglément. À cet égard, l'American Psychiatric Association a recommandé de rester prudent vis-à-vis de l'IA [19]. Il est également nécessaire de mentionner systématiquement l'utilisation de l'IA lorsqu'elle est utilisée et d'obtenir le consentement des patients. L'IA est avant tout un outil utile pour l'analyse des données qu'il faut savoir maîtriser, contextualiser et critiquer [33].

L'empathie montrée par beaucoup d'IA génératives, favorisée par les biais d'acquiescement et de confirmation, explique peut-être la satisfaction des patients quand ils les utilisent. En conséquence, un certain nombre de personnes utilisent les agents conversationnels comme des conseillers, des confident, des amis pour obtenir des conseils pratiques, à des fins de développement personnel ou pour raconter leurs vies intimes. Malheureusement, certains conseils peuvent être inadaptés, erronés, parfois malveillants.

Certaines études ont certes retrouvé que l'IA générative montrait plus d'empathie que les professionnels, mais ces résultats sont très discutés, d'autres études retrouvant des résultats inverses, avec notamment un côté superficiel, stéréotypé, en quelque sorte « robot » [39]. À cet égard, le président de l'AMA a souligné en 2024 que l'intervention humaine devait rester prédominante et que l'IA ne devait certainement pas devenir un prétexte pour se détourner des patients. Des IA prenant mieux en compte la complexité de la reconnaissance des émotions pourraient améliorer les thérapies [40]. Dans tous les cas, il est nécessaire que les thérapeutes informent les patients du fonctionnement des IA, en particulier de la nécessité de ne pas partager leurs données personnelles ni leurs résultats médicaux avec l'IA. Les patients doivent également savoir que les chatbots type chatGPT peuvent faire des erreurs, ne sont

pas toujours actualisés, ne répondent pas toujours aux données actuelles de la science, confortent les opinions plus qu'ils ne conseillent et ne sont liés à aucune clause de confidentialité. Pour minimiser le risque d'erreurs, il est important de formuler les requêtes ou instructions à l'IA (prompts) de la façon la plus exhaustive et précise possible avec tous les éléments utiles de contexte [33].

Conclusions : garder un esprit critique

Avec l'IA, il faut se garder des thuriféraires et des contempteurs. Après une période d'engouement, les limites de l'IA dans le domaine médical, y compris en addictologie, sont de mieux en mieux connues. L'IA est un outil puissant, notamment en termes de gain de temps, y compris pour soulager les cliniciens de certaines tâches administratives chronophages, à condition de garder un esprit critique vis-à-vis des réponses générées par l'IA. En addictologie, les études existantes sont surtout des études pilotes ou de faisabilité. Des études rigoureuses randomisées, contrôlées, sur de grands nombres de patients, restent à mener pour montrer l'efficacité de l'IA. À long terme, on peut imaginer un rôle important de l'IA dans la médecine de précision, notamment en intégrant les paramètres biologiques et ceux issus de l'imagerie cérébrale. L'efficacité de l'IA reste étroitement liée au degré de maîtrise de ceux qui l'utilisent : pour bien utiliser l'IA, il faut en comprendre les forces autant que les limites.

Liens d'intérêts l'auteur déclare ne pas avoir de lien d'intérêts en rapport avec cet article.

Références

1. Blease C, Worthen A, Torous J. Psychiatrists' experiences and opinions of generative artificial intelligence in mental healthcare: An online mixed methods survey. *Psychiatry Res* 2024 ; 333 : 115724.
2. Iqbal U, Tanweer A, Rahmanti AR, Greenfield D, Lee LT, Li YJ. Impact of large language model (ChatGPT) in healthcare: an umbrella review and evidence synthesis. *J Biomed Sci* 2025 ; 32 (1) : 45.
3. Tassinari DL, Pozzolo Pedro MO, Pozzolo Pedro M, Negrão AB, Abrantes do Amaral R, Malbergier A, et al. Artificial Intelligence-driven and technological innovations in the diagnosis and management of substance use disorders. *Int Rev Psychiatry* 2025 ; 37 (1) : 52-8.
4. Bonnell LN, Littenberg B, Wshah SR, Rose GL. A Machine Learning Approach to Identification of Unhealthy Drinking. *J Am Board Fam Med* 2020 ; 33 (3) : 397-406.
5. Palumbo SA, Adamson KM, Krishnamurthy S, Manoharan S, Beiler D, Seiwell A, et al. Assessment of probable opioid use disorder using electronic health record documentation. *JAMA Netw Open* 2020 ; 3 (9) : e2015909.
6. Kolding S, Lundin RM, Hansen L, Østergaard SD. Use of generative artificial intelligence (AI) in psychiatry and mental health care : a systematic review. *Acta Neuropsychiatr* 2024 ; 37 : e37.
7. Amer M, Gittins R, Millana AM, Scheibein F, Ferri M, Tofighi B, et al. Are Treatment Services Ready for the Use of Big Data Analytics and AI in Managing Opioid Use Disorder? *J Med Internet Res* 2025 ; 27 : e58723.
8. Gullapalli B, Luo Y, Rahman T, Garland EL. Opioid misuse detection from cognitive and physiological data with temporal fusion deep learning. *Drug Alcohol Depend* 2025 ; 274 : 112774.
9. Afshar M, Resnik F, Joyce C, Oguss M, Dligach D, Burnside ES, et al. Clinical implementation of AI-based screening for risk for opioid use disorder in hospitalized adults. *Nat Med* 2025 ; 31 (6) : 1863-72.
10. Beaulieu T, Knight R, Nolan S, Quick O, Ti L. Artificial intelligence interventions focused on opioid use disorders: A review of the gray literature. *Am J Drug Alcohol Abuse* 2021 ; 47 (1) : 26-42.
11. Dervaux A. Evaluation d'un patient alcoolodépendant : l'intelligence artificielle peut-elle avoir un intérêt ? *Presse Med* 2018 ; 47 : 638-9.
12. Mahbub M, Dams GM, Srinivasan S, Rizy C, Danciu I, Trafton J, et al. Decoding substance use disorder severity from clinical notes using a large language model. *Npj Ment Health Res* 2025 ; 4 (1) : 5.
13. Burkhardt G, Adorjan K, Kambeitz J, Kambeitz-Ilankovic L, Falkai P, Eyer F, et al. A machine learning approach to risk assessment for alcohol withdrawal syndrome. *Eur Neuropsychopharmacol* 2020 ; 35 : 61-70.
14. Doctor KP, McKeever C, Wu D, Phadnis A, Plawecki MH, Nurnberger Jr JI, et al. Deep learning diagnosis plus kinematic severity assessments of neurodivergent disorders. *Sci Rep* 2025 ; 15 (1) : 20269.
15. Walsh CG, Ripperger MA, Novak L, Reale C, Anders S, Spann A, et al. Risk model-guided clinical decision support for suicide screening : a randomized clinical trial. *JAMA Netw Open* 2025 ; 8 (1) : e2452371.
16. Berly A, Manaoui C, Dervaux A. L'intelligence artificielle peut-elle aider à estimer le risque de récidive dans les comportements violents ? *Médecine et Droit* 2020 ; 105-9.
17. Symons M, Feeney GFX, Gallagher MR, Young RM, Connor JP. Predicting alcohol dependence treatment outcomes: a prospective comparative study of clinical psychologists versus 'trained' machine learning models. *Addiction* 2020 ; 115 (11) : 2164-75.
18. Heinz MV, Price GD, Singh A, Bhattacharya S, Chen CH, Asyyed A, et al. A longitudinal observational study with ecological momentary assessment and deep learning to predict non-prescribed opioid use, treatment retention, and medication nonadherence among persons receiving medication treatment for opioid use disorder. *J Subst Use Addict Treat* 2025 ; 173 : 209685.
19. American Psychiatric Association. *The Basics of Augmented Intelligence: Some Factors Psychiatrists Need to Know Now*. 2023. <https://www.psychiatry.org/news-room/apa-blogs/the-basics-of-augmented-intelligence> (accédé le 16/08/2025).
20. Agence nationale de la performance sanitaire et médico-sociale (ANAP). *Déployer l'IA en toute confiance : stratégies et bonnes pratiques*. www.anap.fr (consulté le 21 novembre 2025).
21. Maddox TM, Embi P, Gerhart J, Goldsack J, Parikh RB, Sarich TC. Generative AI in Medicine – Evaluating Progress and Challenges. *N Engl J Med* 2025 ; 392 (24) : 2479-83.
22. Torous J, Linardon J, Goldberg SB, Sun S, Bell I, Nicholas J, et al. The evolving field of digital mental health: current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality. *World Psychiatry* 2025 ; 24 (2) : 156-74.
23. Giorgi S, Isman K, Liu T, Fried Z, Sedoc J, Curtis B. Evaluating generative AI responses to real-world drug-related questions. *Psychiatry Res* 2024 ; 339 : 116058.
24. Zhong W, Luo J, Zhang H. The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. *J Affect Disord* 2024 ; 356 : 459-69.
25. Heinz MV, Mackin DM, Trudeau BM, et al. Randomized Trial of a Generative AI Chatbot for Mental Health Treatment. *NEJM AI* 2025 ; 2 : Aloa2400802.
26. Suffoletto B, Clark DB, Lee C, Mason M, Schultz J, Szeto I, Walker D. Development and preliminary testing of a secure large language model-based chatbot for brief alcohol counseling in young adults. *Drug Alcohol Depend* 2025 ; 272 : 112697.
27. Miller WR, Rollnick S. *L'entretien motivationnel. Aider la personne à engager et réaliser le changement*. 3^e éd. Paris : InterEditions, 2024.
28. Whelan R, Watts R, Orr CA, Althoff RR, Artiges E, Banaschewski T, et al. IMAGEN Consortium. Neuropsychosocial profiles of current and future adolescent alcohol misusers. *Nature* 2014 ; 512 (7513) : 185-9.
29. Afzali MH, Sunderland M, Stewart S, Masse B, Seguin J, Newton N, et al. Machine-learning prediction of adolescent alcohol use : a cross-study, cross-cultural validation. *Addiction* 2019 ; 114 (4) : 662-71.
30. Albaugh MD, Ottino-Gonzalez J, Sidwell A, Lepage C, Juliano A, Owens MM, et al. IMAGEN Consortium. Association of Cannabis Use During Adolescence With Neurodevelopment. *JAMA Psychiatry* 2021 ; 78 (9) : 1-11.
31. Zhang Z, Robinson L, Whelan R, Jollans L, Wang Z, Nees F, et al. Machine learning models for diagnosis and risk prediction in eating disorders, depression, and alcohol use disorder. *J Affect Disord* 2025 ; 379 : 889-99.
32. Williams CYK, Miao BY, Kornblith AE, Butte AJ. Evaluating the use of large language models to provide clinical recommendations in the Emergency Department. *Nat Commun* 2024 ; 15 (1) : 8236.

33. Abdulnour RE, Gin B, Boscardin CK. Educational Strategies for Clinical Supervision of Artificial Intelligence Use. *N Engl J Med* 2025 ; 393 (8) : 786-97.
34. McCoy LG, Manrai AK, Rodman A. Large language models and the degradation of the medical record. *N Engl J Med* 2024 ; 391 (17) : 1561-4.
35. Omar M, Sorin V, Collins JD, Reich D, Freeman R, Gavin N, *et al.* Multi-model assurance analysis showing large language models are highly vulnerable to adversarial hallucination attacks during clinical decision support. *Commun Med (Lond)* 2025 ; 5 (1) : 330.
36. Monteith S, Glenn T, Geddes JR, Whybrow PC, Achtyes E, Bauer M. Artificial intelligence and increasing misinformation. *Br J Psychiatry* 2024 ; 224 (2) : 33-5.
37. Haupt CE, Marks M. FTC Regulation of AI-Generated Medical Disinformation. *JAMA* 2024 ; 332 (23) : 1975-6.
38. Rickman S. Evaluating gender bias in large language models in long-term care. *BMC Med Inform Decis Mak* 2025 ; 25 (1) : 274.
39. Lee HS, Wright C, Ferranto J, Buttmer J, Palmer CE, Welchman A, *et al.* Artificial intelligence conversational agents in mental health: Patients see potential, but prefer humans in the loop. *Front Psychiatry* 2025 ; 15 : 1505024.
40. Striegl J, Richter JW, Grossmann L, Bråstad B, Gotthardt M, Rück C, *et al.* Deep learning-based dimensional emotion recognition for conversational agent-based cognitive behavioral therapy. *PeerJ Comput Sci* 2024 ; 10 : e2104.